Proceedings of the LFG 01 Conference

University of Hong Kong, Hong Kong

Editors: Miriam Butt and Tracy Holloway King

2001

CSLI Publications

ISSN 1098-6782

Editors' Note

This year's conference was held in Hong Kong at the University of Hong. We would like to thank Adams Bodomo and the local organizing committee for putting together a superb conference.

The program committee for the LFG'01 at Hong Kong were Rachel Nordlinger and Chris Manning. We would like to thank them for putting together the program that gave rise to this collection of papers. We would also like to thank the executive committee and the reviewers, without whom the conference would not have been possible.

The table of contents lists all the papers presented at the conference and some that were accepted but could not be presented. Some papers were not submitted to the proceedings. For these papers, we suggest contacting the authors directly via the e-mail addresses also provided as part of the table of contents.

Table of Contents

Iofu and Spreading Architecture in LFG

Avery Andrews
Australian National University

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)\

2001

CSLI Publications
http://csli-publications.stanford.edu

Iofu and Spreading Architecture in LFG*

Avery D Andrews
School of Language Studies, The Faculties, ANU

1 Introduction

Andrews and Manning (1999) proposed a revision of LFG architecture whereby the standard conception of f-structure is replaced by a more general feature structure, which includes c-structural and argument-structural information in addition to the standard content of f-structure. Results produced in standard LFG by the many-to-one correspondence between c-structure and f-structure are then accounted for by use of a 'spreading architecture', in which grammatical attributes are systematically shared across multiple levels of the c-structure. This revision of LFG architecture supports straightforward analyses of various kinds of complex predicate structures whose properties are quite problematic under the standard version of LFG, and indeed other explicit linguistic formalisms.

However the proposal has a compatibility problem with another recent development in LFG, the use of 'inside-out functional uncertainty' (iofu) to describe phenomena such as anaphoric dependence (Dalrymple 1993) and case-marking (Andrews 1996, Nordlinger 1998). In this paper I will present the problem and a proposed solution.

2 Spreading Architecture vs. Iofu

For a simple example of iofu, consider the application of the constraint (1a) to the f-structure (1b)

(1) a. ((SUBJ ↑) OBJ)
c.f. (Nordlinger 1998)

b.
$$g{:}\begin{bmatrix} \text{SUBJ} & f{:}\begin{bmatrix}\text{CASE} & \text{ERG}\end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{OBJ} & h{:}\begin{bmatrix}\text{CASE} & \text{ABS}\end{bmatrix} \end{bmatrix}$$

Suppose that in the process of instantiation, the ↑ in (a) is instantiated to the f-structure f in (b). Now the position of the arrow at the end rather than the beginning of the expression (SUBJ ↑) indicates that it is an iofu designator; therefore it designates some f-structure which has f as its SUBJ-value; g in (b) satisfies this condition, and so is a possible value for this designator.[1] So now the whole expression (b) says that g

[0]I would like to thank Chris Manning, Jane Simpson and David Nash for comments and discussion.

[1]Because f-structures allow a single f-structure to be the value of (possibly different) attributes in more than one f-structure, an 'inside out' designator has multiple possible values, that is, it is an instance of functional uncertainty.

has an OBJ-value, and therefore the structure (b) satisfies the constraint (a) (as long as $\uparrow$ designates f, but would not if it were to designate h).

To see the incompatibility with spreading architecture, consider the treatment of simple NP-structure and case-marking under the two approaches. In standard LFG, an NP would have a c-structure with nodes such as NP, N, etc (the labels being specified by an independent function assigning c-structure categories to nodes), and an f-structure related to the c-structure by the many-to-one correspondence ϕ (from c-structure nodes to f-structures). This correspondence is determined by annotations on the c-structure nodes (or in more recent formulations by general principles regulating ϕ). Below is the structure for a simple NP in the Australian language Warlpiri (Simpson 1991), with the ϕ correspondence represented by the shared index f:

(2)

$$\begin{array}{c} \mathrm{NP}_f \\ | \\ \mathrm{N}_f \\ \uparrow=\downarrow \\ | \\ \textit{kurdu-ngku} \end{array} \qquad f{:}\begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{PRED} & \text{‘Child’} \end{bmatrix}$$

By contrast, with spreading architecture, the (equivalents of) the c-structure nodes are connected by a single function to feature-structures which specify all featural information relevant to the node. Since this information will typically be different for mother and daughter nodes, the NP will have a different feature-structure correspondent than the N, leading to an overall structure like this:

(3)

$$\begin{array}{c} \bullet_g \\ | \\ \bullet_f \\ =\{\alpha,\kappa,\rho,\mu\} \\ | \\ \textit{kurdu-ngku} \end{array} \qquad \begin{array}{l} g{:}\begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{CAT} & \text{N} \\ \text{PRED} & \text{‘Child’} \\ \text{BAR} & 1 \end{bmatrix} \\[2ex] f{:}\begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{CAT} & \text{N} \\ \text{PRED} & \text{‘Child’} \\ \text{BAR} & 0 \end{bmatrix} \end{array}$$

Rather than having two c-structure nodes associated with a single f-structure, each c-structure node (represented by a dot with a subscript) is associated with a different feature-structure, but these feature-structures share large numbers of attributes, indeed in this example, all attributes except the bar feature. This sharing is controlled by annotations which refer to classes of attributes called ‘projections’. The annotation on the c-structure in (3) says that attributes belonging to the α (argument-structure and semantic), κ (lexical c-structure category), ρ (grammatical function) and μ (morphosyntactic feature) projections should be shared between the mother and the daughter.[2]

The issue with iofu is that it is not clear how a constraint such as (1a), which is supposed to be co-introduced with a morphological element such as a case-value, can

[2] This annotation is an abbreviation for statements that can be formulated using the set-theoretical concept of restriction, see (Andrews and Manning 1999) for details.

be satisfied in a structure like (3). Following the approach to case-marking proposed in (Nordlinger 1998), a constraint such as (1a) would be introduced together with the ergative marker on the node with feature-structure g in (3). But in order to be satisfied it needs to apply to feature-structure f, since this is what would be the value of SUBJ in a containing structure. But f and g never get equated, so the structures remain different, and the constraint cannot be satisifed. The same problem will apply to other iofu constraints associated with morphological material: without some way to get the iofu constraints to apply to a higher structure, these constraints can't be satisfied, and therefore cannot do useful work in a grammar.

3 A Proposal

The basic idea of my proposal is to have the iofu constraints 'hitch a ride' on shared attributes. Consider the question of what would be designated by an expression such as:

(4) (CASE (↑ CASE))

with ↑ desigating f in (3b). (↑ CASE) will designate the CASE-value of f (an ordinary 'outside-in' designator), which is ERG. So the whole expression will be a functionally uncertain inside-out designator referring to some feature-structure which has ERG as its CASE-value. There are two candidates, f and g, the latter being the one we want. Thus the sharing of the CASE-attribute between the upper and lower node provides a way to get iofu constraints to work. There is however a technical issue we need to deal with, and also a significantly different alternative formulation.

The technical issue is that in the standard formalization of LFG (Kaplan and Bresnan 1982), feature-values are thought of as set-theoretic objects for which there is no difference between types and tokens: all ergative case-values in a structure would be represented by the single set-theoretical object ERG. This interpretation is clearly not compatible with the present proposal: we need to think of grammatical feature values as tokens which may or may not be of the same type (ERG vs. ACC, etc.). The situation is similar to that for the PRED-feature, except that different PRED-tokens cannot merge regardless of whether or not they are of the same type, whereas for grammatical feature values we want tokens to be able to merge if and only if they are of the same type.[3] On this account, the feature-structures designated by (4) would be for which there is an explicit identity between their CASE-value and that of f.

Another possible formulation that we need to consider depends on a possible alternative conception of the projections. In (Andrews and Manning 1999), projections are conceived of as mere sets of attributes, having no direct representation in sentence structures, which express only their effects, in the form of shared values of attributes beetween different parts of the structure. However there is another possible formulation (described in unpublished notes by John Maxwell), that projections might be distinct components of the f-structure. So the μ projection for example might be a distinct subregion of a feature-structure where the morphosyntactic features reside. This

[3] However see (Andrews 1990) for suggestions that things are not so clear-cut.

would make it possible to equate projections (parts of feature-structures) without saying anything about the specific attributes that reside on them. We won't pursue the consequences of this kind of choice here, but will remain with the first conception.

(3) is a rather cumbersome formulation, to simplify it I propose an abbreviatory notation whereby the attribute appearing on both sides of the $\uparrow$ is subscripted to the arrow. The constraint (1a) can then be reforumulated as either (a) or (b) below:

(5) a. $((\text{SUBJ} \uparrow_{\text{CASE}})\ \text{OBJ})$

b. $((\text{SUBJ} \uparrow_{\mu})\ \text{OBJ})$

The interpretation of formulation (a) is straightforward under either conception of projections, whereas if we want interpret projections strictly as sets, the (b) formulation could be understood as functionally uncertain with respect to the 'outside in' portion of the designator as well as the inside out one: find some attribute on the μ projection (functionally uncertain because there may be more than one such attribute, and find an feature-structure with a shared value for that attribute).

So we have a proposal that works, but to get a real result we need to find some advantage that the modified theory has over its predecessor. I claim that this can be found in the possibility of accounting for some of the different systems for distributing case-markers across the members of NPs in Australian languages.

4 Case Distribution

Dench and Evans (1988:3-6) observe the following major patterns of case-distribution in Australian Languages.[4]

(6) a. mark on all members of the NP

b. mark on the last member(s) of the NP

c. mark on head of the NP

d. mark on any member of the NP

(Dench and Evans 1988:3-6)

(a) and (c) are non-problematic for LFG (but see (Andrews 1996) for discussion of the role of the Morphological Blocking Condition for (a)). (d) is likewise probably not a problem, although available descriptions of the languages having it do not give an entirely consistent picture of how this distribution pattern actually works.[5]. In many languages, such as Turkish, (b) co-occurs with (c), since the head is final in the NP. But in certain Australian languages, the case-marker is final in the NP (at least normally), but the head isn't. Rather at least some modifiers are able to appear after the head (and normally do so), but before the case-marker. Examples from Central Australia are afforded by Arrernte, Pitjantjatjara and Warlpiri:

[4] Some of the Nyulnyulan languages, such as Bardi, appear to have a significant preference for a fifth pattern, case-marking (by means of a suffix) on the first member of the np (Gedda Aklif, Claire Bowern, p.c.)

[5] (Stokes 1982) for Nyigina, and (McGregor 1990) for Gooniyandi, both Nyulnyulan

(7) Arrernte (Wilkins 1989):

a. *Kngwelye-le ker arlkwe-ke*
dog-ERG meat eat-PAST
The dog ate (the) meat

b. *Kngwelye yanhe-le ker arlkwe-ke*
dog that-ERG meat eat-PAST
That dog ate the meat

(8) Pitjantjatjara (Bowe 1990:29-31):

a. *Minyma-ngku tjitji nya-ngu*
woman-ERG child saw-PAST
The woman saw the child

b. *Wati nyanga ninti pul̲ka-ngku muttuka palya-nu*
man that clever very-ERG car fix-PAST
That very clever man fixed the car

c. *Tjitji pul̲ka-ngku mal̲anypa nya-ngu*
Child big-ERG younger sibling see-PAST
That big child saw his younger sibling

d. *Wati nyanga-ngku mutuka palya-lpai*
man this-ERG car fix-HABIT
This man fixes cars

(9) Warlpiri (Nash 1986:170)

a. *karnta(-ngku) kurdu-ngku wita-ngku*
female(-ERG) child-ERG small-ERG

b. *karnta(*-ngku) kurdu wita-ngku*
female(-ERG) child small-ERG

small female child (Nash 1986:170)

In Warlpiri, unlike the other two languages, case-markers are not restricted to occurring finally; the operative constraint appears to be that a final segment of the nominals in an NP must bear the case-marker.

Simpson (1991:131-133) proposed a c-structure category-based account of case distribution in Warlpiri, whereby nominal words could be introduced as instances of either N or N^{-1}, the latter being a nominal word incapable of bearing a case-marker. A simplified version of Simpson's proposal is:

(10) $\overline{\mathrm{N}} \quad \rightarrow \quad \mathrm{N}^{-1}* \;\; \mathrm{N}* \;\; \mathrm{N}$

This could be modified to apply to Arrernte and Pitjantjatjara by removing the N* term. Simpson's proposal works for a wide range of NPs in Warlpiri, but there are expressions in all three languages for which it fails. The problem is it links the potential absence of case-marking to the distribution of a lexical category. It therefore cannot explain instances in which a phrasal category is non-final and consequently excluded from the domain of case-marking, whether optional or obligatory.

A Warlpiri example is the following from (Simpson 1983:331):

(11) a. *Maliki kurdu yali-kirlangu-rlu ka wajilipi-nyi mutukayi*
dog child that.rem-POSS-ERG PRES chase-NPAST car(ABS)

b. *Kurdu yali-kirlangu maliki-rli ka wajilipi-nyi mutukayi*
child that.rem-POSS dog-ERG PRES chase-NPAST car(ABS)

The dog of that child is chasing the car

In these examples we have an NP containing another NP embedded as a possessive modifier. (a) is consistent with Simpson's approach, since the nonfinal element in each NP is a single word. But in (b), the first element of the NP is the possessor, expressed as a two-word phrase.[6] The final word of the possessor phrase bears a possessive case-marker applying to the whole possessor phrase, but no instance of the ergative. Although this appears to be an instance of the pattern that (10) is trying to describe, (10) doesn't account for it, since (10) constrains the distribution of lexical rather than phrasal categories.

Similar examples are found in Arrernte and Pitjantjatjara:

(12) Pitjantjatjara (Bowe 1990:40-41)

a. *Wati wara-ku minyma ninit pulka-ngku numa palya-nu*
man tall-GEN woman clever very-ERG bread make-PAST
That tall man's very clever wife made the traditional bread

b. *Wati panya minyma nyanga-ku-ngku malu waka-nu*
man ANAPH woman this-GEN-ERG roo spear-PAST
that husband of this woman speared the kangaroo

(13) Mparntwe Arrernte (Wilkins 1989:156)

a. *Kngwelye kweke artwe kngerre nturre-kenhe-le re-nhe utnhe-ke*
dog small man big very-GEN-ERG he-ACC bite-PAST
The small dog of the very big man bit him

b. *Artwe kngerre nthurre-kenhe knwelye kweke-le*
man large very-GEN dog small-ERG
the small dog of the very large man

[6]To make this point, it is essential to use a phrasal possessor. The reason is that possessive case-markers attached to single worlds can be construed as derivational affixes creating derived adjectives, which can fail to bear affixes as a regular consequence of the application of rule (10).

What we can do to accomodate these kinds of examples is to tweak Simpson's original rule so that it introduces phrasal as well as lexical categories, and uses spreading annotations rather than morphosyntactic categories. A first version might be the following, where the material in brackets is absent in Arrernte and Pitjantjatjara:

(14) $\overline{\mathrm{N}} \quad \rightarrow \quad \mathrm{N}^* \ <\mathrm{N}_{=\{\mu\}}> \ \ \mathrm{N}_{=\{\mu\}}$

This formulation will confine the spreading of morphosyntactic features to the final nominal of the NP, or a final segment of the nominals of the NP. Nominals that don't get the $= \{\mu\}$ annotation will not be able bear a case, because the iofu constraints associated with the introduction of the case-marker won't be able to be satisfied, and when a nominal gets the annotation, the morphological blocking principle will require the appropriate case-marker, if any to appear.

The data involving phrasal modifiers of NPs is however still not accounted for, because (14) only introduces lexical categories, not phrasal ones. What we need is to allow the non-head constituents of an NP to be introduced either as N or as $\overline{\mathrm{N}}$. It would be possible to invent notations to allow this to be done with a formulation along the general lines of (14), but it would be better to break the formulation down into general principles that have the capacity apply across different kinds of phrase-structual configurations.

To begin with we need to make some decisions about the representation of NPs with 'semantic' case-markers, not only adnominal ones such as proprietive and genitive, but also adverbial ones such as allative and locative (the latter can also function adnominally). Simpson (1991) provides an argument, recapitulated in (Andrews 1996:7), that semantic case-markers introduce an extra level of f-structure, needed to host the PRED-value of the semantic case, which is distinct from the PRED-value of the nominal to which the case-marker is attached. The argument can probably be circumvented in modern LFG by the use of glue-logic semantics, but we will accept the conclusion for the sake of putting forth some definite structural proposals. To make the outcome of the proposed principles easier to grasp, I will treat semantically case-marked NPs as instances of a category KP (for Case Phrase), which will in Warlpiri expand directly to an NP (ordinary NP), bearing the grammatical function OBL. In an NP such as *pirli-ngka* 'hill/rock-loc', the resulting structure will then look like this:

(15)

$\bullet_h$
|
$\bullet_f$
(↑ OBL))=↓
|
$\bullet_g$
$= \{\alpha, \kappa, \rho, \mu\}$
|
pirli-ngka

$$h{:}\begin{bmatrix} \text{PRED} & \text{`Loc(OBL)'} \\ \text{CAT} & \text{K} \\ \text{BAR} & 1 \\ \text{OBL} & f{:}\begin{bmatrix} \text{CASE} & \text{LOC} \\ \text{CAT} & \text{N} \\ \text{PRED} & \text{`Rock'} \\ \text{BAR} & 1 \end{bmatrix} \end{bmatrix}$$

$$g{:}\begin{bmatrix} \text{CASE} & \text{LOC} \\ \text{CAT} & \text{N} \\ \text{PRED} & \text{`Rock'} \\ \text{BAR} & 0 \end{bmatrix}$$

Note that the feature-structure of the lower N is not itself bear any function in the feature-structures of the nodes that dominate that N, but the values of many attributes are shared between them.

The PRED feature can be placed into its higher structural position by an annotation such as (16), introduced by the case-marker:

(16) ((OBL $\uparrow_\mu$) PRED)= 'Poss(OBL)'
(to implement case-stacking, Nordlinger's 'functional composition' applies)

Structures with significant resemblances to this proposal for semantically case-marked NPs include Enclitic Phrases in Navajo (Kaufman 1974), the structure proposed for 'Quirky' case-marked NPs in Icelandic by Andrews (1982), and Locative Phrases in Bantu languages such as Chichewa (Bresnan and Mchombo 1995).

On this account, an NP such as the ergative in (11b) would have a c-structure like this (c-structure categories written into the tree for convenience, and many attribute-sharing annotations omitted):

(17)

NP
KP (↑ADJUNCT) = ↓ — N *maliki-rli*
KP: NP (↑OBJ) = ↓
NP: N *kurdu* — Dem *yali-kirlangu*

The following principles will produce the kinds of structures we want for NPs:

(18) a. NP can dominate N or KP

b. N under NP can be annotated $= \{\rho, \alpha\}$ or (↑ADJUNCT) = ↓

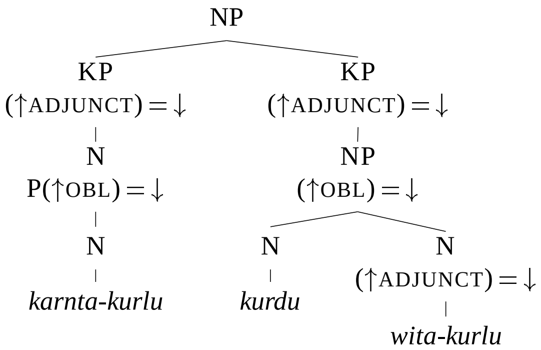
c. KP under NP must be annotated $(\uparrow\text{ADJUNCT}) = \downarrow$

d. the final element (in Warlpiri, elements) must share its μ projection with the mother; the others may not.

Importantly, if the KP and N in (17) are reordered, the lexical entry for the semantic case-marker in (16) will interact with Nordlinger's principle of functional composition so that the ergative marker will appear on the end of the possessor NP, after the genitive marker, as in (11a), and nonetheless apply to the topmost NP.

Another example of how these principles interact is provided by (19):

(19)



[Somebody] with a small female child

What is interesting here is the apparent exception to the generalization that it must be a final segment of the nominals of an NP that is cased-marked. The rules allow this kind of example to work differently than (9) because this example involves an attributive KP, and Warlpiri allows discontinuity of NP and KP. The whole nominal has an unexpressed head and a discontinuous adjunct KP, with the case-marker appearing on the final member of each KP-piece.

The proposed reformulation of iofu-based rules thus provides some empirical coverage that has previously been missing from LFG theories of case.

5 Further Issues

There are a number of further issues to be dealt with. One is the possibility that the supposed case-markers might actually be postpositions, or phrasal clitics of some sort, rather than affixes. On such an analyses, we would treat them as occupying final (head) position in KP or an equivalent node, the syntax per se having no responsibility for their apparent relationship to the nominals they appear on. In Warlpiri, such an analysis is challenged by the possibility of case-distribution; it faces other problems in addition.

In Warlpiri, the ergative and locative case-markers begin with *-ngk* after disallabic stems, *-rl* otherwise, a kind of allomorphy typically associated with affixes rather than clitics. Furthermore in Arrernte and Pitjantjara, different classes of nominals show different case-marking patterns. In Pitjantjatjara, common noun and proper names inflect with an accusative-absolutive pattern while personal pronouns are nominative-accusative (using the otherwise ergative marker *lu* as a nominative marker). Furthermore proper names use *-nya* as an absolutive marker (which is typologically rather usual), while pronouns use it as an accusative marker:

(20) Pitjantjatjara CM Morphology:

	A	S	0
CN: man	wati-ngku	wati	wati
PN: Billy	Billy-lu	Billy-nya	Billy-nya
Pronoun: I	ngayu-lu	ngayu-lu	ngayu-nya

This kind of complexity is normal for affixes, but not for clitics.

Arrernte similarly has an ergative-absolutive system for non-pronominal N and also the first person singular pronoun, but nominative-accusative for the other pronouns. There is in addition an interesting phenomenon involving the third person pronoun *re*, which is also used as a definite article in NPs. While common nominals and the first person singular pronoun in Arrernte are inflected in accordance with an ergative-absolutive pattern, other pronouns, including *re* in its definitizing use, inflect in accordance with a nominative accusative pattern:

(21) a. *kngwelye re(*-rle) ker arlkwe-ke*
dog the(*-ERG) meat eat-PAST
The dog ate the meat

b. *artwe re*(-nhe) kngwelye-le uthwe-ke*
man the-ACC dog-ERG bite-PAST
the dog bit the man

NPs which end in *re* therefore partake in the nominative-accusative case-marking system rather than the ergative-absolutive that is normal for non-pronouns. This can be accounted for if case is being spelled out on the final word of the NP, as determined by its inflectional category, but is not consistent with the idea of case being a clitic. The case-marking cannot thus be seen as an particle which simply appears at the end of the NP, but rather has some kind of significant relationship with its final word.

A loose end in our treatment arises in Warliri. Simpson (1991:275) observes that an adnominal locative NP, unlike the genitive noted previously, must bear a case-marker appropriate for the whole NP:

(22) *pirli-ngka*(-rlu) wati-ngki nganpa luwa-rnu*
hill-LOC*(-ERG) man-ERG us shot-PAST
The man on the hill shot us

Our analysis however predicts that the ergative marker should be optional in (c). A possible solution to this might be that locative is also, and perhaps preferentially, an adverbial case-marker, which will agree in case with an argument to which it ascribes a property (Simpson 1991:2-4-208). Omission of the ergative from *pirli* in (22) would be consistent with an interpretation in which the locative was describing the absolutive argument *nganpa* 'us' in O function, rather than the ergative. If this was the preferred interpation under a version of bi-directional optimality theory, the phenomenon of (22) would follow.

A final issue to consider is the fact that in several of the examples from Arrernte and Pitjantjatjara in (8) and (7), case-markers are appearing after intensifiers of words functioning adjectivally. This sits uncomfortably with the arguments already given that case-markers are inflections. A possible resolution of this discrepancy is that perhaps they can be both. Suppose that the head (K) position in KP was optional, sharing μ with KP. It might then be possible for optimality theoretic principles such as *STRUC, Parse Case etc. to allow case to be manifested as an inflection on the final member of NP under some circumstances, and as a clitic introduce under K in others. However information-sharing of the kind proposed here would be essential in getting such an analysis to work.

Bibliography

Andrews, A. 1982. The representation of case in modern Icelandic. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 427–503. Cambridge, MA: The MIT Press.

Andrews, A. 1990. Unification and morphological blocking. *Natural Language and Linguistic Theory* 8(4):508–558.

Andrews, A. 1996. Semantic case-stacking and inside-out unification. *Australian Journal of Linguistics* 16(1):1–55.

Andrews, A. D., and C. D. Manning. 1999. *Complex Predicates and Information Spreading in LFG*. CSLI Publications.

Bowe, H. 1990. *Categories, Constituents and Constituent Order in Pitjantjatjara*. Routledge.

Bresnan, J. W., and S. Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language and Linguistic theory* 181–254.

Dalrymple, M. 1993. *The Syntax of Anaphoric Binding*. Stanford, CA: CSLI Publications. CSLI Lecture Notes, number 36.

Dench, A., and N. Evans. 1988. Multiple case-marking in australian languages. *Australian Journal of Linguistics* 8:1–48.

Kaplan, R. M., and J. W. Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In J. W. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 173–281. MIT Press.

Kaufman, E. 1974. Navajo spatial enclitics: A case for unbounded rightward movement. *Linguistic Inquiry* 507–534.

McGregor, W. 1990. *A Functional Grammar of Gooniyandi*. Amsterdam, Philadelphia: Benjamins.

Nash, D. G. 1986. *Topics in Warlpiri Grammar*. Garland.

Nordlinger, R. 1998. *Constructive case: Dependent-marking nonconfigurationality in Australia*. CSLI Publications.

Simpson, J. 1983. *Aspects of Warlpiri Morphology and Syntax*. PhD thesis, MIT, Cambridge MA.

Simpson, J. 1991. *Warlpiri Morphology and Syntax: A Lexicalist Approach*. Dordrecht: D. Reidel.

Stokes, B. 1982. *A description of Nyigina, a language of the West Kimberly, Western Australia*. PhD thesis, The Australian National University, Canberra ACT.

Wilkins, D. 1989. *Mparntwe Arrernte (Aranda): Studies in the Structure and Semantics of Grammar*. PhD thesis, The Australian National University, Canberra.

Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi[1]

JOAN BRESNAN, SHIPRA DINGARE, AND CHRISTOPHER D. MANNING

Stanford University

Proceedings of the LFG 01 Conference

The University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

[1]This work is based in part on research supported by the National Science Foundation under Grant No. BCS-9818077 and by a gift to the Stanford Symbolic Systems Program by Ric Weiland. This work omits some material in the LFG 01 conference presentation that will appear in a longer work in preparation by the authors.

ABSTRACT

The same categorical phenomena which are attributed to hard grammatical constraints in some languages continue to show up as statistical preferences in other languages, motivating a grammatical model that can account for soft constraints.

The effects of a hierarchy of person (1st, 2nd > 3rd) on grammar are categorical in some languages, most famously in languages with inverse systems, but also in languages with person restrictions on passivization. In Lummi, for example, the person of the subject argument cannot be lower than the person of a nonsubject argument. If this would happen in the active, passivization is obligatory; if it would happen in the passive, the active is obligatory (Jelinek and Demers 1983). These facts follow from the theory of harmonic alignment in OT: constraints favoring the harmonic association of prominent person (1st, 2nd) with prominent syntactic function (subject) are hypothesized to be present as subhierarchies of the grammars of all languages, but to vary in their effects across languages depending on their interactions with other constraints (Aissen 1999). There is a statistical reflection of these hierarchies in English. The same disharmonic person/argument associations which are avoided categorically in languages like Lummi by making passives either impossible or obligatory, are avoided in the SWITCHBOARD corpus of spoken English by either depressing or elevating the frequency of passives relative to actives. The English data can be grammatically analyzed within the stochastic OT framework (Boersma 1998, Boersma and Hayes 2001) in a way which provides a principled and unifying explanation for their relation to the crosslinguistic categorical person effects studied by Aissen (1999).

1 Categorical Effects of Person on Voice

The effects of the person hierarchy (1) on grammar are categorical in some languages, most famously in languages with inverse systems, but also in languages with person restrictions on passivization. In Lummi (Straits Salish, British Columbia), for example, the person of the subject argument cannot be lower than the person of a nonsubject argument. If this would happen in the active, passivization is obligatory; if it would happen in the passive, the active is obligatory (Jelinek and Demers 1983, 1994).[2]

(1) Person hierarchy:
1st, 2nd ≻ 3rd
(grammatical persons 'local' to speech act outrank others)

(2) Lummi examples:

a. *__ 'The man knows me/you'

b. x̣či-t-ŋ=sən/=sx^{w} ə cə swəyʔqəʔ
know-TR-PASS=1.SG.NOM/=2.SG.NOM by the man
'I am/you are known by the man'

c. x̣či-t-s cə swəyʔqəʔ cə swiʔqoʔəɬ
know-TR-3.TR.SUBJ the man the boy
'The man knows the boy'

d. x̣či-t-ŋ cə swiʔqoʔəɬ ə cə swəyʔqəʔ
know-TR-PASS the boy by the man
'The boy is known by the man'

e. x̣či-t=sən/=sx^{w} cə swəyʔqəʔ
know-TR=1.SG.NOM/=2.SG.NOM the man
'I/you know the man'

f. *__ 'The man is known by me/you'

The same holds in other, unrelated languages such as the Tanoan language Picurís (New Mexico) (Zaharlick 1982, Mithun 1999: 226–228) and the Southern Wakashan language

[2]The 'transitive' stem suffix -*t*, glossed TR, is one of a set that marks degree of volitionality of control of the action; the passive suffix -ŋ, glossed PASS, also marks middles (Jelinek and Demers 1994: 706). With local person arguments the active is obligatory. The Lummi pattern holds for bound pronouns; full nominal phrases designating speaker and hearer are formally 3rd person deictic expressions (Jelinek and Demers 1983: 173; 1994: 714).

Nootka (British Columbia) (Whistler 1985, Emanatian 1988). Although person-driven passives are sometimes viewed as inverses (cf. Klaiman 1991, Jacobs 1994, Forrest 1994, Jelinek and Demers 1983, 1994 on Salish), the syntactic contrasts in (3) have been drawn between person-driven passives and the Algonquian-type inverse exemplified by Plains Cree (Dahlstrom 1984), from Mithun (1999: 222–228):

(3)

Passive:	Inverse:
intransitive	transitive
patient Subject	patient Object
oblique case marking on agent	non-oblique agent
omissibility of indefinite agent	non-omissibility

On the basis of such contrasts, we accept with Mithun the evidence for the existence of person-driven passives. (Mithun 1999: 227 concludes of Picurís, “There is no question that these constructions are formally passive.”)

2 A Theory of Passivization in Optimality Theory

From a classical generative point of view, it is difficult to see why person and voice should interact: after all, the person of arguments has nothing to do with verbal lexical semantic structure or case frames or the syntactic classification of lexical argument roles in LFG’s lexical mapping theory (Bresnan 2001: ch. 14 and references), which have usually been taken to drive passivization. But from a different perspective, such interactions are not surprising.

It is well known that passives have properties of syntactically ‘marked’ constructions (Greenberg 1966, Trask 1979):

(i) Typological distribution: There are many languages without passives.

(ii) Language-internal distribution: Where it occurs, the passive is often more restricted than the active. For example, many languages restrict the passive agent (it may not appear, or may appear only in certain persons); others have a morphologically defective passive paradigm (lacking certain tenses, etc); only subclasses of active transitive verbs may passivize.

(iii) Morphological marking: Passivization is morphologically marked (Haspelmath 1990).

But why should this be? An historical explanation is that actives are basic (unmarked) verb types; passives arise from originally deverbal constructions such as stative adjectives or nominals by a historical process of verbalisation (Trask 1979, Estival and Myhill 1988, Haspelmath 1990, Garrett 1990). But the historical explanation does not answer the question: *Why* are actives the basic/unmarked verb types, rather than passives?

The intuition shared by many linguists and adopted by Aissen (1999) is that agents make better subjects than patients do. Semantically 'active' (proto-agent) arguments harmonically align with the most prominent syntactic argument positions; semantically 'inactive' (proto-patient) arguments harmonically align with the least prominent syntactic positions. The ultimate reasons for this alignment lie in the pragmatics of discourse and the cognitive biases of humans (see Givón 1979, 1983, Kuno and Kaburaki 1977, Kuno 1987, Ariel 1991, Warren and Gibson 2001, MacWhinney in progress for discussion of several proposals).

The detailed effects of harmonic alignment on grammars can be explicitly modelled in Optimality Theory (OT). In phonology, the sonority hierarchy aligns with syllable structure so that the most sonorous sounds are attracted to syllable peaks and the least sonorous sounds to syllable margins (see Kager 1999 for a synthetic overview). Aissen (1999) proposes that syntax is analogous, with the most agentive semantic arguments attracted to the structural 'peak' of the clause, the Subject, and the least agentive arguments to the non-Subjects. Harmonic alignment is formally defined for a binary scale and an n-ary scale (Prince and Smolensky 1993: 136). The binary scale refers to a structure (of the syllable or clause, for example), and the n-ary scale refers to a substantive dimension such as sonority or proto-agentivity. Harmonic alignment produces two aligned Harmonic scales, one showing how the elements of the n-ary scale are distributed with respect to the high element of the binary scale, the other showing how they are distributed with respect to the low element.

Aissen's (1999) theory of harmonic alignment in syntax is illustrated in (4).

(4)	Prominence scales:	Harmonically aligned scales:	OT constraint subhierarchies:
	$\textsc{s} > \textsc{o}$	$\mathrm{S}_{ag} \succ \mathrm{S}_{pt}$	$^*\mathrm{S}_{pt} \gg {}^*\mathrm{S}_{ag}$
	agent > patient	$\mathrm{O}_{pt} \succ \mathrm{O}_{ag}$	$^*\mathrm{O}_{ag} \gg {}^*\mathrm{O}_{pt}$

The prominence scales on the left are analogous to the structural hierarchy of the syllable and the sonority hierarchy, respectively.[3] The upper and lower ends of these prominence scales are harmonically aligned as shown in the middle. On the right these alignments are expressed in OT terms as subhierarchies of markedness constraints prefixed with the '*' ('avoid') operator and inverted so that the most disharmonic combinations will receive the worst constraint violations.

[3]Aissen (1999) assumes a binarized relational hierarchy, adopting the binary scale Subject > Nonsubject, which encapsulates both $\textsc{s} > \textsc{o}$ and $\textsc{s} > \textsc{obl}$. She also assumes a role hierarchy based on proto-role theory (Dowty 1991, Asudeh 2001).

Other constraints may be interleaved in constraint subhierarchies, enhancing or suppressing their effects, but the relative ranking of the constraints in a subhierarchy is fixed across languages.[4] Harmonic alignment of the person hierarchy (1) with the relational hierarchy (see n. 3) yields further constraint subhierarchies, which may interact with the harmonic alignments in (4):

(5) $*S_3 \gg *S_{1,2}$, $*O_{1,2} \gg *O_3$, $*Obl_{1,2} \gg *Obl_3$

The markedness of the passive compared to the active follows from the universal subhierarchy $*S_{pt} \gg *S_{ag}$. For a semantically transitive verbal input, the active and not the passive is the optimal expression, all else being equal:

(6)

	input: v(ag,pt)	$*S_{pt}$	$*S_{ag}$
	passive: S_{pt}, Obl_{ag}	*!	
☞	active: S_{ag}, O_{pt}		*

But if the active is the optimal expression of a semantically transitive input, how can passivization occur at all? The answer is of course that other constraints favor the passive: avoiding or 'backgrounding' the agent (Shibatani 1985, Thompson 1987), avoiding subjects newer than non-subjects in the discourse (Birner and Ward 1998), placing the topic in subject position to enhance topic continuity (Givón 1983, Thompson 1987, Beaver 2000), etc. Tableau (7) illustrates this outcome for English, taking Birner and Ward's theory that passive subjects tend to be discourse-newer than non-subjects as the basis for the constraint $*S_{newer}$, an instance of Aissen's (1999) schematic constraint $*S_x$.

(7) English avoids subjects newer than non-subjects ($*S_{newer}$):

	input: v(ag/new, pt)	$*S_{newer}$	$*S_{pt}$	$*S_{ag}$
	active: S_{ag},O_{pt}	*!		*
☞	passive: S_{pt},Obl_{ag}		*	

Tableau (8) illustrates this outcome for Lummi.[5]

(8) Lummi avoids third person subjects ($*S_3$):

	input: v(ag/3,pt/1)	$*S_3$	$*S_{pt}$	$*S_{ag}$
	active: S_{ag},O_{pt}	*!		*
☞	passive: S_{pt},Obl_{ag}		*	

[4]In syntactically ergative languages (Manning 1996), the preference for agentive subjects must be overridden.

[5]This analysis of Lummi differs somewhat from that given in Aissen (1999); it was derived from the Lummi data by the Gradual Learning Algorithm. See below for further discussion.

Crosslinguistic variation comes from reranking (see Aissen 1999 for details). In languages without passives, the constraint *S_{pt} is undominated by any of these countervailing constraints. In general, the same constraints are hypothesized to be present in all grammars, but are more or less active depending on their ranking relative to other constraints. Thus Lummi falls back on *S_{newer} with third person agent and patient:

(9)

	input: v(ag/3/new,pt/3)	*S_3	*S_{newer}	*S_{pt}	*S_{ag}
	active: S_{ag},O_{pt}	*	*!		*
☞	passive: S_{pt},Obl_{ag}	*		*	

And English suppresses the relation/person constraints (*S_3, etc.) by low ranking:

(10)

	input: v(ag/3, pt/1)	*S_{newer}	*S_{pt}	*S_{ag}	*S_3
☞	active: S_{ag},O_{pt}			*	*
	passive: S_{pt},Obl_{ag}		*!		

We know this because the disharmonic combinations are still grammatical in English, unlike Lummi: *She met me, She'll be met by you*.[6]

3 Statistical Person/Voice Interactions in English

In the OT framework of the present study, following Aissen (1999), the active and passive are viewed as alternative candidate expressions of the same input person/role combinations. Evidence of person/voice interactions in English has been given previously (Svartvik 1966, Estival and Myhill 1988; Seoane Posse 2000, DeLancey 1981, Kuno and Kaburaki 1977, Kuno 1987, cf. Kato 1979, reviewed in Dingare 2001). But for the present study what is needed is information about the systematic choices made rather than information about the distributions of subsets within passives or actives. Prior studies generally fail to provide the full joint distribution, from which we can reconstruct the conditional frequencies needed.[7] We have therefore examined the parsed SWITCHBOARD corpus, a carefully designed database of spontaneous telephone conversations spoken by over 500 American English speakers, both male and female, from a great variety of speech communities (Godfrey et al. 1992, Marcus et al. 1993). The conversations average 6 minutes in length, collectively amounting to 3 million words of text. We have used the parsed portion of this corpus, which contains 1 million words.

[6] We note that Aissen's (1999) constraint subhierarchy *$O_{ag} \gg$ *O_{pt}, taken as a whole, penalizes transitivity and would therefore favor passives over actives if ranked high enough. We defer discussion of this problem to future work.

[7] Estival and Myhill (1988) provide exactly the kind of information needed for animacy and definiteness, but they provide person frequencies only for the patient role.

We have found that the same disharmonic person/argument associations which are avoided categorically in languages like Lummi by making passives either impossible or obligatory, are avoided in the SWITCHBOARD corpus of spoken English by either depressing or elevating the frequency of passives relative to actives. Compared to the rate of passivization for inputs of third persons acting on third persons (1.2%), the rate of passivization for first or second person acting on third is substantially depressed (0%) while that for third acting on first or second (2.9%) is substantially elevated:

(11) English person/role by voice (full passives)

action:			# Act:	# Pass:	% Act:	% Pass:
1,2	→	1,2	179	0	100.0	0.0
1,2	→	3	6246	0	100.0	0.0
3	→	3	3110	39	98.8	1.2
3	→	1,2	472	14	97.1	2.9

The leftmost column in (11) gives the four types of inputs (local person acting on local, local acting on nonlocal, etc.). We estimate the number of times each input was evaluated as the number of actives plus passives with that person/structure association. We then calculate the rate of passivization as the number of times that input was realized as passive.

Though the percentage of full passives (with *by* phrases) in spoken conversational English is very small, the person/voice effects are highly significant (χ^2 = 115.8, $p < 0.001$; Fisher exact test, $p < 0.001$). Similar significance levels result if short passives are included, but we omit them because the person of the agent is not always clear. See Dingare (2001) for further analysis and detailed methodological discussion.

In sum, the 'hard' grammatical constraints on person/voice interactions seen in languages like Lummi, Picurís, and Nootka continue to show up as statistical preferences in English.

4 Why is English like Lummi?

It is "a mainstay of functional linguistics" that "linguistic elements and patterns that are frequent in discourse become conventionalized in grammar" (from a publisher's blurb on Bybee and Hopper 2001). On this view, Lummi is simply at an extreme point from English along the continuum of conventionalization that connects frequentistic preferences in usage to categorical grammatical constraints. But as noted by Bresnan and Aissen (to appear), it remains unclear in a conventional generative syntax by what mechanisms usage preferences can harden into grammatical conventions:

> Classical generative theories of formal grammar are designed with mathematically discrete and logically deterministic formal architectures. On these

theories, frequentistic processes (such as the conventionalization of usage preferences) must belong either to grammar-external 'performance' along with speech errors and memory limitations, or to external choices among competing dialect grammars. Yet neither of these alternatives is an adequate model of variation and change, as first pointed out by Weinreich, Labov, and Herzog (1968). The same is true of the variable effects of markedness hierarchies on syntax.

— Bresnan and Aissen (to appear)

Stochastic Optimality Theory offers a useful approach to this phenomenon. It is one of a family of new optimization-based theories of grammar that can provide a unified account of categorical, variable, and gradient data (see Anttila in press, Manning to appear, and references).

5 Stochastic Optimality Theory

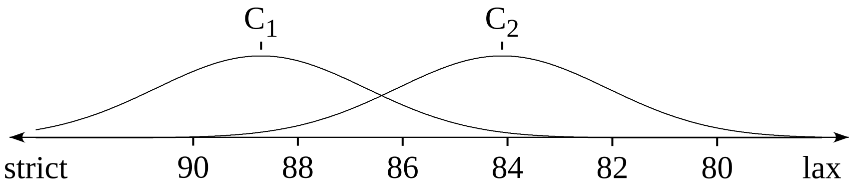
Stochastic OT (Boersma 1998, 2000, Boersma and Hayes 2001) differs from standard OT in two essential ways:

(i) **ranking on a continuous scale:** Constraints are not simply ranked on a discrete ordinal scale; rather, they have a value on the continuous scale of real numbers. Thus constraints not only dominate other constraints, but they are specific distances apart, and these distances are relevant to what the theory predicts.

(ii) **stochastic evaluation**: at each evaluation the real value of each constraint is perturbed by temporarily adding to its ranking value a random value drawn from a normal distribution. For example, a constraint with the mean rank of 99 could be evaluated at 98.12 or 100.3. It is the constraint ranking that results from these sampled values that is used in evaluation; it is referred to as the 'effective rank' here.

Figure 1 shows two constraints, C_1 and C_2. Note that the scale is inverted to match the standard OT convention that leftward is stronger in the constraint ranking. The ranks of these constraints are the means of their varying effective rankings, and are marked at the peaks of the bell curves; thus normally, $C_1 \gg C_2$. Nevertheless, on some evaluations the effective rank of C_1 will fall in the lower end of its normal distribution at the same time that the effective rank of C_2 falls in the higher end of its distribution, and a ranking reversal will occur, with $C_2 \gg C_1$. If C_1 and C_2 crucially conflict, such ranking reversals will create an alternative output for the same input, giving rise to variation.

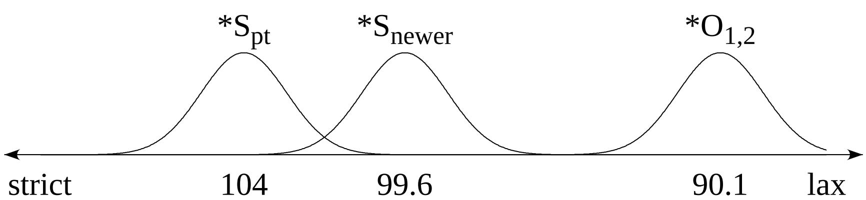
Figure 1: Constraint ranking on a continuous scale with stochastic evaluation



An OT grammar with stochastic evaluation can generate both categorical and variable outputs and can be learned from variable data by the GLA (Gradual Learning Algorithm, Boersma 1998, Boersma and Hayes 2001). Categorical outputs arise when crucially ranked constraints are distant. As the distance between constraints increases, interactions become vanishingly rare.[8] Variable outputs arise when crucially ranked constraints are closer together.

We can illustrate these ideas with the English 'pragmatic passive' grammar in Figure 2. The ranking values of the constraints (= the means of their normal distributions) are given on the X axis. The constraint $*S_{pt}$ penalizes passives, but it is close enough to the constraint $*S_{newer}$ to allow discernable variation. When ranking reversal occurs, as shown in the tableaux of (12), an alternative output occurs.

Figure 2: The English-type 'pragmatic passive'



[8] A distance of five standard deviations gives an expected reversal rate of less than 0.02% (Boersma and Hayes 2001: 50). Units of measurement are arbitrary. The standard deviation of ranking variation here is fixed at 2.0, so a ranking distance of 10 units between constraints is taken to be effectively categorical.

(12) Alternative outputs of the constraint ranking in Figure 2

	input: v(ag/new, pt)	$*S_{pt}$	$*S_{newer}$	$*O_{1,2}$
☞	active: S_{ag},O_{pt}		*	
	passive: S_{pt},Obl_{ag}	*!		

	input: v(ag/new, pt)	$*S_{newer}$	$*S_{pt}$	$*O_{1,2}$
	active: S_{ag},O_{pt}	*!		
☞	passive: S_{pt},Obl_{ag}		*	

In this stochastic grammar pragmatic or discourse-driven passivization is a statistical tendency, but not a categorical property of the output. Passives avoid subjects newer than non-subjects, but passivization is infrequent, and actives with new subjects also occur.

Where do the real number ranking values in a stochastic grammar come from? Starting from an initial state grammar in which all constraints have the same ranking values (arbitrarily set to be 100.0), the GLA is presented with learning data consisting of input-output pairs having the statistical distribution of, say, English. For each learning datum (a given input-output pair), the GLA compares the output of its own grammar for the same input; if its own output differs from the given output, it adjusts its grammar by moving all the constraints that disfavor its own output upward on the continuous ranking scale by a small increment, in order to make them apply more strictly, and moving all constraints that disfavor the given output downward along the scale by a small decrement, to relax their effects.[9] The adjustment process applies recursively to constraint subhierachies in order to preserve their local ordering relations.

6 Stochastic Grammars for English and Lummi

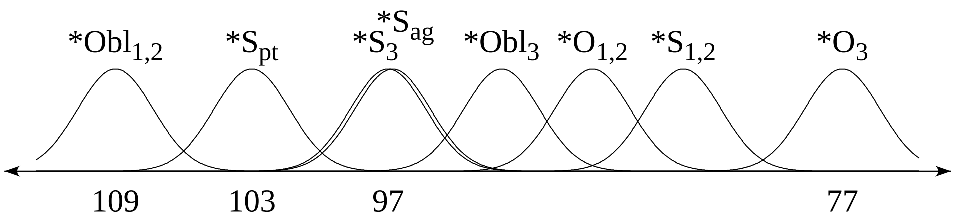
A partial stochastic grammar for the English person/voice interactions is given in Figure 3 and its output distribution in (13).

(13) Output distribution of the grammar in Figure 3

input:			% Active:	% Passive:
1,2	→	1,2	100.00	0.00
1,2	→	3	100.00	0.00
3	→	3	98.80	1.20
3	→	1,2	97.21	2.79

[9]The increment/decrement value is called the 'plasticity' and may be assumed to vary stochastically and to change with age (Boersma 2000).

Figure 3: Partial stochastic grammar of English



The constraint rankings and output distribution of this grammar were determined by simulation using the Praat system (Boersma and Weenink 2000), which includes an implementation of the Gradual Learning Algorithm.[10] Because of the stochastic components, the learned rankings and output distributions of grammars learned from the same distribution vary. The figures given are based on only one learned grammar; averaging over many grammars would better guarantee representativeness.

The constraints used are Aissen's (1999) constraint subhierarchies derived by harmonic alignment as outlined above. However, the grammar unrealistically omits the effects of the *S_{newer} constraint, which has a major influence on passivization in English (Birner and Ward 1998). Additionally, five of these constraints which are less active in our data were also omitted from the simulations for perspicuity: *Obl_{ag}, *Obl_{pt}, *O_{ag}, and *O_{pt}.

Observe that although the passive avoidance constraint *S_{pt} dominates the person-avoidance constraint *S_3, the two constraints are only 6 units apart (less than the near-categorical distance of 10; see n. 8), and will therefore produce low frequency variable outputs for some inputs. For inputs where only the agent is third person, passive outputs will occasionally be favored by *S_3, as shown in the tableau in (14):

(14) An (infrequent) effect of *S_3 on passive outputs:

	input: v(ag/3,pt/1)	*S_3	*S_{pt}	*S_{ag}
	active: S_{ag},O_{pt}	*!		*
☞	passive: S_{pt},Obl_{ag}		*	

[10]For learning the constraint rankings, a distribution of input-output pairs of person/voice combinations was specified according to the proportions given in our data in (11). The GLA learned directly from these distributions using the default settings in the Praat system for plasticity and replications. The relative ranking of constraints (the means) in the subhierarchies was maintained.

When both agent and patient are third person, the $*S_3$ constraint cannot decide between active and passive, and the decision passes to other constraints.[11]

The highest ranked constraint in Figure 3 is $*Obl_{1,2}$, which penalizes local person passive agents. It is more than 10 units (see n. 8) above any constraint that would disfavor an active (namely, $*O_{1,2}$ for an input with local-person patient and $*O_3$ for an input with third-person patient). These rankings reflect the zero frequency of local person passive agents in our data. Local person passive agents have been described as unacceptable in English (Kuno and Kaburaki 1977). Though they may be dispreferred or contextually marked, they are grammatical in spoken English. Among the examples cited by Kato (1979) are those in (15):

(15) a. I said, "Me watch it! Fuck that! Let him watch it." He was hired by me. I could fire him if I didn't like him. (Studs Terkel, *Working*)

b. When somebody says to me, "You're great, how come you're *just* a waitress?" *Just* a waitress. I'd say, "Why, don't you think you deserve to be served by me?" (Studs Terkel, *Working*)

With more training data and a more complete constraint set which includes factors of topicality and focus, our model should learn grammars that produce passives with local person agents. Note that if the ranking value of $*Obl_{1,2}$ in the grammar of Figure 3 were lowered from 109 to 104, the output of local person passives would increase to one-tenth of one percent, 0.1%, while barely changing the frequency of other outputs.

In sum, stochastic OT can capture the soft influence of person on English passivization that exists beneath the level of grammaticality judgments. Disharmonic person/argument combinations are grammatical but avoided, affecting the frequency of passivization.

Unfortunately we lack a parsed SWITCHBOARD corpus for Lummi. Nevertheless, it is possible to show by simulation how the descriptions of passive/voice interactions in Lummi grammar can also be captured by a stochastic OT grammar. We interpret the descriptions of Lummi from Jelinek and Demers (1983, 1994) by means of a simple distribution. Where a sentence type is described as ungrammatical, we assign it 0% relative frequency; where it is described as obligatory, we assign it 100%; and where it is described as optional, we assign it 50%:

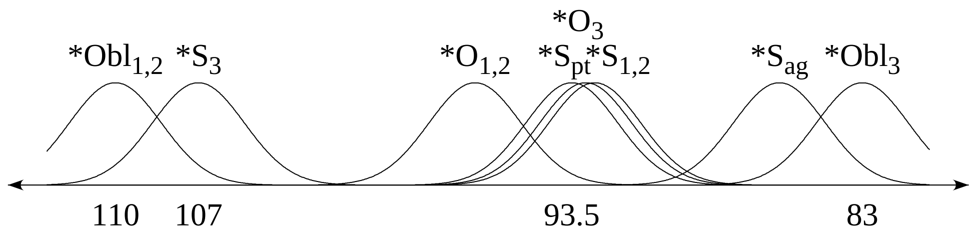
[11]For this input it will be $*S_{ag}$ that permits passive ouputs, with slightly less frequency than the passive outputs produced by $*S_3$, which is ranked marginally higher. In a less limited grammar other constraints would fill this role.

(16) Simulated Lummi (Straits Salish) input/output distribution

input:			% Active:	% Passive:
1,2	→	1,2	100.00	0.00
1,2	→	3	100.00	0.00
3	→	3	50.00	50.00
3	→	1,2	0.00	100.00

The simulated input/output distribution in (16) is then used to generate training data for the GLA, as before. The initial state of the grammar and the training regime (n. 10) are exactly the same as for English. A partial Lummi grammar learned by the GLA is shown in Figure 4.

Figure 4: Partial stochastic grammar of Lummi



Note that in contrast to the English grammar in Figure 3, the syntactic person-avoidance constraint $*S_3$ in the Lummi grammar is more than 10 units (5 standard deviations) above $*S_{pt}$. This ranking yields the obligatory passivization of inputs with local person patients and non-local person agents, capturing the categorical influence of person on Lummi passivization. The output distribution of the grammar in Figure 4 matches that in (16).

It might be thought that ranking on the continuous scale of real numbers is powerful enough to learn any distribution. In fact, under the present theory there are no stochastic OT grammars for 'anti-Lummi' or 'anti-English' distributions, which reverse the generalizations embodied in our data. Greater relative frequency of passives for first or second person acting on third would imply that third person subjects are avoided less than first or second person subjects. If so, then $*S_{1,2}$ must dominate $*S_3$ for a greater proportion of evaluations. But that ranking violates the subhierarchy in (5), which in stochastic OT requires the *mean* ranking values of these constraints to occur in the reverse order.

Thus, stochastic OT grammars are limited to subspaces of distributions that conform to the theory embodied in the constraint set. They are not general-purpose statistical analyzers and they have no special memory for frequencies (Boersma 2000).

7 Conventionalization and Frequency

Stochastic OT grammars allow us to place the person/voice interactions in English and Lummi at points on a continuum of conventionalization that connects frequentistic preferences in usage to categorical grammatical constraints. If this general perspective is correct, then we would expect to find languages at intermediate points on this same continuum. In the domain of person/voice interactions, Squamish (Coast Salish, British Columbia) may be a case in point.

Squamish and Lummi differ in their treatment of first person patient and nonlocal agent combinations: passivization is obligatory in Lummi but optional in Squamish; with second person patients passivization is obligatory for both languages. Within the ordinal OT framework, Aissen (1999) analyzes the difference between Lummi and Squamish by this constraint ranking:

(17) Lummi constraint ranking: $\ldots\ {*}O_1, {*}O_2 \gg \ldots\ {*}S_{pt}$

Squamish constraint ranking: $\ldots\ {*}O_2 \gg \ldots {*}S_{pt} \gg {*}O_1 \ldots$

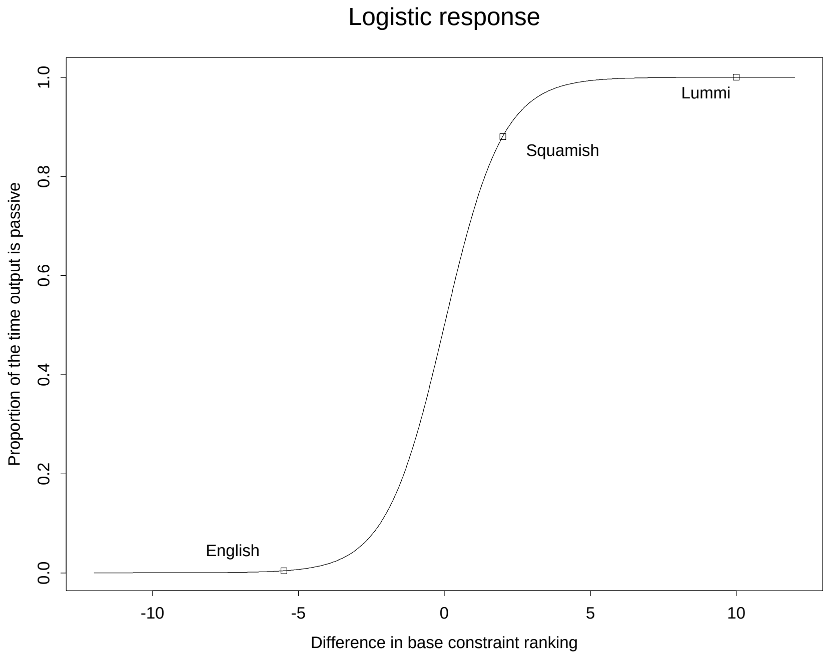
Given independent evidence that languages differ in whether first or second person is dominant (DeLancey 1981), Aissen assumes that the mutual ranking of the local-person avoidance constraints is not fixed by the subhierarchy, but subject to crosslinguistic variation.

However, it is not fully informative to say, as Aissen (1999) does, that passivization with third person agents and first person patients is optional in Squamish. In terms of what is preferred rather than what is merely possible, Squamish is described as being much the same as Lummi, “except that third person acting on first may be active, though commonly passive” (Klokeid 1969: 11). Thus in Squamish as in English, passives of the type *I was fooled by her* are optional alternatives to actives with disharmonic local-person objects: *She fooled me*. But in spoken English, as we have seen, such passives are exceedingly infrequent, far less common than the corresponding actives, while in Squamish they are more frequent than the corresponding actives.

In stochastic OT the rerankings postulated by Aissen (1999) take place on a continuous scale and imply *changes in frequency* as well as changes in grammaticality. The high rate of passivization with first person patients in Squamish shows that a constraint favoring passive, such as ${*}O_1$ is still ranked considerably above a constraint favoring active, such as ${*}S_{pt}$, on our continuous constraint ranking scale. Squamish and Lummi are closely related Coast Salish languages. It is plausible that we are observing a change in progress: the two languages represent different points in the changing categoricity of person effects on the passive, reflected in the ranking of the person-avoidance constraints for first and second person.

If a process of historical change is modeled by the movement in strength of a constraint along the continuous scale, as implied by the stochastic OT model, then (all else being equal) smooth changes in the relative frequencies of usage are predicted. Note however that although the change is smooth, it is not predicted to be linear. Rather, if a constraint reranking is crucial to the choice between two outputs, the prediction is that we should see an 'S' curve between the proportion of occurrences of the two outputs, of the sort that has been widely remarked on in historical and socio-linguistics (Weinreich, Labov, and Herzog 1968, Bailey 1973, Kroch 2001).

Figure 5: A conventionalization continuum.



More technically, assuming that the difference in ranking of two constraints which are crucial to the choice between two outputs A and B is changing linearly, then the proportion of output A is given by a logistic curve (Figure 5): when the constraints are at least 5 standard deviations apart, the proportion of the disfavored output is negligible; at 2 standard deviations, the rate reaches about 8%, but then it increases much more rapidly to 50% of each output when the two constraints are equally ranked.

These considerations suggest that classical grammatical descriptions in terms of what is 'possible' or 'grammatical' are overly idealized, concealing grammatically significant statistical structure beneath the idealization of linguistic intuitions of grammaticality.

8 Conclusion

The same categorical phenomena which are attributed to hard grammatical constraints in some languages continue to show up as statistical preferences in other languages, motivating a grammatical model that can account for soft constraints.

This observation is not new. Givón (1979: 26–31) already made this point forcefully over twenty years ago. What is new here is our demonstration that the stochastic OT framework can provide an explicit and unifying theoretical framework for these phenomena in syntax.

References

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language & Linguistic Theory* 17: 673–711.

Anttila, Arto. In press. Variation and phonological theory. In J. Chambers, P. Trudgill and N. Schilling-Estes (eds.), *Handbook of language variation and change*. Oxford: Blackwell.

Asudeh, Ash. 2001. Linking, optionality, and ambiguity in Marathi. In P. Sells (ed.), *Formal and Empirical Issues in Optimality-Theoretic Syntax*, pp. 257–312. Stanford, CA: CSLI Publications.

Bailey, Charles-James N. 1973. *Variation and Linguistic Theory*. Arlington, Virginia: Center for Applied Linguistics.

Beaver, David. 2000. The optimization of discourse. Stanford: Stanford University Department of Linguistics MS. Draft on-line, Stanford University: http://www.stanford.edu/~dib/paper.htm.

Birner, Betty J. and Gregory Ward. 1998. *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins.

Boersma, Paul. 1998. *Functional Phonology. Formalizing the interaction between articulatory and perceptual drives*. Amsterdam: University of Amsterdam Ph.D. thesis.

Boersma, Paul. 2000. Learning a grammar in Functional Phonology. In J. Dekkers, F. v. d. Leeuw, and J. v. d. Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*, pp. 465–523. Oxford: Oxford University Press.

Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32: 45–86.

Boersma, Paul and David Weenink. 2000. Praat computer program. On-line, Institute of Phonetic Sciences, University of Amsterdam: http://www.fon.hum.uva.nl/praat/.

Bresnan, Joan and Judith Aissen. To appear. Optimality and functionality: Objections and refutations. *Natural Language & Linguistic Theory* 20.1 Draft on-line, Stanford University: http://www-lfg.stanford.edu/bresnan/download.html.

Bresnan, Joan. 2001. *Lexical-functional Syntax*. Oxford: Blackwell.

Bybee, Joan and Paul Hopper (eds). 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.

Dahlstrom, Amy. 1984. *Plains Cree Morphosyntax*. Berkeley: University of California, Berkeley Ph.D. dissertation. Published in 1991. New York: Garland.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57: 626–57.

Dingare, Shipra. 2001. *The effect of feature hierarchies on frequencies of passivization in English*. Stanford, California: Stanford University Department of Linguistics Master's thesis. On-line, Rutgers Optimality Archive: http://ruccs.rutgers.edu/roa.html. ROA-467-0901.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.

Emanatian, Michele. 1988. The Nootka passive revisited. In W. Shipley (ed.), *In Honor of Mary Haas. From the Haas Festival Conference on Native American Linguistics*, pp. 265–91. Berlin: Mouton de Gruyter.

Estival, Dominique and John Myhill. 1988. Formal and functional aspects of the development from passive to ergative systems. In Shibatani (ed.), pp. 441–91.

Forrest, Linda B. 1994. The de-transitive clauses in Bella Coola: Passive vs. inverse. In Givón (ed.), pp. 147–68.

Garrett, Andrew. 1990. The origin of NP split ergativity. *Language* 66: 261–296.

Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.

Givón, Talmy (ed.) 1983. *Topic Continuity in Discourse. A Quantified, Text-based Study.* Amsterdam: John Benjamins.

Givón, Talmy. (ed.) 1994. *Voice and Inversion*. Amsterdam: Benjamins.

Godfrey J., E. Holliman and J. McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of ICASSP-92*, San Francisco, pp. 517–520.

Greenberg, Joseph. 1966. *Language Universals. With Special Reference to Feature Hierarchies.* The Hague: Mouton.

Haspelmath, Martin. 1990. The grammaticalization of passive morphology. *Studies in Language* 14: 25–72.

Jacobs, Peter. 1994. The inverse in Squamish. In Givón (ed.), pp. 121–45.

Jelinek, Eloise and Richard Demers. 1983. The agent hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics* 49: 167–85.

Jelinek, Eloise and Richard Demers. 1994. Predicates and pronominal argments in Straits Salish. *Language* 70: 697–736.

Kato, Kazuo. 1979. Empathy and passive resistance. *Linguistic Inquiry* 10: 149–52.

Klaiman, M. H. 1991. *Grammatical Voice*. Cambridge: Cambridge University Press.

Klokeid, Terry J. 1969. Notes on the comparison of Wakashan and Salish. *University of Hawaii Working Papers in Linguistics* 7: 1–19.

Kroch, Anthony. 2001. Syntactic change. In M. Baltin and C. Collins (eds.), *The Handbook of contemporary syntactic theory*, pp. 699–729. Oxford: Blackwell Publishers.

Kuno, Susumu. 1987. *Functional Syntax*. Chicago: University of Chicago Press.

Kuno, Susumu and Etsuko Kaburaki. 1977. Empathy and syntax. *Linguistic Inquiry* 8: 627–72.

Manning, Christopher. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford, California: CSLI Publications.

Manning, Christopher. To appear. Probabilistic approaches to syntax. In R. Bod, J. Hay and S. Jannedy (eds.), *Probability Theory in Linguistics*. Cambridge, Massachusetts: The MIT Press. Draft on-line, Stanford University: http://www-nlp.stanford.edu/∼manning/papers/.

Marcus, Mitchell, Beatrice Santorini, and May Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn treebank. *Computational Linguistics* 19: 313–330.

Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.

Plank, Frans. (ed.) 1979. *Ergativity. Towards a Theory of Grammatical Relations*. New York: Academic Press.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCS Technical Report #2. Piscataway, NJ: Rutgers University Center for Cognitive Science.

Seoane Posse, Elena. 2000. The passive as an object foregrounding device in Early Modern English. In R. Bermudez-Otero, D. Denison, R. M. Hogg, C. B. McCully (eds.), *Generative Theory and Corpus Studies: A Dialogue from 10 ICEHL*, pp. 211–32. Berlin: Mouton de Gruyter.

Shibatani, Masayoshi. 1985. Passives and related constructions: A prototype analysis. *Language* 61: 821–48.

Shibatani, Masayoshi (ed.). 1988. *Passive and Voice*. Amsterdam: Benjamins.

Svartvik, Jan. 1966. *On Voice in the English Verb*. The Hague: Mouton.

Trask, R. L. 1979. On the origin of ergativity. In Plank (ed.), pp. 385–404.

Thompson, Sandra A. 1987. The passive in English: A discourse perspective. In R. Channon and L. Shockey (eds.), *In Honor of Ilse Lehiste*, pp. 497–511. Dordrecht: Foris Publications.

Warren, Tessa and Edward Gibson. 2001. The influence of referential processing on sentence complexity. Cambridge, Massachusetts: MIT Department of Brain and Cognitive Sciences MS. Submitted for publication.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (eds.), *Directions for historical linguistics. A symposium*, pp. 95–188. Austin: University of Texas Press.

Whistler, Kenneth W. 1985. Focus, perspective, and inverse person marking in Nootkan. In J. Nichols and A. C. Woodbury (eds.), *Grammar Inside and Outside the Clause. Some Approaches to Theory from the Field*, pp. 227–65. Cambridge: Cambridge University Press.

Zaharlick, Amy. 1982. Tanoan studies: Passive sentences in Picurís. *Ohio State University Working Papers in Linguistics* 26: 34–48.

MODELLING POSSESSOR CONSTRUCTIONS IN LFG: ENGLISH AND HUNGARIAN

Erika Chisarik and John Payne

Department of Linguistics, University of Manchester, Manchester M13 9PL, England

chisarikerika@hotmail.com
John.Payne@man.ac.uk

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

English and Hungarian are languages with more than one structurally distinct possessor construction. In this paper, it is argued that noun phrases in such languages require the postulation of two unrestricted argument functions: SUBJ and NCOMP. The SUBJ function in noun phrases parallels the sentential SUBJ function, but NCOMP is special to noun phrases and distinct from the sentential OBJ function, not least in the range of semantic roles it encompasses. For example, NCOMP like SUBJ but unlike OBJ permits agent roles. In English, as in Hungarian, a single unrestricted argument of an appropriate role can in principle map either to SUBJ or to NCOMP. The two functions however stand in a hierarchical relation. In both languages, the structurally higher possessor construction maps onto SUBJ and the structurally lower construction maps onto NCOMP. Also, given a thematic hierarchy for noun phrase arguments analogous to that required for clauses, when two unrestricted arguments are present, there is an asymmetry in that the higher role must map to SUBJ and the lower to NCOMP.

1 Introduction

English and Hungarian are examples of languages with more than one basic possessor construction, as illustrated in [1] and [2]:[1]

[1] a. [**the king's**] daughter [genitive]
b. the daughter [**of the king**] [*of*-oblique]

[2] a. [**a király-nak**] a lány-a [dative]
ART king-DAT ART daughter-SUBJ[3]
'the king's daughter'

b. [**a király**] lány-a [genitive]
ART king daughter-NCOMP[3]
'the king's daughter'

We will refer to the English constructions in [1a] and [1b] as the genitive and the *of*-oblique respectively. In the genitive construction, the possessor is an NP in the genitive case which is a pre-head dependent and which simultaneously functions as a definite determiner. In the *of*-oblique, the possessor is a PP headed by the preposition *of* which is a post-head dependent and lacks any determiner function.

The Hungarian constructions in [2a] and [2b] are standardly known as the dative possessor construction and the nominative possessor construction (e.g. Szabolcsi 1994, Laczkó 2000). In the dative construction, the possessor is a dative-marked NP which is structurally a predeterminer, preceding for example the definite article in [2a]. The definite article can be replaced by other determiners, for example the indefinite article in *[**a király-nak**] egy lány-a* (ART king-DAT a daughter-SUBJ[3]) 'a daughter of the king'. By contrast, somewhat like the English genitive, the so-called nominative possessor in [2b] is in complementary distribution with the definite article and, in the absence of any specific indication of the indefiniteness of the noun phrase, functions as a definite determiner. In both constructions, the possessum formally agrees with the possessor.

In a recent paper (Payne & Chisarik 2001), we have argued that the case of the possessor in [2b] is strictly not to be analysed as nominative, but rather as a new genitive resulting from reanalysis of the definite article *a*/*az* as a case prefix. Possessive pronouns, for example, have a distinct genitive paradigm:

[3] a. az-én lány-om
GEN-I daughter-NCOMP[1SG]
'my daughter'

b. a-mi lány-unk
GEN-we daughter-NCOMP[1PL]
'our daughter'

c. a-te lány-od
GEN-you daughter-NCOMP[2SG]
'your(sg) daughter'

d. a-ti lány-otok
GEN-you daughter-NCOMP[2PL]
'your(pl) daughter'

e. az-ő lány-a
GEN-he/she daughter-NCOMP[3SG]
'his/her daughter'

f. az-ő lány-uk
GEN-they daughter-NCOMP[3PL]
'their daughter'

Compare nominative *én* 'I', *te* 'you(sg)', *ő* 'he/she', *mi* 'we', *ti* 'you(pl)', *ők* 'they' with their corresponding genitive forms in [3]. Especially notable is the third-person plural, which is identical to the third-person singular in the genitive, but has the distinctive plural marking *–k* in the nominative. In this analysis, personal proper names such as *Mari* also optionally and with dialectal variation have a distinct genitive form: *a-Mari lány-a* or *Mari lány-a* (GEN-Mari daughter-NCOMP[3]) 'Mari's daughter'. One indication that reanalysis of the article has taken place is the compatibility of these genitive forms with indefiniteness of the noun phrase. In an example like [4], the new genitive is compatible with the indefinite determiner *egyik* 'a/one':[2]

[4] az-én egyik lány-om
GEN-I a daughter-NCOMP[1SG]
'a daughter of mine'

We will therefore refer to this construction henceforth as the genitive construction.

2 Grammatical functions

What grammatical function or functions should be associated with the possessor constructions in [1] and [2]? A fairly standard view in LFG is that possessors are linked to a distinct POSS function (e.g. Laczkó 1995, 2000; Sadler 2000, Bresnan 2000, Falk 2001). In some works (e.g. Sadler 2000), the POSS function is explicitly stated to be a subspecies of the SUBJ function, i.e. unrestricted and discourse-oriented. The goal of this paper is to show that this is not the whole story. It is indeed necessary to postulate an unrestricted and discourse-related function associated with possessors. Because of the close parallels with clause structure we will identify this function with SUBJ, rather than employing the term POSS. This is the function associated with the structurally higher genitive possessor in English and the structurally higher dative possessor in Hungarian.

The lower *of*-oblique possessor in English and the genitive possessor in Hungarian cannot however be associated with the same function: we will argue that the function required is unrestricted, but not discourse-oriented. It cannot be the unrestricted object function OBJ, firstly because nouns are generally barred from taking object complements, and secondly because of the semantic range it encompasses. This range includes inter alia agent-like roles, which standard mapping principles for clause structure naturally prevent from mapping to OBJ. An English noun phrase in which the *of*-oblique has an agent role would be *the bad performance of the team in yesterday's match*, where *the team* has the same role as the clausal subject in *the team performed badly in yesterday's match*. Also, as we will show in section 5, the range of semantic roles permitted by the required function is, taking the evidence of English and Hungarian, in principle even wider than that permitted by the SUBJ function. It would not seem desirable therefore to identify this function in LFG terms as a restricted oblique function (the solution essentially proposed by Rappaport 1983). We argue that what is needed is a new complement function which we will term NCOMP.

If the basic argument functions are decomposed not just in terms of [±r] (restricted) and [±o] (objective) (Bresnan 2000: 308), but also in terms of [±d] (discourse-related), and SUBJ is the only [+d] function, then NCOMP naturally fills the [–r, –o] slot in the set of [–d] functions. The range of permitted argument functions is then shown in table [5]:

[5]

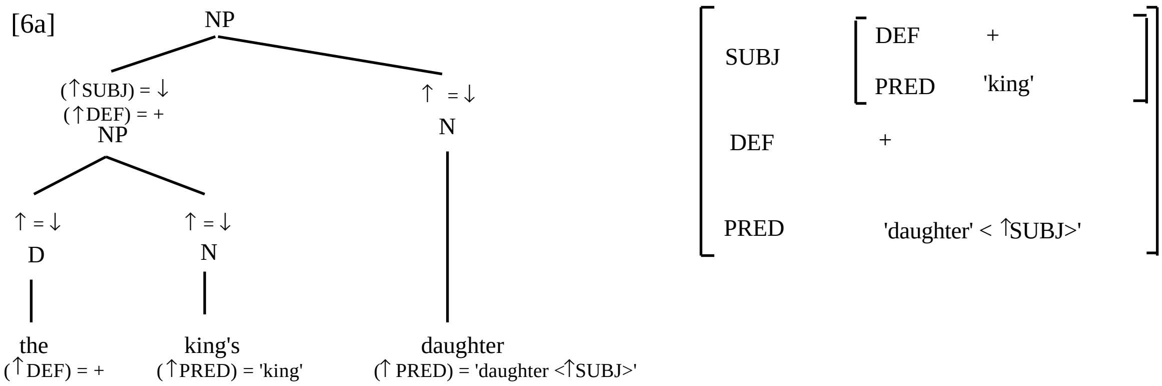
	+d	–d	
	–r	–r	+r
–o	SUBJ	NCOMP	OBL_θ
+o		OBJ	OBJ_θ

The [+o] functions are excluded from noun phrases, and NCOMP is excluded from clauses, where only a single [–r, –o] function (SUBJ) is permitted. The peculiarity of NCOMP is that it has a semantic range covering roles in noun phrase structure which match those of both SUBJ and OBJ in clause structure, but when it co-occurs with SUBJ, NCOMP is hierarchically subordinate to SUBJ.

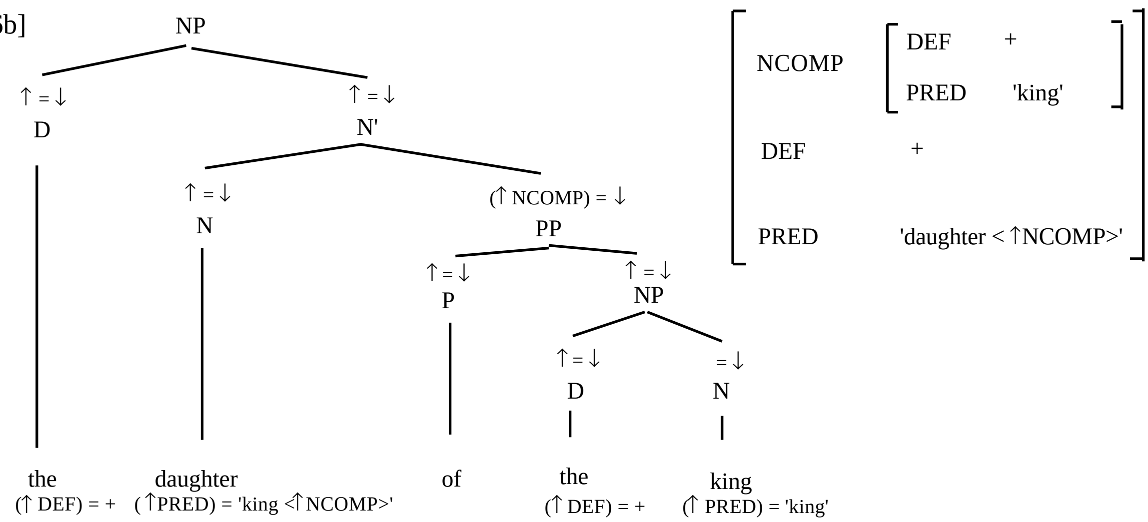
3 Syntax (English)

The analysis of the English examples [1a] and [1b] is then straightforward. We assume an NP analysis rather than a DP analysis, but the arguments for this are orthogonal to the analysis of possessor constructions and we will not rehearse them here. For a full discussion of the issues as they relate to English, though not in an LFG framework, see Payne & Huddleston (to appear, 2002). Consider the c-structures and corresponding f-structures in [6a] and [6b]:

[6a]



[6b]



The genitive NP *the king's* in [6a] acts as a definite determiner: the definiteness of the matrix NP follows from the ($\uparrow$DEF) = + annotation on the genitive NP. Through its ($\uparrow$SUBJ) = $\downarrow$ annotation, the genitive NP maps straightforwardly to the SUBJ function in f-structure. The semantic role here is one of kin relationship. The same semantic role of kin relationship is also associated with the NCOMP function in [6b], where *of the king* is an *of*-oblique. Here we take the ($\uparrow$NCOMP) = $\downarrow$ annotation to be associated with the PP node. Note that the preposition *of* in [6b] is then treated as having no f-structure annotation itself, and consequently contributes nothing to the f-structure representation. This reflects the longstanding idea that the preposition *of* in these constructions has no semantic value and serves merely to create a syntactically oblique PP complement for the head noun. The analysis of the preposition *of* in the NCOMP construction would then contrast with the analysis of prepositions heading PPs which map to semantically restricted OBL_θ functions, the value of θ being individually supplied by the preposition concerned.

A number of factors are involved in the selection of the SUBJ rather than the NCOMP function. The syntactic and pragmatic factors include the following (Jucker 1993; Payne & Huddleston (to appear, 2002)):

Pronoun vs Non-Pronoun

Pronouns are strongly preferred as SUBJ rather than NCOMP:

[7]	a	her car	a' ??the car of her
	b.	her only portrait	b'.the only portrait of her

It is tempting and plausible to relate this preference to the discourse-oriented nature of SUBJ: the unmarked role of an overt pronoun in English is to function as a topic, and pronouns gravitate to the discourse-related function within NP. In [7a], for example, it is typically necessary to identify the possessor from the discourse before we can identify the car. But this is not an absolute constraint: in [7b] the pronoun can equally be an NCOMP. Semantic factors are involved in this example: *her* in [7b'] is likely to be relatively low on the thematic hierarchy, i.e. the person depicted in the portrait, but in [7b] it can also be a creator or owner.

Weight

Other things being equal, relatively short, head-final NPs favour the SUBJ function and relatively long NPs with post-head dependents favour the NCOMP function. This is clearly related to the pre-head position of SUBJ and the post-head position of NCOMP in English rather than any deep property of the functions themselves. Compare the examples in [8]:

[8] a. John's speech
a' ?the speech of John

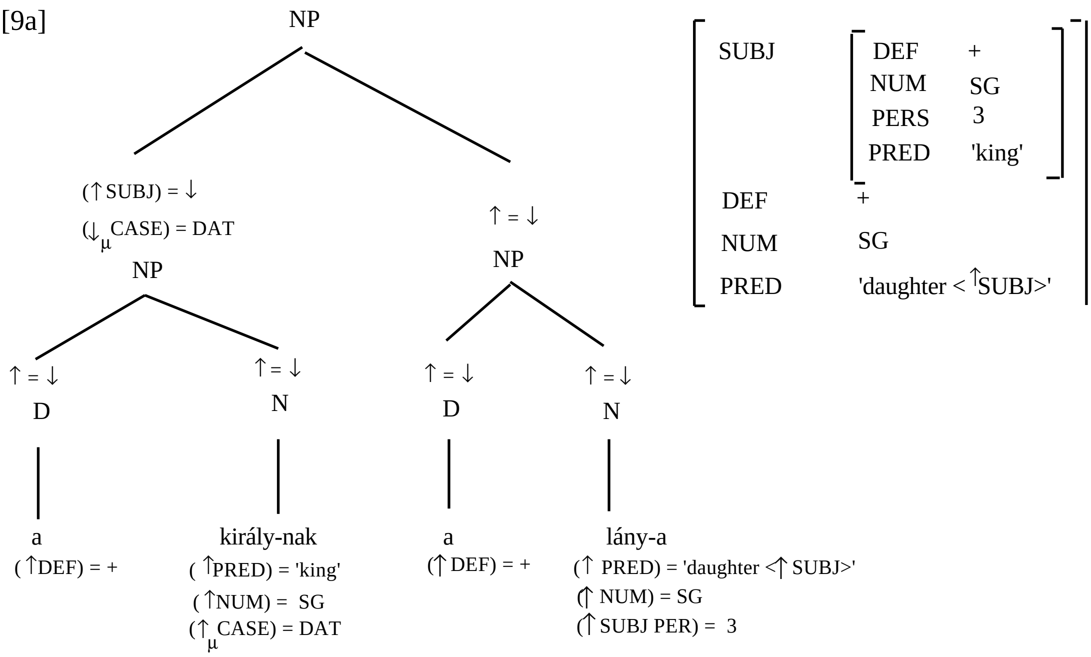
b. ?the delegate from the Republic of Chad's speech
b'. the speech of the delegate from the Republic of Chad

The possessor in [8a] is the short, head-final NP *John*, and the pre-head SUBJ position is preferred to the post-head NCOMP position in [8a']. By contrast, NCOMP position is preferred for the relatively long NP *the delegate from the Republic of Chad*, which itself contains a post-head dependent *from the Republic of Chad*. It can be noted that [8b'] is preferable to [8a'] even when the semantic role involved, that of agent, is high on the thematic hierarchy.

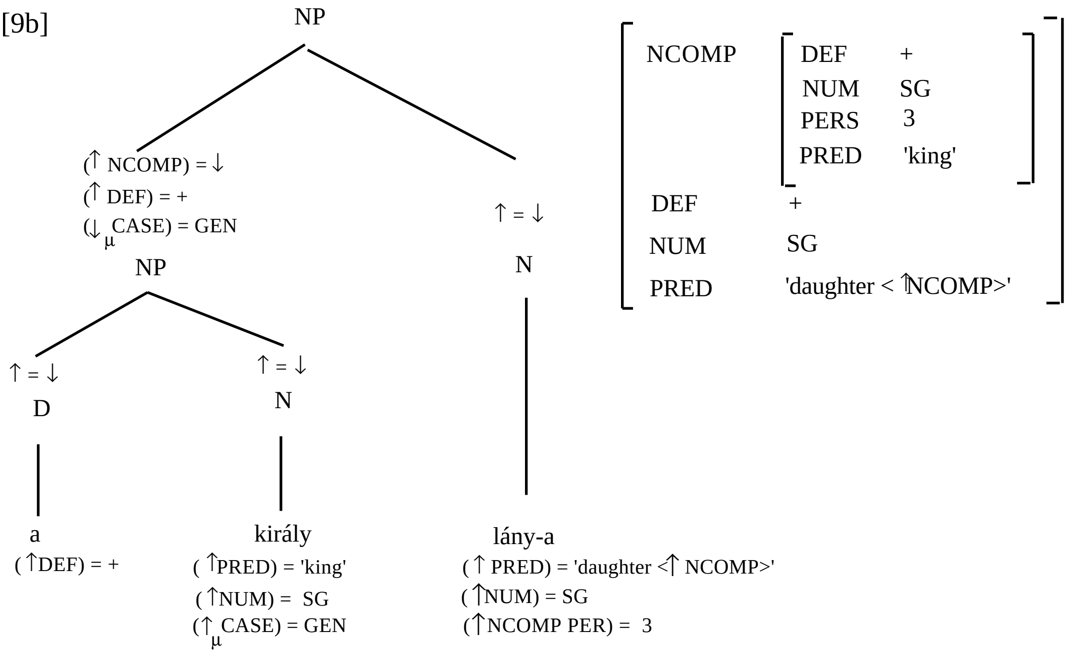
4 Syntax (Hungarian)

The representation of the Hungarian examples [2a] and [2b] is similar to the representation of the English noun phrases in [6a] and [6b] , but it is also necessary to take into account the agreement of the possessum:[3]

[9a]



[9b]



In [9a], the dative possessor is analysed as a predeterminer, forming a new NP-level constituent from an NP which already contains the definite article *a*. It is therefore the article rather than the possessor which is annotated with the (↑DEF) = + equation and determines the definiteness of the matrix NP. Evidence for the constituent structure assumed here comes from coordinate constructions: a single possessor can act as predeterminer to conjoined NPs, as shown in [10]:

[10]	a	**király-nak**	$[_{NP}$ [	a	fi-a]	és	[a	lány-a]]
	ART	king DAT		ART	son-SUBJ[3]	and	ART	daughter-SUBJ[3]
	‘the king’s son and daughter’							

By contrast, in [9b] the genitive possessor stands in complementary distribution with the definite article and acts, as in English, as a definite determiner. The genitive NP itself therefore bears the (↑DEF) = + equation. Although the dative possessor and genitive possessor cannot co-occur, the fact that the dative possessor is invariably a predeterminer rather than a determiner suggests that it is structurally higher than the genitive possessor. This is reflected in the mapping of the dative possessor to the SUBJ function in f-structure, and the mapping of the genitive possessor to NCOMP. It should be noted that, from a universal point of view, these mappings do not depend on the particular cases employed, or whether there is exponence through adpositions or cases. The case called genitive in English maps to SUBJ, while the case called genitive in Hungarian maps to NCOMP. The exponence of NCOMP in English is prepositional, but in Hungarian it involves morphological case. These are purely surface phenomena: the mapping of the possessor NP to SUBJ or NCOMP is rather determined by the hierarchical syntactic and semantic relationships involved, and by the discourse-related nature of SUBJ.

The agreement of the head noun with the possessor in Hungarian is straightforwardly handled by the annotations ↑SUBJ PERS = 3 and ↑NCOMP PERS = 3 on the agreeing head nouns in [9a] and [9b] respectively. It should be noted that the possessum *lány-a* ‘daughter’ in examples [9a] and [9b] requires only that the possessor NP be third person: there is no specification of number. The same form in *-a*, which is identical to the form which shows agreement with singular third-person pronouns as shown in example [3a], co-occurs both with singular and plural possessors if these are headed by non-pronouns.

There is however a small but significant difference in agreement between the SUBJ and NCOMP functions: plural NCOMPs are incompatible with a possessum distinctively marked for a plural possessor, but SUBJs are not. Compare [11a-d]:

[11]	a.	a	lány-ok	macská-ja	
		ART	girl-PL	cat-NCOMP[3]	
		‘the girls’ cat’			
	b.	*a	lány-ok	macská-juk	
		ART	girl-PL	cat-NCOMP[3PL]	
	c	a	lány-ok-nak	a macská-ja	
		ART	girl-PL-DAT	cat-SUBJ[3]	
	d.	a	lány-ok-nak	a	macská-juk
		ART	girl-PL-DAT	ART	cat-SUBJ[3PL]

Example [11a], with a plural NCOMP and general third-person agreement in *–a*, is grammatical, but [11b], with distinctively plural third-person agreement in *–juk*, is excluded. Both forms of agreement are however permitted with a plural SUBJ as in [11c] and [11d]. A full treatment of the agreement system requires firstly that a proper distinction be made between pronoun incorporation and agreement, and secondly that agreement with pronouns and non-pronouns is differentially treated, perhaps invoking binding features along the lines of Bresnan (2000: 146). The point made here is simply that, although the agreement paradigms of head nouns with SUBJ and NCOMP arguments largely overlap, they are not completely identical. It is therefore an advantage rather than a defect of the present analysis that the lexicon must contain two representations of head nouns such as *lány-a* ‘daughter’ in [9a] and [9b], one with the annotation (↑PRED) = ‘daughter <↑SUBJ>’ and one with the annotation (↑PRED) =

'daughter <↑NCOMP>'. The fact that the paradigms overlap can be handled by lexical redundancy rules.

As in English, the two constructions display syntactic and pragmatic differences. These include:

Syntactic type

A number of NP types cannot appear in the NCOMP construction, for example NPs containing demonstratives:

[12]		*[ez	a	lány]	ruhá-ja	
		this.*GEN	ART	girl.GEN	dress-SG.NCOMP[3]	
		'this girl's dress'				
	b.	[en-nek	a	lány-nak]	a	ruhá-ja
		this-DAT	ART	girl-DAT	ART	dress-SG.SUBJ[3]
		'this girl's dress'				

In Payne & Chisarik (2001), we handle such restrictions by assuming a paradigm gap. The demonstrative, which agrees in case with the NP, has a dative but not a genitive form.

Prodrop

The possessum forms can occur on their own, with pro-drop. For example, *a lány-om* (ART daughter-SUBJ[1SG]) is the unmarked form for 'my daughter'. Not only are overt topic pronouns omitted in the SUBJ function, but it is actually ungrammatical to insert them: **nekem a lány-om* (I-DAT ART daughter-SUBJ[1SG]).[4] This can be straightforwardly handled using the mechanism proposed by Bresnan (2000: 146), viz *lány-om* is lexically specified as follows:

[13] *lány-om*: ↑PRED = 'daughter <↑ SUBJ>
↑NUM = SG
↑NCOMP PRED = 'PRO'
↑NCOMP NUCL = –
↑NCOMP PERS = 1
↑NCOMP NUM = SG

When the annotation ↑NCOMP PRED = 'PRO' is present, the form *lány-om* itself creates a predicate with the value 'PRO' in f-structure and this is incompatible with an overt pronoun.

By contrast, overt pronouns do occur in the NCOMP function, in which case the pronoun is focussed: *az-én lány-om* (GEN-I daughter-NCOMP[1SG] = 'MY daughter'. In this case, *lány-om* must be lexically specified as follows:

[14] *lány-om*: ↑PRED = 'daughter <↑ NCOMP>
↑NUM = SG
↑NCOMP NUCL = –
↑NCOMP PERS = 1
↑NCOMP NUM = SG

Here the absence of the annotation ↑NCOMP PRED = 'PRO' will require the form *lány-om* to co-occur with an NCOMP whose predicate value is independently specified, and the remaining annotations will force this to be a first person singular non-reflexive pronoun.

The distribution of prodrop and overt pronoun forms here is not accidental. Prodrop is naturally compatible with a topic discourse-orientation of the subject function, while, in Hungarian at least, focus is not. Focussing through the use of stressed pronominal forms is compatible with any non-subject argument function, including NCOMP.

5 Semantic roles

The formation of the possessum of an ordinary, non-relational noun is a lexical process:

[15] macska: N, $\uparrow$PRED = 'cat', $\lambda y[cat(y)]$

macská-ja: N, $\uparrow$PRED = 'cat <'$\uparrow$NCOMP>', $\lambda x \lambda y[\pi(x,y)\ \&\ cat(y)]$
$\uparrow$NCOMP PERS = 3

Semantically, this process forms a 2-place predicate from a 1-place predicate, introducing an arbitrary relation π from the set of possible relations Π between possessor x and possessum y Barker (1995). The new two place predicate will combine with a suitable argument (e.g. *Mari*) to form a construction (e.g. *Mari macská-ja)* which again has the semantics of a 1-place predicate. In the case of relational nouns, e.g. *lány* 'daughter', we can think of the relation involved directly as a member of Π: *lány-a* translates as $\lambda x \lambda y[daughter(x,y)]$. The lexical mapping rules then simply associate a [–r, –o] function with the argument x, and this can freely map either onto SUBJ or onto NCOMP for most relations π.

If we consider the set of possible relations Π, abstracting away from the syntactic and pragmatic differences between the SUBJ and NCOMP constructions in English and Hungarian, the range of semantic relations permitted by each construction is strikingly similar in the two languages. Natural subgroups for the SUBJ function are (adapted from Payne & Huddleston (to appear, 2002)):

[16]

English SUBJ	Hungarian SUBJ	Relation[5]	
Mary's back	Mari-nak a hát-a		BODYPART
Mary's older sister Mary's boss Mary's friend	Mari-nak a nővér-e Mari-nak a főnök-e Mari-nak a barát-ja	*KIN* *SUPERIOR* *EQUAL*	RELATION
Mary's team	Mari-nak a csapat-a		MEMBER
Mary's debut Mary's book	Mari-nak a bemutatkozás-a Mari-nak a könyv-e	*PERFORMER* *CREATOR*	AGENT
Mary's new house	Mari-nak az új ház-a		OWNER
Mary's honour	Mari-nak a becsület-e		CHARACTER
Mary's anger	Mari-nak a harag-ja		EXPERIENCER
Mary's letter	Mari-nak a level-e		RECIPIENT
Mary's biography Mary's surgery	Mari-nak az életrajz-a Mari-nak a rendelő-je	*DESCRIPTUM* *UNDERGOER*	THEME
the sun's rays the conflict's origin	a nap-nak a sugar-ai a viszály-nak a forrás-a	*EMITTER* *ORIGIN*	SOURCE
the flood's consequence	az árvíz-nek a következmény-e		RESULT
the cathedral's spire	a templom-nak a torny-a		PART

All these types of relation are equally possible for the NCOMP function, straightforwardly in Hungarian, and with sufficiently heavy NPs in English, e.g.

[17]	a.	the green eyes of the girl sitting opposite me	BODYPART
	b.	the sister ot the man who had been arrested	RELATION
		the debut of the young flautist from Abergavenny	AGENT
	c.	the new house of the Vice-Chancellor elect	OWNER

The sheer range of semantic relations involved, encompassing for example both agent and theme, leads us to treat the NCOMP function as an unrestricted function.[6] What is more, there are some further relations which in both languages are available to the NCOMP function but ungrammatical in the SUBJ function. These include relations of quality and apposition (for further possible examples see Chisarik 1999):

[18]	a.	a	boldogság	perc-ei	
		ART	happiness.GEN	minute-PL.NCOMP [3]	
		'the minutes of happiness'			
	b.	*a	boldogság-nak	a	perc-ei
		art	happiness-dat	ART	minute-PL.SUBJ[3]
[19]	a.	Budapest	város-a		
		Budapest	city-NCOMP[3]		
		'the city of Budapest'			
	b.	*Budapest-nek	a	város-a	
		Budapest-DAT	ART	city-SUBJ[3]	

The situation is analogous to that in English, where similar relationships are excluded from the genitive construction: **happiness's minutes*/**Budapest's city*.

The SUBJ relations in both languages are therefore a subset of the NCOMP relations in both languages, and the two functions, although both unrestricted, are not semantically equivalent. It remains of course an open question whether this pattern has any universal validity.[7]

6 Multiple possessors

The SUBJ and NCOMP functions are distinct, and therefore in principle not incompatible. Semantically, it is possible to iterate the function which forms two-place relations, giving possessum nouns with both functions. In English, this process occurs quite straightforwardly:

[20] photo: N, $\uparrow$PRED = 'photo',
$\lambda y[\mathit{photo}(y)]$

photo: N, $\uparrow$PRED = 'photo <'$\uparrow$NCOMP>',
$\lambda x[\lambda y[\pi(x,y)\ \&\ \mathit{photo}(y)]]$

photo: N $\uparrow$PRED = 'photo <$\uparrow$SUBJ, $\uparrow$NCOMP>',
$\lambda x[\lambda z[\lambda y[\pi_2(z,y)\ \&\ \pi_1(x,y)\ \&\ \mathit{photo}(y)]]]$

In the case of the noun *photo*, the relation π_2 will typically be OWNER or CREATOR, and π_1 will be DEPICTUM:

[21] Mary's photo of Bill

The lexical mapping rules in English associate a [–r, –o] function both with the z argument (NCOMP) and the x argument (SUBJ). Clearly there is a hierarchical correspondence between the order in which the arguments are combined (π_2 higher than π_1) and some version of the thematic hierarchy. The thematic hierarchy for argument mapping in clause structure is assumed to be as in [22] (Bresnan 2000: 307):

[22] *agent* > *beneficiary* > *experiencer*/*goal* > *instrument* > *patient*/*theme* > *locative*

From [21] we have OWNER < DEPICTUM (a subspecies of theme), and AGENT < DEPICTUM. A full analysis of multiple possessor constructions in English and other languages which permit SUBJ and NCOMP to co-occur will therefore require that the hierarchy be augmented with typical NP roles such as OWNER.

Hungarian now presents an interesting puzzle: it is impossible to have both a genitive and a dative possessor in the same noun phrase. As shown by Laczkó (1995, 2000), event nominalizations require the theme argument to be expressed as either a dative or a genitive possessor, but the agent must be treated as an adjectivalised postpositional modifier:

[23]	a.	a	váza	Edit	által-i	összetör-és-e
		ART	vase.GEN	Edith	by-ADJ	smash-NOM-NCOMP[3]
		‘the smashing of the vase by Edith’				
	b.	*Edit-nek	a	váza		összetör-és-e
		Edith-DAT	ART	vase.GEN		smash-NOM-NCOMP[3]

There is therefore no mapping in Hungarian equivalent to the mapping involved in the English *Edith’s smashing of the vase*.

On the basis of their incompatibility and similarity in function, analyses in the Government and Binding framework (Szabolcsi 1994, É.Kiss 1999 etc.) treat the two possessor constructions in Hungarian as related by movement. The LFG analysis of Laczkó (1995, 2000) naturally eschews this movement analysis, but treats the two possessor positions as functionally equivalent and subsumed under a single function POSS: it is then natural to account for the incompatibilty of the two types of possessor using the principle of coherence (e.g. Bresnan 2000). However, we have argued that SUBJ and NCOMP are distinct functions, and coherence cannot therefore be used to account for this language-specific incompatibility. The difference between English and Hungarian requires an alternative explanation.

Two possible solutions suggest themselves: Firstly, we might assume that there is a paradigm gap: no lexical form can realise a lexical entry annotated ↑PRED = ‘<↑SUBJ, ↑NCOMP>‘. This however seems seems arbitrary, since such lexical entries are required for the multiple possessor constructions of English. Nor does it seem plausible to relate such a paradigm gap to agreement: there seems to be no reason why a form which agreed both with SUBJ and NCOMP would be needed. After all, in transitive clauses it is typical for a verb to agree with SUBJ and not OBJ.

The second solution, which we prefer, is to suggest that two [–r] arguments are blocked in Hungarian noun phrases by an extension of the asymmetrical object parameter of Bresnan and Moshi (1990). We will call this the asymmetrical possessor parameter to reflect its applicability to noun phrases:

[24] *Asymmetrical Possessor Parameter*

$$\begin{array}{cc} {*}\theta & \theta \\ | & | \\ -r & -r \end{array}$$

In the case of nominalisation, all unrestricted arguments automatically acquire the value [–o]. The nominalisation of a verb with a [–o] agent and a [–r] theme will then have its agent mapped in noun phrase structure to either [–o, –r] or [–o, +r], and its theme to [–o, –r]. In a language such as Hungarian which does not permit two [–o, –r] arguments, there is no alternative but to express the theme as SUBJ or NCOMP, and to treat the agent as an oblique. In English, which does permit two [–o, –r] arguments, the higher role will map to SUBJ and the lower to NCOMP.

7 Conclusion

In this paper, we have argued that the existence of structurally distinct possessor constructions in languages such as English and Hungarian necessitates the postulation of more than one possessor function. We identify the structurally higher position of the possessor with the SUBJ function and postulate a new function NCOMP for the structurally lower position. The SUBJ function has topic-like discourse properties, whereas NCOMP does not. We also argue that NCOMP is an unrestricted function, since the range of semantic roles it encompasses in both languages is at least as broad as that of SUBJ. English and Hungarian differ in that SUBJ and NCOMP can co-occur in English, but are mutually exclusive in Hungarian. We account for this by an asymmetrical possessor parameter, an extension to noun phrases of the asymmetrical object parameter. When SUBJ and NCOMP co-occur, as in English, the superordinate nature of SUBJ is, it seems, reflected in the operation of a thematic hierarchy.

The analysis we have given is of course a fragment based on English and Hungarian. It remains a subject for further research to see how the analysis might apply to other languages with multiple possessor constructions, and indeed to languages with a single possessor construction. We have treated SUBJ and NCOMP as unrestricted, but the evidence of the *a friend of Mary's* construction in English hints that the possibility of restricted possessor constructions should not be excluded.

Notes

[1] In [2] we anticipate the assignment of the dative possessor to the SUBJ function and the so-called nominative possessor to the NCOMP function.

[2] Note that the determiner *egyik* 'a/one' in example [4] is not the same as the indefinite article *egy* 'a', though it is clearly related to it morphologically. The genitive possessor is in complementary distribution with the indefinite article as it is with the definite article.

[3] While the agreement features of person and number naturally feed into the f-structure representation, we have made the decision that the case of the possessor NP is an arbitrary property of each construction and itself makes no contribution to f-structure. Some feature percolation mechanism is then needed to ensure that the syntactic case of the possessor NP matches the case of the head noun where this case has its exponence, and there are no obvious mechanisms within the existing LFG framework for achieving this. One possible move is to employ m-structure representations along the lines of Butt et al (1996), and this is what lies behind the representation of case in [9a] and [9b]. We employ this representation simply to avoid the presence of case in f-structure: whether the use of m-structures is ultimately the appropriate mechanism for achieving this is a question we will leave open.

[4] Although *nekem a lány-om* (I-DAT ART daughter-SUBJ[1SG]) with a dative pronoun is considered ungrammatical by most speakers, some speakers do use it. For these speakers, the annotation (↑NCOMP PRED) = 'PRO' in [13] is optional.

[5] In table [16], *PERFORMER* and *CREATOR* are taken to be subspecies of the more general AGENT relation, and so on.

[6] Unrestricted functions naturally permit expletive elements which have no semantic value (Bresnan 2000: 308). It is well-known that expletives are excluded from English NPs: compare the expletive *it* in *It is thought that the universe will expand forever* with **its thought that the universe will expand forever*. In Hungarian, as pointed out by Kenesei (1994:319) expletives are permitted: *annak a gondolat-a, hogy a fogorvos-hoz kell mennie* (that.DAT ART thought-SUBJ[3], that ART dentist-ALL must go) 'the thought that he had to go to the dentist'. The expletive is the distal demonstrative, and since demonstratives in general lack genitive case, expletives only occur in the SUBJ function.

[7] We do not of course exclude the possibility that further functions beyond SUBJ and NCOMP may be needed to capture the full range of possessor constructions in the world's languages. English for example has a third possessor construction, the genitive oblique construction illustrated by NPs like *a friend of Mary's*. The range of semantic relations permitted by this construction is somewhat restricted compared with those permitted by the SUBJ and NCOMP constructions (it basically encompasses RELATION, OWNER and AGENT), and therefore the use of restricted oblique functions might indeed be more appropriate in this case.

References

Barker, Chris 1995. *Possessive Descriptions.* Stanford: CSLI Publications.

Bresnan, Joan 2000. *Lexical-Functional Syntax.* Oxford: Oxford University Press.

Bresnan, Joan & Lioba Moshi 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21, 147-185.

Butt, Miriam, M.E.Niño & F.Segond 1996. Multilingual processing of auxiliaries in LFG. In D.Gibbon (ed) *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference.* Bielefeld, 111-122.

Chisarik, Erika 1999. *The Syntactic Structure of Hungarian and English Possessive Noun Phrases: a Comparative Study.* Doctoraalscriptie, Leiden University, The Netherlands.

Falk, Yehuda 2001. *Lexical-Functional Grammar: An Introduction to Constraint-Based Syntax.* Stanford: CSLI Publications.

Jucker, Andreas H 1993. The genitive versus the of-construction in newspaper language. *Anglistik und Englischunterricht* 49, 121-136.

Katalin.É.Kiss 1999. The Hungarian noun phrase is like the English noun phrase. In Gábor Alberti & István Kenesei (eds) *Approaches to Hungarian, Volume 7*: Papers from the Pécs Conference, 119-149. Szeged: JATE Press.

Kenesei, István 1994. Subordinate clauses. In Ferenc Kiefer & Katalin.É.Kiss (eds) *Syntax and Semantics 27: The Syntactic Structure of Hungarian*, 275-354. San Diego: Academic Press.

Laczkó, Tibor 1995. *The Syntax of Hungarian Noun Phrases: a Lexical Functional Approach*. Frankfurt am Main: Peter Lang.

Laczkó, Tibor 2000. Derived nominals, possessors, and lexical mapping theory in Hungarian DPs. In Miriam Butt & Tracy Holloway King (eds) *Argument Realization*, 189-227. Stanford: CSLI Publications.

Payne, John & Erika Chisarik 2001. The so-called “nominative” possessor construction: a new genitive? Paper presented at the 5th International Conference on the Structure of Hungarian, Budapest 24-26 May 2001.

Payne, John & Rodney Huddleston (to appear, 2002). Nouns and noun phrases. In Rodney Huddleston & Geoffrey K.Pullum The Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.

Rappaport, Malka 1983. On the nature of derived nominals. In Lori Levin, Malka Rappaport & Annie Zaenen (eds) *Papers in LFG,* 113-142. Bloomington: Indiana University Linguistic Club.

Sadler, Louisa 2000. Noun phrase structure in Welsh. In Miriam Butt & Tracy Holloway King (eds) *Argument Realization*, 73-109. Stanford: CSLI Publications.

Szabolcsi, Anna 1994. The noun phrase. In Ferenc Kiefer & Katalin.É.Kiss (eds) *Syntax and Semantics 27: The Syntactic Structure of Hungarian*, 179-274. San Diego: Academic Press.

XLFG-an LFG parsing scheme for French

∗Lionel Clément, ⋄Alexandra Kinyon
∗UFRL - University of Paris 7 ⋄IRCS - University of Pennsylvania
Paris, FRANCE Philadelphia, USA
clement@linguist.jussieu.fr kinyon@linc.cis.upenn.edu
http://talana.linguist.jussieu.fr/˜lionel
http://www.cis.upenn.edu/˜kinyon

Proceeding of the LFG01 Conference
University of Hong Kong, Hong Kong
Miriam Butt and Tracy Holloway King (Editors)
2001
CSLI Publications
http://csli-publications.standford.edu/

Abstract

We present XLFG, a new parser for LFG developed by Lionel Clément (Clément 2000), (Clément 2001b). This parser is freely available for pedagogical as well as for research purpose. In this paper, we explain how the tool has been used to develop and evaluate a medium-size LFG for French. We also show how we have exploited the similarities between the LFG and the Tree Adjoining Grammar Frameworks in order to integrate into XLFG a rule-based parse-ranker for French.

Introduction

Several parsers have already been implemented for Lexical Functional Grammars (LFG), for example (Kaplan and Maxwell 1994), (Andrews 1990). Some of these parsers have even been developed for French (Zweigenbaum 1991) (Vappillon 1997). However, these tools are sometimes not publicly available, or not flexible enough to implement additions to the "base" LFG formalism (e.g. the source code may not be available). This is why we have decided to implement XLFG, a new parser for LFG. Our goal was to provide a tool that is user-friendly enough to be used in a pedagogical environment, reasonably fast, portable to most Operating Systems, and flexible enough so that additions could be made to the base LFG framework in order to test new linguistic hypothesis. In the first part of this paper, we provide an overview of XLFG, we explain how XLFG was used to develop an LFG medium-size grammar for French, and how it can be used for pedagogical purpose. In a second part, we explain how XLFG has been used to investigate the relationship between the Tree-Adjoining Grammars (TAG) framework, and the LFG framework. More specifically, we exploit the similarities between TAG "derivation trees" and LFG "functional structures" to incorporate into XLFG a parse-ranker for disambiguation purpose. Finally, we present a preliminary evaluation of the tool on the TSNLP test-suite for French. We assume the reader is familiar with the LFG framework and refer to (Bresnan 1982), (Kaplan 1989) for an introduction.

1 Overview of XLFG

The XLFG project started in 1996. XLFG, a wide-coverage parser for lexical functional grammars (LFG), is written in C, using an LR algorithm (Knuth 1965). We have tried to put an emphasis on the user-friendliness, thanks to a Graphical User Interface (GUI) developed in Tcl/Tk and to provide a fast and ergonomical way to encode and test LFG grammars (Clément 2000). XLFG is highly portable and is publicly available for academic and research purpose. Both the program and the source code can be downloaded on the www for most platforms[1].
The input of the parser consists in 3 files:

- A template file
- A lexicon file
- A rule file

For each sentence parsed, the output of the parser consists in:

- One or several couples of constituent and functional structure(s)

Moreover, XLFG can be used both online and offline. When used online, a constituent structure and an functional structure are displayed in the GUI[2]. One can see on figure 1 an example of output when parsing the sentence *Le chien aboie (The dog barks)*.

When used offline, one can obtain different types of output (e.g. a Latex file), which may be further postprocessed (for example to create figures, to input to a parse-ranker, to build a Treebank ...).

At this point, the following items have been implemented:

- Computation of constituent and functional structures
- Functional descriptions

[1] http//talana.linguist.jussieu.fr/˜lionel/xlfg/
[2] When a sentence is ambiguous, several constituent and functional structures are then displayed.

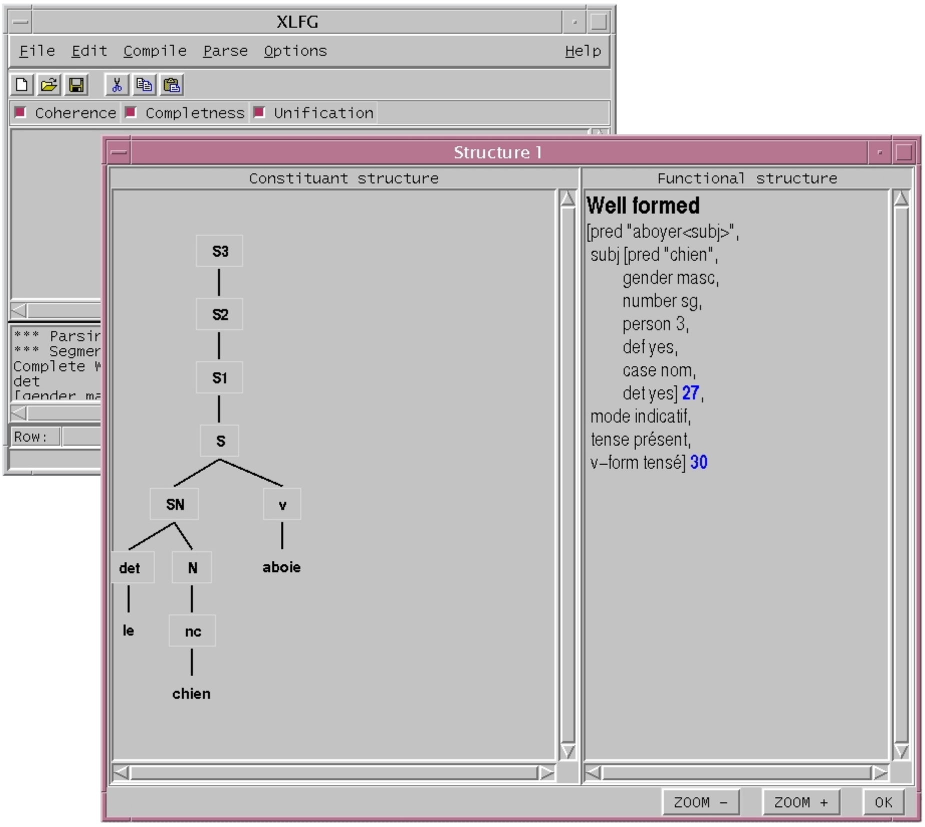


Figure 1: Sample output *Le chien aboie (the dog barks).*

- unification (with "re-entrance")
- defining equations
- constraining equations
- set of functional structures
- existential constraint

- Lexical rules
- Functional uncertainty

The goal of this work twofold:

- XLFG may be used as a pedagogical tool
- XLFG can easily be augmented for research purpose

1.1 XLFG: a pedagogical tool

Since XLFG can be used as a pedagogical tool to learn or teach the LFG framework, it was important to scrupulously follow the base LFG model as it was defined in the 80's in (Bresnan 1982), (Kaplan 1989).

XLFG can be used to teach the LFG framework to linguistics students "hands-on".

It has been used at the University of Paris 7 to teach LFG to MA students in computational linguistics. It has also been used at the University of Provence in an introductory course to syntactic formalisms. XLFG may be used during lab sessions (often on Unix machines), but also for homework assignments after being downloaded by students (who often use Microsoft Windows). Therefore, portability was important.

In addition, it is possible for lecturers to establish an online connexion using a client-server model. The server provides a large, robust wide-coverage lexicon. Thus, the lecturer and the students can focus on the syntactic phenomena under investigation without having to deal with lexical coverage issues.

To emphasize the pedagogical aspect, XLFG also parses ungrammatical sentences, as long as a "good" constituent structure is found (i.e. sentences with a good constituent structure, but possibly an ill-formed functional structure). It then expicitely shows in the output why the functional structure is ill-formed (i.e. uniqueness, coherence and/or completeness, as well as unification problems). Some examples are shown on figures 2,3 and 4).

In addition, it is possible to relax conditions on constraint equations as well as on existential equations.

1.2 Application to French

XLFG was used to develop and test a grammar for French, similarly to what has been achieved by (Abeillé et al. 2000a) for Tree Adjoining Grammars. The following syntactic phenomena are currently handled: clitics, passive, sentential complements, infinitival complements (control and raising verbs), some idiomatic expressions, some comparatives and coordinations (including gapping phenomena), compound tenses, negation (both for finite and infinitival clauses), some long distance dependencies.

Our grammar, which is hand-crafted, currently comprises 78 syntactic rules. Our lexicon currently comprises 450,000 inflected forms.

The following examples show how the lexicon is encoded (inflected forms) some syntactic rules for XLFG. The inflected forms of the lexicon are computed from morphological tables and lexical properties on lemmas (such as sub-categorization, lexical transformations, functional descriptions of raising verbs, etc.).

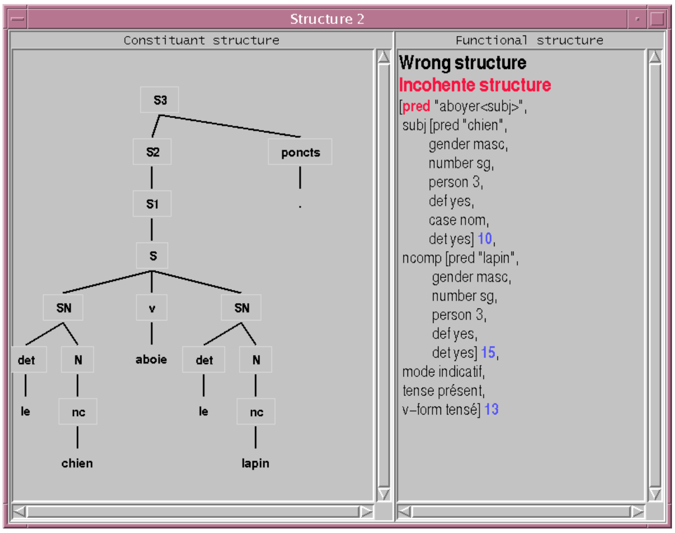


Figure 2: An incoherent structure **Le chien aboie le lapin (*The dog barks the rabbit).*

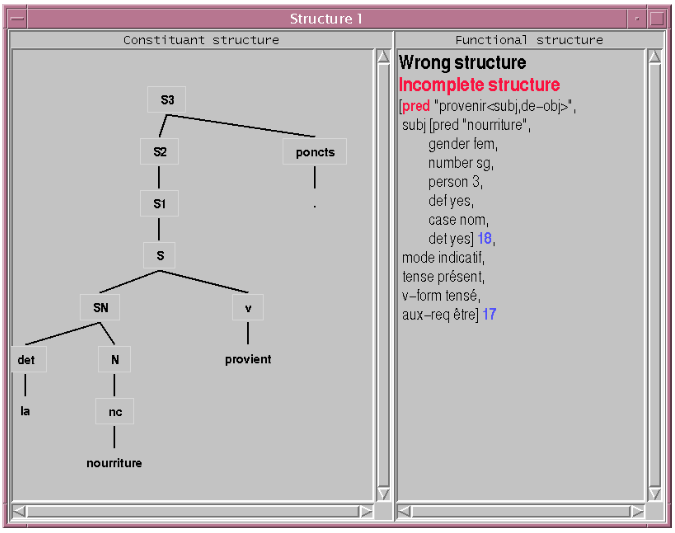


Figure 3: An incomplete structure **La nourriture provient (*The food comes from).*

Extract of the French lexicon (inflected forms)[3]

```
saisonné v [pred="saisonner<subj>",@Kms,@avoir,@active]  ;
```

[3] In features, @Kms, @Kfs, etc. are abbreviations. For agreement features, @Kms stands for "v-form=past-p, gender=masc, number=sg"; For auxiliary verbs, @avoir stands for "aux-req = avoir,

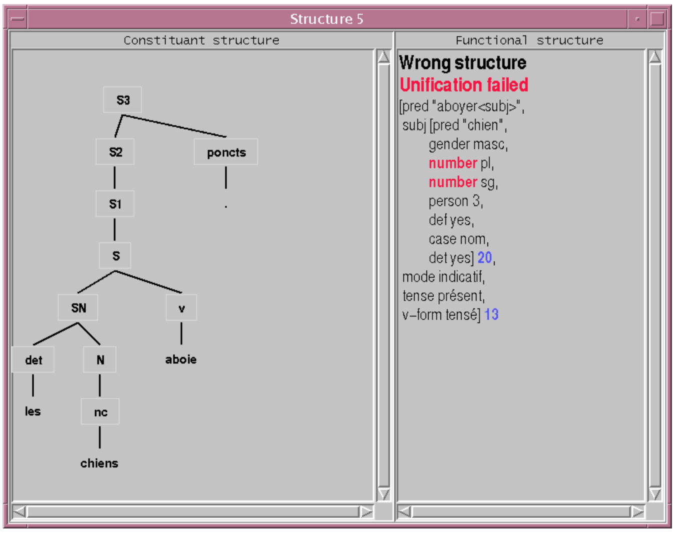


Figure 4: A structure where unification fails because of agreement **Les chiens aboie (*The dogs barks).*

```
saisonnée v [pred="saisonner<subj>",@Kfs,@avoir,@active]  ;
saisonnées v [pred="saisonner<subj>",@Kfp,@avoir,@active]  ;
saisonnés v [pred="saisonner<subj>",@Kmp,@avoir,@active]  ;
saisons nc [pred="saison",@fp] ;
saisîmes v [pred="saisir<subj,(obj)>",@Jp1]  ;
saisîmes v [pred="saisir<subj,(de-obj)>obj",@Jp1] @Pron ;
saisît v [pred="saisir<subj,obj>",@Ts3]  ;
saisît v [pred="saisir<subj,(de-obj)>obj",@Ts3] @Pron ;
saisîtes v [pred="saisir<subj,obj>",@Jp2]  ;
saisîtes v [pred="saisir<subj,(de-obj)>obj",@Jp2] @Pron ;
sait v [pred="savoir<subj,(obj/comp/vcomp)>",@Ps3] @CompInd @Ctrl-
Subj ;
saké nc [pred="saké",@ms] ;
sakés nc [pred="saké",@mp] ;
sala v [pred="saler<subj,obj>",@Js3]  ;
salace a [pred="salace<subj>",@s]  ;
salaces a [pred="salace<subj>",@p]  ;
```

For non feature abbreviations, @Pron stands for the functional description: "(↑ obj case =c rfl)", @CtrlSubj for "(↑ vcomp subj = ↑ subj)" for subject raising of control verbs.

Extract of the rules for a French grammar[4]**.**

```
// sample: Wh extraction
// A quelle fille Jean pense-t-il˜? (Which girl does J think about)
S3 -> SP S2 (poncts);
($$ top = $1)
($1 qu =c +)
($$ = $2)
($$ (($1 pcas)-obj) = $1)
($$ = $3)
(($$ mode) = indicative/conditional);

// sample: ‘‘surcomposed past’’
// Jean a eu déjeuné (J has had luch)
// Jean n’a pas eu moins déjeuné
_VERB -> (_CLITIC) aux (claff) (cln) (advneg) aux (SADV) v;
($$ tense = $2 tense)
($$ mode = $2 mode)
($$ subj number = $2 number)
($$ subj person = $2 person)
($2 form-aux = avoir)
($$ subj = $2 subj)
($$ = $3)
($$ subj = $4)
($5 neg = $$ neg)
($5 < ($$ adjunct))
($6 form-aux = avoir)
($6 v-form = past-p)
($$ = $8)
($$ aux-req = $6 form-aux)
($$ aspect = perfect);
```

The previous two syntactic rules show how one can encode rewriting rules as well as functional descriptions.

Although the goal of XLFG is not to develop a robust and fast parser, but rather to develop LFG grammars and test linguistic hypothesis in a user-friendly and portable environment, the system is nonetheless fast. For instance, a sentence such as *L’europe espère que le marché national améliorera l’influence de la Commission. (Europe hopes that the national market will improve the influence of the commission)* was parsed in 2 seconds on a Super Sparc, yielding 73 analyses.

[4] In this rules, $$ is the notation for $\uparrow$, and $\$n$ for $\downarrow$ of the n^{th} term. $<$ is the notation for $\in$. We mark $(\uparrow x = y)$ for $(\uparrow x) = y$

As always when parsing, a trade-off must be found between trying to be wide-coverage and obtaining an acceptable syntactic ambiguity rate.

To increase the coverage, we are currently adding more rules to the grammar, as well as more entries to the lexicon. Moreover, to deal with certain specific syntactic phenomena such as ellipsis (e.g. *Jean mange des pommes et Marie des cerises (John eats apples and M. cherries)*, *Jean est plombier et fier de l'être (John is a plumber and proud of it)*), we have augmented the LFG formalism with the notion of *lexical capture*, and have implemented this feature into XLFG. Due to space limitations, we do not develop this notion here but refer to (Clément 1996) and future publications for more details.

To limit the ambiguity rate, we needed to integrate a disambiguation strategy into XLFG. We chose to depart from methods that have been traditionally used for LFG, such as Optimality Theory (Frank et al. 2000), or solutions based on probabily models (Riezler et al. 2000). Instead, we have observed that LFG functional structures and TAG so-called "derivation trees" are intuitively very similar (modulo re-entrant features)[5].

This point is developed in the next section.

2 Adapting TAG based disambiguation principles to LFG

2.1 Why not use a probabilistic disambiguation model ?

In order to disambiguate (LFG parses, as well as parses from other frameworks), one can resort to probabilistic models. However, we chose not to, both for theoretical as well as for practical reasons.

- From a practical point of view, probabilistic models are not language nor domain independent, they are costly because one needs large training data (i.e. treebank). Unfortunately, such training data is not available for most languages other than English. For French, there is no treebank to train on (although one is currently being built by (Abeillé et al. 2000b), (Clément 2001a)). Therefore, we could not resort to a probabilistic disambiguation model in order to disambiguate.

[5] It is also interesting to note that both functional structures and derivation trees are in turn quite close the notion of dependency tree, but we don't develop this point here.

- From a theoretical point of view, probabilistic models have no "explanatory power"(they may work, but this does not say anything about why it works). In human sentence processing, the importance of Lexical preferences is widely accepted (e.g. (Trueswell 1996)); For instance, a given verb may prefer to attach a PP as an argument. But there is little data available regardings these human preferences. And finding frequency effects in language comprehension does not automatically condemn all structural approaches: If one considers the sentence *John put the book that you were reading in the library*, although *put NP1 in NP2* is obviously a common subcategorization frame for *put*, the sentence nonetheless seems incomplete, although it is syntactically well-formed.

Therefore, since we were not convinced that probabilistic disambiguation methods were sufficient from a theoretical point of view, and since we did not have any data to train a probabilistic disambiguation model on, we had to turn to rule-based methods[6].

2.2 Why not express preference rules on C-structures ?

Traditionnally, rule-based disambiguation principles have been formulated on the shape of constituent trees. For instance (Kimball 1973) formulated the right association principle, which allows to retrieve the correct attachment in a sentence such as *Tom said that Joe left yesterday* (were yesterday attaches to left rather than to Tom). Similarly, (Frazier and Fodor 1978) formulated a minimal attachement principle which states that in case where several constituent structures are associated to a given sentence, the constituent tree with the fewest number of nodes should be favored. For a sentence such as *Tom bought the flowers for Sue*, this principle allows one to retrieve the preferred attachment, where *for Sue* attaches to *bought* rather than to *flowers*. However, as argued in (Kinyon 1999) this type of principles, formulated on constituent trees were deemed unsatisfactory for several reasons:

[6] We do not of course reject the idea of lexical preference. In fact, in addition to structural preferences (see infra), we have added some weights to lexical items to model some lexical preferences.

- Principles such as Minimal Attachment make some assumption about how a constituent tree should be built[7].
- The interaction between these principles is unclear.
- It is difficult to integrate semantic and/or pragmatic information into such principles.
- These principles do not establish a distinction between argument and modifier attachment.
- Counter-examples are easy to find: e.g. the validity of right association is debated for Spanish (Cuetos and Mitchell 1988) and Dutch (Brysbaert and Mitchell 1996).

Therefore, we did not want to express disambiguation principles in terms of constituent structure.

On the other hand, some widely accepted disambiguation principles may not be expressed on constituent tree, but are easy to express on "dependency-like" structures:

- Prefer the idiomatic interpretation of a sentence over its literal interpretation.
- Prefer to attach an element as an argument rather than as a modifier.
- Prefer to attach an argument to its closest potential governor.

Principle 1 allows the idiomatic interpretation of a sentence such as *John breaks the ice* to be preferred over its literal interpretation.
Principle 2 allows one to prefer to attach *to be honest* as argument of *prefer*, rather than as a sentence modifier in a sentence such as *John prefers his daughter to be honest*[8].
Principle 3 allows to prefer to attach *of the demonstration* as argument of *organizer*, rather than as an argument of *suspect* in a sentence such as *John suspects the organizer of the demonstration*.

These disambiguation principles were succesfully used to implement a parse-ranker for French (Kinyon 2000) for TAGs, expressed on the number of nodes in TAG "derivation trees", and thus we wanted to see if this was transposable to LFG, and expressable on "f-structures".

[7]Most of these principles were formulated within an X-bar type of linguistic framework. And principles such as Minimal attachment rely for a large part on some X-bar characteristic concerning the number of nodes a given constituent tree will have.

[8]The modifier reading being identical to *To be honest, John prefers his daughter*.

2.3 Similarities between TAG and LFG

The similarity between LFG and TAG was to the best of our knowledge first investigated in (Kameyama 1986). Surprisingly, very little has been written on this topic since. We will not provide an introduction to TAGs here (and refer to (Joshi 1994) and (Abeillé and Rambow 2000) for an overview). For our purpose, one just needs to know that in TAG, the elementary structures, called "elementary" trees, combine via 2 operations: substitution (for "initial" elementary trees) and adjunction (for "auxiliary" elementary trees). Figure 5 shows how one can parse the sentence *Jean ne voit pas Marie (John does not see Mary)* using a Lexicalized TAG consisting in four elementary trees[9]. In this example, the initial trees "α-Jean" and "α-Marie" were substituted respectively as subject and object in the initial tree "α-voit", and the auxiliary tree "β-ne-pas" was adjoined as a modifier into the initial tree "α-voit". The result one obtains from combining the trees consists in:

- A **derived** tree, which is a constituent structure.
- A **derivation** tree, close to a dependency structure, which is built by keeping the historic of operations (i.e. substitution and adjunction) performed to obtain the derived tree.

The important thing to note is that both LFG and TAG yield a constituent structure (categorial structure for LFG, derived tree for TAG) and a *dependency-like structure* (functional structure for LFG, derivation tree for TAG)[10]. So the derivation tree from figure 6(a) can be seen as equivalent to the F-structure on figure 6(b). This similarity between TAG derivation trees and LFG functional structure does not of course extend to the case of reentrant features, which turn f-structures into a graph rather than a tree, but reentrant features are in practice limited to bounded phenomena such as control[11].

[9] Agrement is dealth with by associating features to nodes in elementary trees (not shown on figure for readability). α denote inital trees, β auxiliary trees, plain lines represent an adjunction (i.e. the insertion of an tree into another tree), dotted line a substitution.

[10] Other similarities exist, for instance (Rambow 1996) noted the similarity between *d-links* used in some formalisms related to TAGs, and the LFG notion of functional uncertainty but this is beyond the scope of our work.

[11] So turning LFG functional structures into TAG derivation tree may imply a "loss" of information, but this information can be encoded using co-indices on derivation tree leaves.

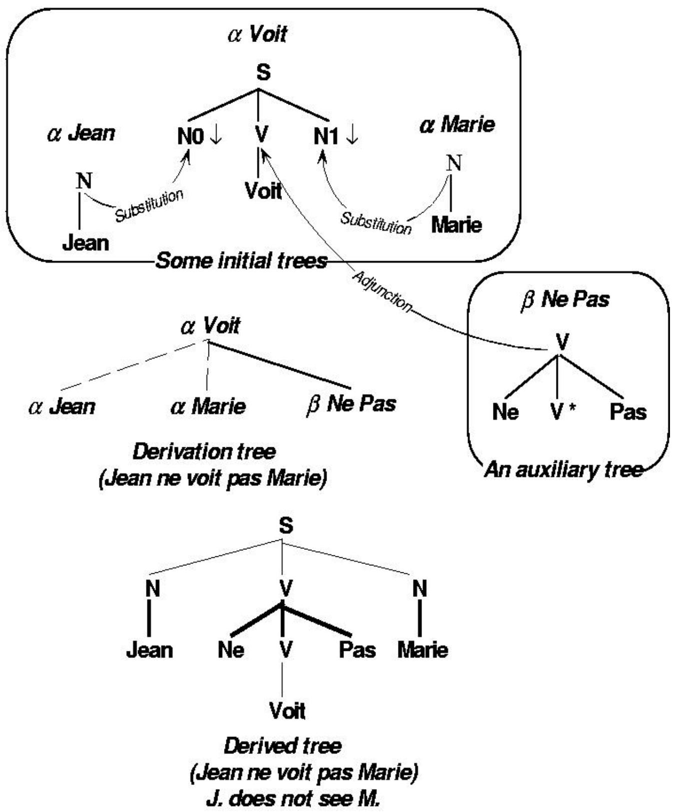
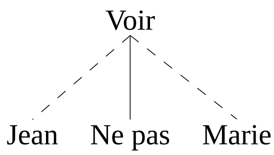


Figure 5: Jean ne voit pas Marie (*John doesn't see Mary.*).

2.4 Expressing disambiguation principles in terms of f-structure

Principle 1, which says that the idiomatic interpretation of a sentences should be favored over its literal interpretation, was modeled in TAGs by preferring the derivation tree (i.e. the dependency-like structure) with the fewest number of nodes. To implement Principle 1 into XLFG, we only had to make sure that the functional structure which resorts to the most constrained lexical item would be favored (here *break1* which constrains its object to be *ice*, rather than *break2* which is unconstrained for its object) (cf fig 7).

Principle 2, which says that the attachment of an element as an argument rather than as a modifier should be preferred was expressed on TAG derivation trees by preferring the derivation trees with the fewest



a

$$\begin{bmatrix} \text{pred} & \text{‘voir}\langle\text{subj, obj}\rangle\text{’} \\ \text{subj} & \begin{bmatrix} \text{pred} & \text{‘Jean’} \\ & \dots \end{bmatrix} \\ \text{obj} & \begin{bmatrix} \text{pred} & \text{‘Marie’} \\ & \dots \end{bmatrix} \\ \text{adjunct} & \left\{ \begin{bmatrix} \text{pred} & \text{‘pas’} \\ & \dots \end{bmatrix} \right\} \end{bmatrix}$$

b

Figure 6: Derivation tree and equivalent F-structure for *Jean ne voit pas Marie (John doesn't see Mary)*.

number of β-tree nodes[12]. This principle was also easy to encode on LFG functional structure, simply by counting the number of local arguments functions of a predicator (i.e. *subj*, *obj*, *à-obj*, *de-obj*, ...), and favoring the functional structure with the highest count (cf fig. 8).

Finally, Principle 3 which states that one should prefer to attach arguments as close as possible to their governor, was implemented in TAGs by computing on each derivation tree the sum of the distances between each node in the tree and their son. This allows one to prefer *to Paul* as an argument of *give* rather than as an argument of *says* in example 1a, since the distance between *to Paul* and *give* is shorter than the distance between *to Paul* and *say*. Whereas,conversely, in sentence 1b, *to whom* will be argument of say, rather than argument of *give*, for the same reason.

(1) a. John says that Mary gives flowers to Paul.
    b. To whom does John say that Mary gives flowers.

This principle was adapted to LFG by computing the sum on the functional structures, seen as dependency-like structures: For each node in a given structure, one computes the sum of the distances between the predicator and all of its daughters, according the linear order of

[12]TAG β trees (i.e. auxiliary trees), are mainly used to encode modifiers, whereas TAG α trees (i.e. initial trees), are mainly used to encode arguments.

the sentence[13]. The functional structure with the lowest sum will be favored (cf fig. 9).

As was done for TAGs, the 3 principles are applied in a sequential order. Thus there are no conflicts. The parser computes all possible functional structures for a given sentence. It then applies the 3 principles and outputs the n-best functional structures according to these principles, along with their corresponding constituent structure.

In addition, we do not commit exclusively to a rule-based disambiguation system: since a weight is associated to each lexical entry in the lexicon, this allows us, in addition of favoring the idiomatic interpretation of a sentence, to encode lexical or semantic preferences as well, thus allowing to mix rule-based general structural preference principles with more fine-grained statistical lexical preferences in order to obtain the best possible disambiguation. For instance, this may be used to prefer “function words” over“lexical word” (e.g. to prefer *être (to be), avoir(to have)* as auxiliaries rather than full lexical verbs.).

$$\begin{bmatrix} \text{subj} & \begin{bmatrix} \text{pred} & \text{‘John’} \\ \text{num} & \text{sg} \\ \text{pers} & 3 \end{bmatrix} \\ \text{pred} & \text{‘to break the ice}\langle\text{subj}\rangle\text{obj’} \\ \text{obj} & \begin{bmatrix} \text{pred} & \text{‘ice’} \\ \text{num} & \text{sg} \\ \text{gend} & \text{fem} \end{bmatrix} \end{bmatrix} \text{(a)} \qquad \begin{bmatrix} \text{subj} & \begin{bmatrix} \text{pred} & \text{‘John’} \\ \text{num} & \text{sg} \\ \text{pers} & 3 \end{bmatrix} \\ \text{pred} & \text{‘break}\langle\text{subj, obj}\rangle\text{’} \\ \text{obj} & \begin{bmatrix} \text{pred} & \text{‘ice’} \\ \text{num} & \text{sg} \\ \text{gend} & \text{fem} \end{bmatrix} \end{bmatrix} \text{(b)}$$

Figure 7: Preferred (a) and dispreferred (b) functional structures for the idiomatic and literal interpretations of *John breaks the ice*.

[13]We take into account only the nodes where *Pred* is associated to a lexical item (and not e.g. to a morpheme or discourse unit).

$$
\left[\begin{array}{ll}
\text{subj} & \left[\begin{array}{ll}\text{pred} & \text{‘John’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\
\text{pred} & \text{‘to prefer}\langle\text{subj, obj, vcomp}\rangle\text{’}\\
\text{vcomp} & \left[\begin{array}{ll}\text{pred} & \text{‘to be}\langle\text{subj,acomp}\rangle\text{’}\\ \text{subj} & [\ \ ]\\ \text{acomp} & \left[\begin{array}{ll}\text{pred} & \text{‘honest}\langle\text{subj}\rangle\text{’}\\ \text{subj} & [\ \ ]\end{array}\right]\end{array}\right]\\
\text{obj} & \left[\begin{array}{ll}\text{pred} & \text{‘daughter’}\\ \text{num} & \text{sg}\\ \text{gend} & \text{masc}\\ \text{poss} & \text{yes}\end{array}\right]
\end{array}\right] \text{(a)}
\qquad
\left[\begin{array}{ll}
\text{subj} & \left[\begin{array}{ll}\text{pred} & \text{‘John’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\
\text{pred} & \text{‘to prefer}\langle\text{subj, obj}\rangle\text{’}\\
\text{mod} & \left\{\left[\begin{array}{ll}\text{pred} & \text{‘to be}\langle\text{acomp}\rangle\text{’}\\ \text{acomp} & \left[\begin{array}{ll}\text{pred} & \text{‘honest}\langle\text{subj}\rangle\text{’}\\ \text{subj} & [\ \ ]\end{array}\right]\end{array}\right]\right\}\\
\text{obj} & \left[\begin{array}{ll}\text{pred} & \text{‘daughter’}\\ \text{num} & \text{sg}\\ \text{gend} & \text{masc}\\ \text{poss} & \text{yes}\end{array}\right]
\end{array}\right] \text{(b)}
$$

Figure 8: Preferred (a) and dispreferred (b) functional structure for *John prefers his daughter to be honest* with *to be honest* as an argument (a) or as a sentence modifier(b).

$$
\left[\begin{array}{ll}
\text{subj} & \left[\begin{array}{ll}\text{pred} & \text{‘John’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\
\text{pred} & \text{‘to say}\langle\text{subj, obj}_{to}\text{, comp}\rangle\text{’}\\
\text{comp} & \left[\begin{array}{ll}\text{pred} & \text{‘to give}\langle\text{subj, obj}\rangle\text{’}\\ \text{subj} & \left[\begin{array}{ll}\text{pred} & \text{‘Mary’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\ \text{obj} & \left[\begin{array}{ll}\text{pred} & \text{‘flower’}\\ \text{num} & \text{pl}\\ \text{gend} & \text{fem}\end{array}\right]\end{array}\right]\\
\text{obj}_{to} & \left[\begin{array}{ll}\text{pred} & \text{‘Paul’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]
\end{array}\right] \text{(a)}
\qquad
\left[\begin{array}{ll}
\text{subj} & \left[\begin{array}{ll}\text{pred} & \text{‘John’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\
\text{pred} & \text{‘to say}\langle\text{subj, comp}\rangle\text{’}\\
\text{comp} & \left[\begin{array}{ll}\text{pred} & \text{‘to give}\langle\text{subj, obj}_{to}\text{, obj}\rangle\text{’}\\ \text{subj} & \left[\begin{array}{ll}\text{pred} & \text{‘Mary’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\ \text{obj}_{to} & \left[\begin{array}{ll}\text{pred} & \text{‘Paul’}\\ \text{num} & \text{sg}\\ \text{pers} & \text{3}\end{array}\right]\\ \text{obj} & \left[\begin{array}{ll}\text{pred} & \text{‘flower’}\\ \text{num} & \text{pl}\\ \text{gend} & \text{fem}\end{array}\right]\end{array}\right]
\end{array}\right] \text{(b)}
$$

Figure 9: Preferred (a) and dispreferred (b) functional structure for *John says that Mary gives flowers to Paul*.

3 Preliminary Evaluation of XLFG

In order to test both the coverage of XLFG as well as the validity of the rule-based disambiguation principles we have implemented, we used 1,200 grammatical sentences for French from the test suite TSNLP (Estival and Lehman 1997). We parsed these 1,200 sentences with XLFG and our grammar for French (approximately 78 syntactic rules, 450,000 lexical items) without any disambiguation. We developed a tool to browse through all the analysis for each sentence (i.e. the different pairs of possible constituent structure / functional structure assigned to each sentence by the parser). We used the output of the parser and the annotation tool to manually build a gold-standard. We then reparsed the 1,200 sentences, this time with the disambiguation module turned on and compared this output to the gold standard. Without disambiguation, XLFG produced one or more correct parse(s) for 81.5% of the sentences (i.e. 972 of the 1,200 sentences), with an average of 3.36 functional structures per sentence. 80.3% of these sentences were ambiguous (i.e. 781 ambiguous sentences out of 972 parsed sentences). With disambiguation, XLFG still produced one or more correct parse(s) for 81.5% of the sentences, but this time with an average of 1.29 functional structures per sentence. Only 38% of the parsed sentences were now ambiguous (i.e. 373 out of 972 parsed sentences): 408 sentences (i.e. 52.2%) had been totally disambiguated, and had only one functional structure associated to them. And another 40 sentences (i.e. 5%) had been partially disambiguated, meaning that they were associated with fewer functional structures then when they were initially parsed without the disambiguator. More importantly, no correct functional structure was discarded.

The 20% of sentences that did not receive any parse in the first place were due either to unknown lexical items, or to the fact that our grammar did not deal with some syntactic phenomena. TSNLP is a test-suite, so the sentences are built artificially and resort on purpose to a limited lexicon. It would be interesting in the future to enrich both the grammar and lexicon of XLFG to make it more robust and perform a new evaluation on *real text*.

It is interesting to note though that when tested on the same TSNLP sentences XLFG obtained basically the same results as FTAG, a wide coverage grammar for French (Abeillé et al. 2000a) which is semi-automatically generated, and consists in 5,000 elementary trees and approximately 600,000 lexical entries: approx 80% of the sentences get at least one correct parse. Again, the goal of XLFG is not to be a robust wide-coverage parser, but, however, these results are very encouraging, especially since our LFG grammar contained far fewer rules than its TAG counterpart.

Conclusion

We have presented XLFG, an LFG parser freely available for the research community. We have shown how this tool can be used as a pedagogical tool, as well as to test and implement new linguistic hypothesis. We have so far manually developed a medium-size grammar for French of approximately 78 rules and 450,000 lexical entries. We have also implemented a rule-based disambiguation module which exploits the similarities between LFG functional structures and TAG derivation trees and have presented a preliminary evaluation of the parser and the disambiguator. In the future, we hope to augment the coverage of our grammar for French, and to develop grammars for other languages. In addition, we plan to investigate more systematically the link between the LFG framework and the TAG framework, based on the work initiated by (Kameyama 1986).

References

Abeillé, A., M. Candito, and A. Kinyon. 2000a. The current status of FTAG. In *TAG+5*, Paris.

Abeillé, A., L. Clément, and A. Kinyon. 2000b. Building a Treebank for French. In *Proceedings LREC'2000*, Athens.

Abeillé, A., and O. Rambow (Eds.). 2000. *Tree Adjoining Grammars: Formalism, linguistic analysis and processing.* CSLI publications.

Andrews, A. 1990. Functional closure in LFG. Technical report, The Austrilian National University.

Bresnan, J. 1982. *The mental representation of Grammaticals Relations.* MIT Press.

Brysbaert, M., and D. Mitchell. 1996. Modifier attachment in sentence parsing: Evidence from Dutch. *Quaterly Journal of experimental psychology* 49a:664–695.

Clément, L. 1996. Un analyseur syntaxique pour le français en LFG - quelques pistes pour l'implémentation de la coordination. Master's thesis, University Paris 7.

Clément, L. 2000. XLFG - Une plate-forme de développement de Grammaires Lexicales Fonctionnelles. In *TALN 2000*, Lausanne.

Clément, L. 2001a. *Construction et exploitation d'un corpus syntaxiquement annoté pour le français*. PhD thesis, University of Paris 7.

Clément, L. 2001b. XLFG a parser to learn LFG framework. In *NAACL Demos*, Pittsburgh.

Cuetos, M., and D. Mitchell. 1988. Cross linguistic differences in parsing: restrictions on the use of the late closure strategy in Spanish. In *Cognition 30*, 73–105.

Estival, D., and Lehman. 1997. TSNLP: des jeux de phrases test pour le TALN. *TAL* 8:1.

Frank, A., T. Holloway King, J. Kuhn, and J. Maxwell. 2000. Optimality Theory Style Constraint Ranking in Large-scale LFG Grammars. In *LFG'98*, Australia.

Frazier, L., and J. Fodor. 1978. The sausage machine: a two stage parsing model. In *Cognition 6*.

Joshi, A. 1994. Introduction to tree-adjoining grammars–Preface to the Special Issue on Tree Adjoining Grammars. *Computational Intelligence*.

Kameyama, M. 1986. Characterising Lexical Functional Grammar (LFG) in terms of Tree Adjoining Grammar (TAG). Technical report, Dept. of Computer and Information Science. University of Pennsylvania, Philadelphia.

Kaplan, R. 1989. The formal Architecture of Lexical Functionnal Grammar. In *Journal of Informations Science and Engineering.*

Kaplan, R., and J. Maxwell. 1994. *Grammar Writer's Workbench*. Xerox Corporation. Version 2.0.

Kimball, J. 1973. Seven principles of surface structure parsing in Natural Language. In *Cognition 2*.

Kinyon, A. 1999. Some remarks on the psycholinguistic relevance of LTAGs. In *CLIN'99*, Utrecht.

Kinyon, A. 2000. Are structural principles useful for automatic disambiguation ? In *COGSCI'00*, Philadelphia.

Knuth, D. E. 1965. On the translation of langages from left to right. In *Information and Control*.

Rambow, O. 1996. Word order, clause union, and the formal machinery of syntax. In M. Butt and T. H. King (Eds.), *On-line Proceedings of the First LFG Conference, Rank Xerox, Grenoble, August 26–28, 1996.*

Riezler, S., D. Prescher, J. Kuhn, and M. Johnson. 2000. Lexicalized Stochastic Modeling of ConstraintBased Grammars using Log-Linear Measures and EM Training. In *ACL'00*, Hong Kong.

Trueswell, J. 1996. The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language* 35:566–585.

Vappillon, J. 1997. Un atelier de Génie Linguistique pour l'analyse syntaxique fondée sur le formalisme LFG. In *JST Francil*, Avignon.

Zweigenbaum, P. 1991. Un analyseur syntaxique pour grammaires lexicales fonctionnelles. *T.A. Informations.*

WEAK CROSSOVER AND THE ABSENCE OF TRACES

Mary Dalrymple, Ronald M. Kaplan, and Tracy Holloway King
Xerox PARC

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong
Miriam Butt and Tracy Holloway King (Editors)
2001
CSLI Publications
http://csli-publications.stanford.edu/

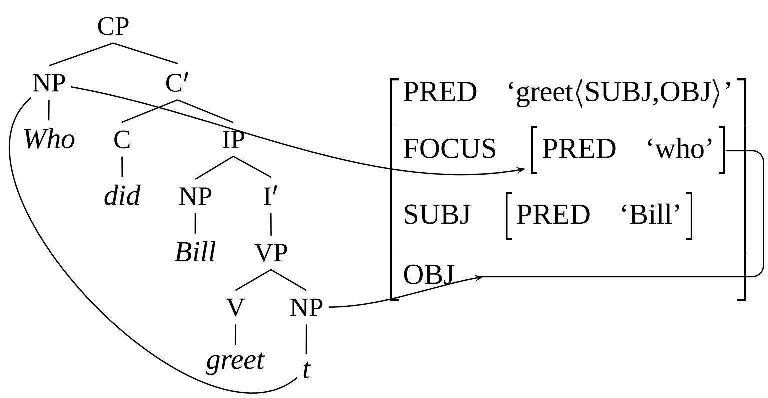
Abstract

We provide a new definition of the linear prominence constraints between pronouns and operators (wh-words and quantifiers) which correctly rules out examples that violate weak crossover. Previous analyses of weak crossover relied on the presence of a trace in the extraction site of a wh-question; in contrast, our analysis enables a traceless account of examples previously cited in support of traces. Since no other incontrovertible evidence for traces has been put forward, our account allows a return to the traceless and therefore more constrained account of long-distance binding in LFG proposed by Kaplan and Zaenen (1989).

1 Long-distance dependencies in LFG

Traces were introduced in transformational grammar (Chomsky, 1973) in order to give a phrasal account for a wide range of long-distance dependency phenomena: a long-distance dependency between a displaced element in a wh-question or relative clause was assumed to arise by movement of the displaced element, which leaves behind a ‘trace’ in its original position. The original LFG treatment of long distance dependencies (Kaplan and Bresnan, 1982) was based on an adaptation of the transformational/phrase-structure scheme. The relation between the displaced wh-phrase and its within-clause function was defined in terms of a relation between the wh-word and an empty c-structure constituent, a trace, within the clause. In example (1), the relation between the NP node dominating *who* and the NP node dominating its trace, represented by t, ensures that the wh-phrase *who* is both the FOCUS and the OBJ of the sentence:

(1)



Subsequently, however, Kaplan and Zaenen (1989) proposed that constraints on long distance dependencies are best stated in functional and not phrasal terms. As such, *functional uncertainty* offers a more accurate and direct characterization of the long distance dependency. A rule like the one in example (2) establishes two roles for the NP daughter of CP: it is the FOCUS, and it plays a grammatical role defined by the functional uncertainty path COMP* OBJ:

(2) $$\mathrm{CP} \longrightarrow \begin{array}{cc} \mathrm{NP} & \mathrm{C}' \\ (\uparrow \ \mathrm{FOCUS}) = \downarrow & \uparrow = \downarrow \\ (\uparrow \ \mathrm{COMP^{*}\ OBJ}) = \downarrow & \end{array}$$

In example (3), the path consisting simply of OBJ is chosen, and the FOCUS f-structure also appears as the OBJ of *greet*:

(3) Who did Bill greet?

$$\left[\begin{array}{ll} \mathrm{PRED} & \text{‘greet}\langle\mathrm{SUBJ,OBJ}\rangle\text{’} \\ \mathrm{FOCUS} & \left[\begin{array}{ll} \mathrm{PRED} & \text{‘who’} \end{array}\right] \\ \mathrm{SUBJ} & \left[\begin{array}{ll} \mathrm{PRED} & \text{‘Bill’} \end{array}\right] \\ \mathrm{OBJ} & \underline{\hspace{6em}} \end{array}\right]$$

With this account, long distance dependencies no longer provide arguments that trace-like devices are *necessary* in the theory of grammar. And in fact, a traceless theory is preferred for two reasons: it is more adequate descriptively, and it offers theoretical advantages. We outline these reasons below before discussing examples involving weak crossover and how a traceless theory can account for them.

A traceless theory is descriptively more adequate. That is, it can describe data that the traceful theory cannot, as argued in detail by Kaplan and Zaenen (1989). Kaplan and Zaenen (1989) cite evidence from wandering adverbs in Icelandic, cross-categorial dependencies in English, and cross-conjunct relativization constraints in Japanese that can be accounted for naturally in a traceless, but not a traceful, analysis. More recently, Sag (1998) has provided additional evidence against traces in his analysis of the conjunct constraint and of floating quantifiers and adverbs.

For several reasons, a traceless theory is also theoretically preferred. That is, even if the traceless and traceful theories accounted for exactly the same data, the traceless theory is preferable. First, the traceless theory is less redundant in that there is only a single way of dealing with long distance dependencies; it is a general scientific principle that less redundant theories are preferred to more redundant ones. Second, it is more restrictive in that it has fewer powerful mechanisms and is more tractable mathematically. Third, it avoids spuriously ambiguous analyses for fillers without ‘canonical’ phrase-structure positions. For example, in extraction of obliques and adverbs there is more than one possible extraction site, even in a relatively fixed word order language like English, as shown in (4).

(4) a. Oblique extr.: About what did John talk (t) to Mary (t)?

b. Adverb extr.: How often did John talk (t) to Mary (t) about apples (t)?

Finally, with the traceless theory, the listener's perceptual problem is simpler: the listener must guess only the function of the displaced element, not both a function and a position. Thus, there are both empirical and theoretical reasons to prefer the traceless account.

2 Weak crossover: A counterexample?

Despite the appeal of a traceless analysis of long distance dependencies, there is evidence that seems to suggest that traces cannot be eliminated: *weak crossover* (Postal, 1971; Wasow, 1979) seems to indicate that the linear position encoded by a trace does affect grammaticality. The sentence in (5) exemplifies a weak crossover violation:

(5) *Who$_i$ did his$_i$ mother greet t?
(cannot mean: Whose$_i$ mother greeted him$_i$?)

The name 'crossover' comes from the transformational analysis of wh-question formation: in a crossover violation, the wh-phrase 'crosses over' a coreferential pronoun when it is moved to the beginning of a sentence. In later analyses, the violation has been reformulated in terms of the position of the trace: a crossover violation ensues when, as in (5), a coreferential pronoun precedes the trace. We concentrate here on *weak crossover* violations, those in which the pronoun precedes but does not c-command the trace.

Crossover effects are found in *operator binding*, in particular, binding by WH-operators and quantifiers (Reinhart, 1983):

(6) a. *Who$_i$ did his$_i$ mother greet?

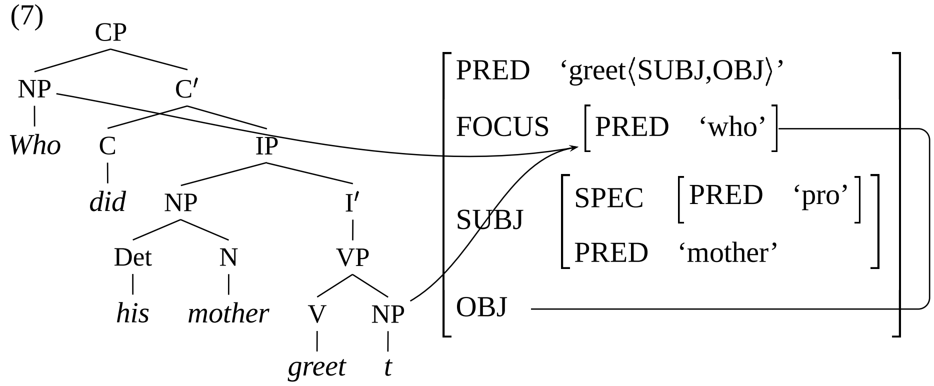
b. *His$_i$ mother greeted everyone$_i$.

Our main focus here is on the question of the necessity of traces, and our discussion will center around operator binding in wh-questions. We make only some brief remarks below about quantifier binding in German.

Some recent LFG accounts of operator binding and weak crossover propose a return to a theory of long distance dependency with traces (Bresnan, 1994, 2001;

Berman, 2000); these theories adopt the view that the relation between the pronoun and a trace of the displaced wh-phrase is the important one for characterizing weak crossover violations. These theories propose a representation like the one in example (7) for a question like *Who did his mother greet?*, and rule out coreference between the pronoun *his* and the operator *who* by reference to the relation between the position of the pronoun and the position of the trace of *who*:

(7)



In the following, we provide a reformulation of the constraints on the relation between a pronoun and an operator that allow an account of these examples *without* traces.

3 Prominence

Bresnan (1994, 1995, 2001) proposes that operator binding relations are constrained by (at least) two dimensions of *prominence*: syntactic prominence depends on the functional rank of the operator and the pronoun it binds, and linear prominence depends on the linear order between an operator and a pronoun that it binds. We first discuss Bresnan's definitions and then provide a revised set of definitions of these two prominence dimensions.

3.1 Syntactic and linear prominence

Bresnan (2001) defines operator binding requirements in the following way:

(8) **Syntactic Prominence** (A unit containing) the pronoun may not be higher than (a unit containing) the operator on the grammatical function hierarchy:
SUBJ > OBJ > COMP > …

Linear Prominence The pronoun must not f-precede the operator.

Bresnan further proposes that languages can vary in which of these constraints apply: some languages impose both kinds of prominence constraints, some require only one or the other, some require a disjunction of the two.

We believe that Bresnan's basic insight is correct: prominence requirements between an operator and a pronoun are multifaceted, and different languages can impose different kinds of prominence requirements between an operator and the pronoun it binds. It turns out, however, that a simple redefinition of the linear prominence condition makes many of the same predictions as the definitions above, including for the crucial cases of weak crossover, without assuming traces.

3.2 Rethinking linear prominence

The intuition behind our reformulation is that linear precedence requirements between an operator and a pronoun are determined by the overt material which indicates the *syntactic role* of the displaced phrase. This intuition follows the proposal of Sag (1998), though it differs in detail.

A theory with traces accounts for the unacceptability of an example like **Who did Sue talk about his mother to?* by reference to the relation between the pronoun *his* and the trace in the object position of the pronoun *to*, as in (9a). In contrast, our theory (like Sag's) refers to the relation between the position of the pronoun and the position of the stranded preposition *to*. The same results obtain, but without positing a trace, as in (9b).

(9) a. Traces: *Who$_i$ did Sue talk about his$_i$ mother to t_i?

b. No traces: *Who$_i$ did Sue talk about his$_i$ mother [to]?

We first define the notion of *coargument* to encompass the arguments as well as the adjuncts of a single predicate:

(10) **Coarguments**:

the arguments and adjuncts of a single predicate.

We also assume the definition of f-precedence in (11), though our proposed redefinition of linear precedence does not depend on adopting this particular formulation of f-precedence as opposed to other definitions of f-precedence that have been adopted in the literature (e.g. Bresnan, 1995).

(11) **F-precedence**:

f_1 f-precedes f_2 if and only if all c-structure nodes corresponding to f_1 precede all nodes corresponding to f_2. (e.g. Kaplan, 1989)

We now present our revised prominence requirements:

(12) Let CoargOp and CoargPro be coargument f-structures such that CoargOp contains O and CoargPro contains P. Then:

Syntactic Prominence An operator O is more prominent than a pronoun P if and only if CoargOp is at least as high as CoargPro on the functional hierarchy.

Linear Prominence An operator O is more prominent than a pronoun P if and only if CoargOp f-precedes P.

The most important difference between our revised definitions and those in (8) is that linear prominence depends on the f-precedence properties of *CoargOp*, an f-structure containing the operator, *not* on the operator.

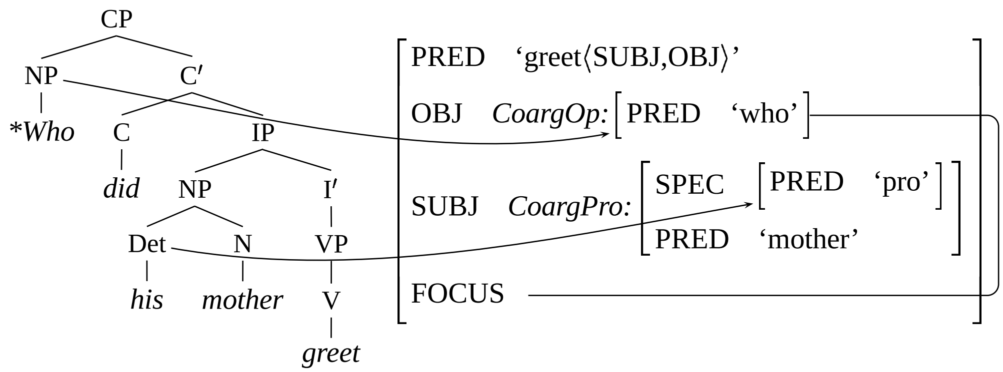
4 Evaluating the proposal

In this section, we show how our revised definitions can account for weak crossover phenomena in English, German, and Malayalam, without positing traces.

4.1 English

Following Bresnan (1995), we assume that in English, the operator must outrank any pronoun it binds in both Linear Prominence and Syntactic Prominence. With this, we correctly predict the unacceptability of example (13):

(13) *Who$_i$ did his$_i$ mother greet?



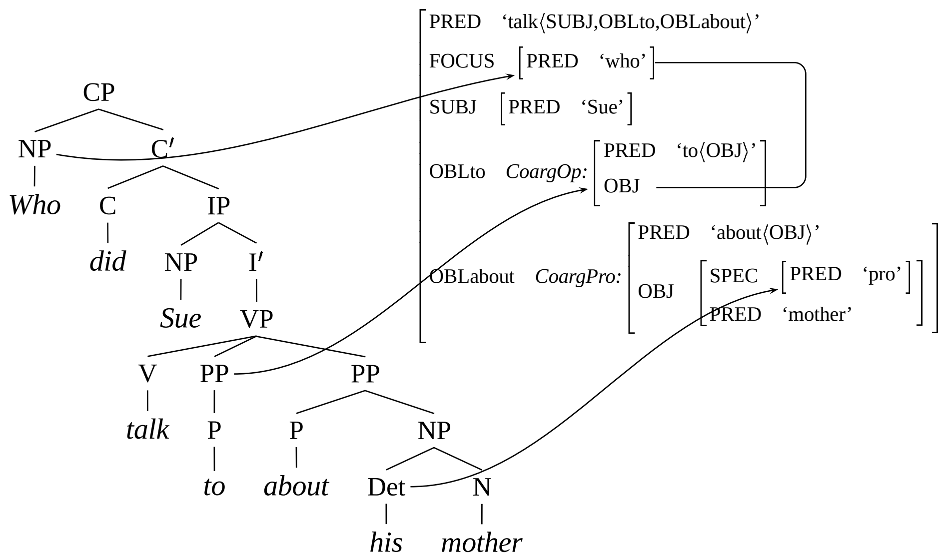
To apply our definitions of prominence, we must first locate the f-structure coarguments that contain the operator and the pronoun. We observe the following:

- SUBJ and OBJ of *greet* are coarguments
- pronoun is *his*
- SUBJ of *greet* contains the pronoun, so it is CoargPro
- OBJ of *greet* contains the operator, so it is CoargOp

On this basis, we check whether both Linear Prominence and Syntactic Prominence are satisfied. In (13), the OBJ *who* f-precedes the pronoun, since every c-structure node corresponding to the OBJ f-structure precedes the pronoun; thus, the Linear Prominence requirement is satisfied. However, the Syntactic Prominence requirement is violated, since CoargPro (SUBJ) outranks CoargOp (OBJ) on the functional hierarchy. This account for the unacceptability of the example.

Next, we examine the grammatical sentence *Who$_i$ did Sue talk to about his$_i$ mother?*:

(14) Who$_i$ did Sue talk to about his$_i$ mother?



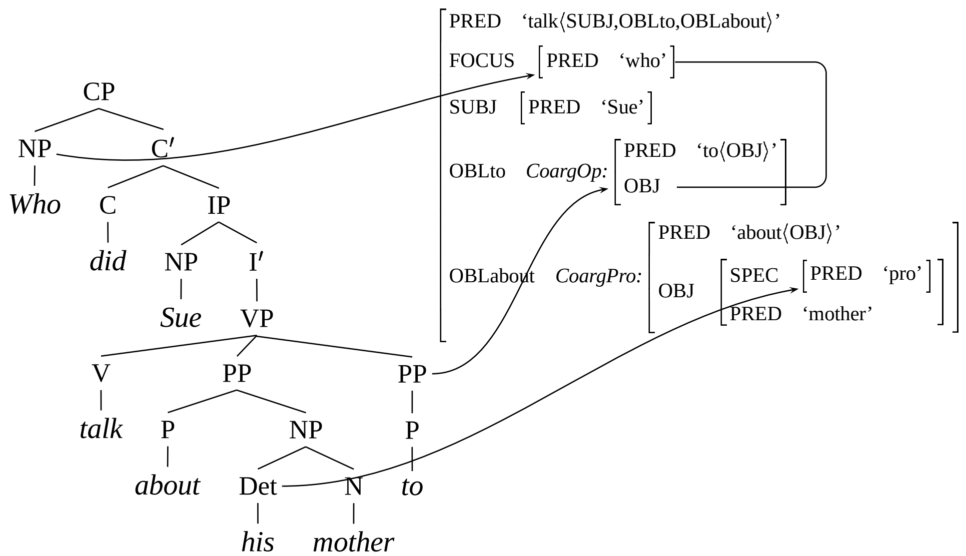
In this example:

- OBLto and OBLabout of *talk* are coarguments
- pronoun is *his*
- OBLabout of *talk* contains the pronoun, so it is CoargPro
- OBLto of *talk* contains the operator, so it is CoargOp

Again, we check to make sure that both Linear Prominence and Syntactic Prominence are satisfied. The syntactic prominence requirement is satisfied: we assume that all obliques occupy the same rank on the functional hierarchy, so that CoargOp (OBLto) is at least as high as CoargPro (OBLabout) on the functional hierarchy. The Linear Prominence requirement is also met. The CoargOp f-structure corresponds to the c-structure nodes PP and P dominating *to*. CoargOp f-precedes the pronoun f-structure, and the Linear Precedence condition is met. Since both requirements are satisfied, the example is correctly predicted to be grammatical.

Finally, we examine the ungrammatical example **Who$_i$ did Sue talk about his$_i$ mother to?*:

(15) *Who$_i$ did Sue talk about his$_i$ mother to?



(15) has the same coarguments as (14). Again, the Syntactic Prominence requirement is met, since the two obliques occupy the same rank on the functional hierarchy. The c-structure nodes corresponding to CoargOp are the PP and P dominating *to*; these nodes do not f-precede the pronoun, the Linear Precedence requirement is not met, and the example is correctly classified as ungrammatical.

Thus, we have seen that traces are not necessary to account for the English weak crossover data once coarguments are taken into account.

4.2 German

Following Berman (2000) and Bresnan (2001), we assume that operator binding in German requires that *either* the Linear Prominence *or* the Syntactic Prominence requirement must be met; unlike English, meeting a single requirement suffices for grammaticality.

First, we consider some patterns of quantifier binding. Both of the examples in (16) are grammatical:

(16)	…dass	[jeder]	[seine Mutter]	mag
	that	everyone-NOM	his mother	likes
		(CoargOp, SUBJ)	(CoargPro, OBJ)	
	…dass	[seine Mutter]	[jeder]	mag
	that	his mother	everyone-NOM	likes
		(CoargPro, OBJ)	(CoargOp, SUBJ)	

'… that everyone$_i$ likes his$_i$ mother'

In example (17), the quantifier *jeder* 'everyone' is the CoargOp SUBJ, and CoargPro is the OBJ:

- SUBJ and OBJ of *mag* 'like' are coarguments
- pronoun is *seine*
- OBJ of *mag* 'like' is CoargPro
- SUBJ of *mag* 'like' is CoargOp

The Syntactic Prominence requirement is met in both examples, since SUBJ outranks OBJ. Since only a single requirement must be satisfied, the Linear Prominence requirement need not hold, and the SUBJ and OBJ may appear in either order.

When the Linear Prominence requirement is met, the Syntactic Prominence requirement need not be met. In example (17), CoargOp is the OBJ *jeden* 'everyone', and CoargPro is the SUBJ *seine Mutter* 'his mother'. Though the Syntactic Prominence requirement is not met, the CoargOp f-precedes the pronoun, and the example is grammatical:

(17)	…dass	[jeden]	seine Mutter	mag
	that	everyone-ACC	his mother	likes
		(CoargOp)		

'… that his$_i$ mother likes everyone$_i$'

- SUBJ and OBJ of *mag* 'like' are coarguments
- pronoun is *seine*
- SUBJ of *mag* 'like' is CoargPro
- OBJ of *mag* 'like' is CoargOp

In example (18), which has the same coargument structure of (17), neither requirement is met, and the sentence is ungrammatical:

(18)	*…dass	[seine Mutter]	[jeden]	mag
	that	his mother	everyone-ACC	likes
		(CoargPro, SUBJ)	(CoargOp, OBJ)	

‘… that his$_i$ mother likes everyone$_i$’

We now turn to examples involving long distance dependencies. In example (19), CoargPro is the SUBJ of ‘say’, *seine Mutter* ‘his mother’. CoargOp is the COMP of ‘say’. Since SUBJ outranks COMP, CoargOp is not more syntactically prominent than CoargPro. And since CoargOp does not f-precede the pronoun, the linear prominence requirement does not hold either. Since neither requirement holds, the sentence is classified as ungrammatical. Note that the reason for the ungrammaticality of example (19) is not that there is a trace in the subordinate clause; instead, the subordinate clause itself is treated as the important constituent in determining precedence requirements between the operator and the pronoun.

(19)	*jeden/wen	meinte	[seine Mutter],	[habe sie getröstet]
	everyone/who	said	his mother	has she consoled
			(CoargPro, SUBJ)	(CoargOp, COMP)

‘Everyone$_i$, his$_i$ mother said that she consoled./
Who$_i$ did his$_i$ mother say that she consoled?’

- SUBJ and COMP of *meinte* ‘say’ are coarguments
- pronoun is *seine*
- SUBJ of *meinte* ‘say’ is CoargPro
- COMP of *meinte* ‘say’ is CoargOp

In example (20), we evaluate the prominence conditions with respect to the coargument SUBJ and OBJ of *getröstet* ‘consoled’: the SUBJ *seine Mutter* ‘his mother’ is CoargPro, and the OBJ is CoargOp. The Syntactic Prominence condition does not hold, since the CoargPro SUBJ outranks the CoargOp OBJ. But the Linear Prominence condition holds, since the CoargOp f-precedes the pronoun. Thus, the operator binding conditions for German are met, and the sentence is predicted to be grammatical:

(20) [jeden/wen] sagte sie, habe seine Mutter getröstet
everyone/who said she has his mother consoled
(CoargOp)
'Everyone$_i$, she said that his$_i$ mother consoled./
Who$_i$ did she say that his$_i$ mother consoled?'

- SUBJ and OBJ of *getröstet* 'consoled' are coarguments
- pronoun is *seine*
- SUBJ of *getröstet* 'consoled' is CoargPro
- OBJ of *getröstet* 'consoled' is CoargOp

Thus, we see that in German simplex and complex clauses, there is no need to posit traces to account for the weak crossover data.

4.3 Malayalam

Finally, we briefly examine the behavior of 'null pronouns', phonologically unrealized pronominal elements. As shown by Mohanan (1982) and Bresnan (2001), only the Linear Prominence condition is relevant in Malayalam: the CoargOp must f-precede the pronoun, but no syntactic prominence condition is relevant. Example (21) is ungrammaticalwith an overt pronoun:

(21) *in̲n̲ale awar̄e s̄akaar̄icca st̲riikaḷe in̲n̲ə [oor̄oo kuṭṭiyum] sahaayiccu
yesterday they scolded woman today each child helped
(CoargOp)
'Today each child$_i$ helped the woman who scolded them$_i$ yesterday.'

- SUBJ and OBJ of *sahaayiccu* 'helped' are coarguments
- pronoun is *awar̄e*
- OBJ of *sahaayiccu* 'helped' is CoargPro
- SUBJ of *sahaayiccu* 'helped' is CoargOp

In this example, CoargOp is the quantifier phrase *oor̄oo kuṭṭiyum*. CoargOp does not f-precede the overt pronoun *awar̄e*, and so the Linear Prominence requirement is not met.

However, with a null pronoun, the sentence is grammatical:

(22)	in̲n̲ale	∅	s̄akaar̄icca	st̲riikaḷe	in̲n̲ə	[oor̄oo	kuṭṭiyum]	sahaayiccu
	yesterday	(they)	scolded	woman	today	each (CoargOp)	child	helped

‘Today each child$_i$ helped the woman who scolded them$_i$ yesterday.’

The same coargument information holds for (22) as for (21). However, the null pronoun does not correspond to any c-structure nodes, and so according to the definition of f-precedence in (11), it vacuously f-precedes and is f-preceded by every other f-structure. In particular, it vacuously satisfies the Linear Prominence condition. This explains the difference between the overt and null pronouns in examples (21) and (22).

5 Empirical differences between trace and traceless accounts

Though the data from English, German, and Malayalam discussed above are classified correctly by our theory of prominence as well as by theories that assume traces, there are empirically testable differences between the two approaches; in particular, the availability of certain sets of data would allow us to distinguish between our traceless account and the accounts of Bresnan (1994), Bresnan (2001), and Berman (2000), which assume traces. Though we have not yet been able to find the crucial data that would allow us to decide between the two theories, we provide characterizations of some of these differences in the following, in hopes that some of these crucial data can be found and examined.

5.1 Language 1

We first examine what we will call Language 1, a language in which the Syntactic Prominence condition does not apply. More specifically, Language 1 has the following characteristics:

- Fixed word order
- Wh-phrases are displaced to the beginning of the sentence
- Only Linear Prominence is relevant

In Language 1, our theory predicts that an example like (23) would be grammatical, since CoargOp precedes the pronoun, and by hypothesis the Syntactic Prominence condition does not apply:

(23)	[who$_i$]	did	[his$_i$ mother]	see (t_i)
	(CoargOp, OBJ)		(CoargPro, SUBJ)	

In contrast, the traceful theory predicts ungrammaticality, since the pronoun outranks the (trace of the) operator in Linear Prominence.

5.2 Language 2

We next examine Language 2, a language in which the object precedes the subject. Language 2 has the following characteristics:

- Fixed word order; Object precedes subject
- Wh-phrases are displaced to the beginning of the sentence
- Both Linear Prominence and Syntactic Prominence must hold

In Language 2, our theory predicts grammaticality for example (24):

(24)	[who$_i$]	saw his$_i$ book	(t_i)
	(CoargOp, SUBJ)	(CoargPro, OBJ)	

Here the Syntactic Prominence requirement holds, since CoargOp is SUBJ and CoargPro is OBJ. The Linear Prominence requirement is also met, since CoargOp precedes the pronoun. In contrast, the traceful theory predicts ungrammaticality, since the pronoun precedes the trace.

5.3 Language 3

Finally, we examine Language 3, a language in which either prominence requirement must hold. Language 3 has the following characteristics:

- Fixed word order; Subject precedes object
- Wh-phrases are displaced to the beginning of the sentence
- Either Linear Prominence or Syntactic Prominence must hold

Our theory predicts grammaticality for example (25):

(25)	[what$_i$]	saw its$_i$ owner	(t_i)
	(CoargOp, OBJ)	(CoargPro, SUBJ)	

The Syntactic Prominence requirement is not met in this example, since CoargOp is OBJ and CoargPro is SUBJ. However, the Linear Prominence requirement is met, since CoargOp precedes the pronoun. The traceful theory predicts that this example is ungrammatical, however, since once again the pronoun precedes the trace.

6 Summary

Our traceless theory of weak crossover accounts for the data on weak crossover that motivated previous researchers to posit traces in long distance dependencies. We accomplished this by providing a new definition of Linear Prominence which applies to f-structure units that contain the operator. Since we are aware of no other evidence supporting the existence of traces in long distance dependencies, we believe that traces remain unmotivated in the theory of grammar.

Acknowledgments

We are grateful to Yehuda Falk and the audience at LFG2001 for helpful discussion and comments.

References

Berman, Judith. 2000. *Topics in the Clausal Syntax of German*. Ph.D. thesis, University of Stuttgart, Stuttgart, Germany.

Bresnan, Joan. 1994. Linear order vs. syntactic rank: evidence from weak crossover. In Katie Beals, Jeannette Denton, Bob Knippen, Lynette Melnar, Hisami Suzuki, and Erika Zeinfeld (editors), *Papers from the Thirtieth Regional Meeting of the Chicago Linguistic Society*. University of Chicago: Chicago Linguistic Society.

Bresnan, Joan. 1995. Linear order, syntactic rank, and empty categories: On weak crossover. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell, and Annie Zaenen (editors), *Formal Issues in Lexical-Functional Grammar*, pp. 241–274. Stanford University: CSLI Publications.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell Publishers.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson and Paul Kiparsky (editors), *A Festschrift for Morris Halle*. New York: Holt, Reinhart, and Winston.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell, III, and Annie Zaenen (editors). 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford University: CSLI Publications.

Kaplan, Ronald M. 1989. The formal architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen (editors), *Proceedings of ROCLING II*, pp. 3–18. Reprinted in Dalrymple et al. (1995, pp. 7–27).

Kaplan, Ronald M. and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan (editor), *The Mental Representation of Grammatical Relations*, pp. 173–281. Cambridge, MA: The MIT Press. Reprinted in Dalrymple et al. (1995, pp. 29–130).

Kaplan, Ronald M. and Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin and Anthony Kroch (editors), *Alternative Conceptions of Phrase Structure*, pp. 17–42. Chicago University Press. Reprinted in Dalrymple et al. (1995, pp. 137–165).

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In Joan Bresnan (editor), *The Mental Representation of Grammatical Relations*, pp. 504–589. Cambridge, MA: The MIT Press.

Postal, Paul M. 1971. *Cross-Over Phenomena*. New York: Holt, Rinehart, and Winston.

Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. Chicago: The University of Chicago Press.

Sag, Ivan A. 1998. Without a trace. MS, Stanford University.

Wasow, Thomas. 1979. *Anaphora in Generative Grammar*. Ghent: E. Story.

Addresses:
{dalrymple,kaplan,thking}@parc.xerox.com
Xerox PARC, Palo Alto, CA 94304 USA

Constituent Structure and Grammatical Functions in the Hebrew Action Nominal

Yehuda N. Falk
The Hebrew University of Jerusalem

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http//csli-publications.stanford.edu/

I would like to thank Joan Bresnan for suggesting to me that I look into Hebrew action nominals. Versions of this paper were presented at the 2001 conference of the Israel Association of Theoretical Linguistics (Jerusalem) and the LFG2001 conference (Hong Kong). I have benefitted from discussions (in person and by e-mail) with Joan Bresnan, Miriam Butt, Carmen Dobrovie-Sorin, Rob Malouf, and Irit Meir.

1. Overview

Hebrew noun phrases in general, and action nominals in particular, pose interesting theoretical and descriptive problems.

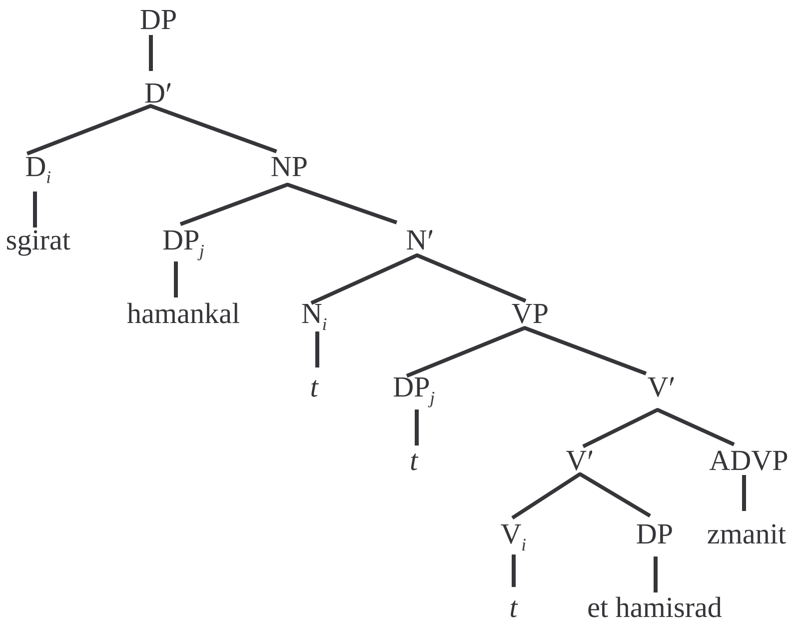
(1) a. ha- iša šel ha- politikay
the- wife of the- politician
b. ešet ha- politikay
wife.CONSTR the- politician
c. išt- o šel ha- politikay
wife- his of the- politician
'the politician's wife'

(2) a. ibud ha- mumxim et hakolot yadanit
processing the- experts ACC votes manually
'the experts' manual processing of the ballots'
b. sgirat ha- mankal et ha- misrad zmanit
closure.CONSTR the- director ACC the- office temporarily
'the director's temporary closure of the office'

In (1) we see the three types of genitive constructions in Hebrew. In the free genitive (1a), the possessive nominal is marked by the preposition *šel*. In the construct state genitive (1b), the possessive immediately follows the head noun, and the head noun appears in a special morphological form traditionally called the construct state. The double genitive (1c) combines a pronominal suffix on the head noun (in a form phonologically related to the construct state) with a PP headed by *šel*. In (2) we see a clearly nominal head in a construct state genitive construction, followed by an accusative object and an adverb.

These Hebrew noun phrase patterns have been discussed in a series of studies in the Government/ Binding (GB) and Minimalist Program (MP) variants of transformational theory. The analyses, although they differ from each other in details, all posit head movement from N to a functional head position (generally identified with D), and most of them hypothesize V-to-N head movement in the action nominal.

(3)



The claim to be made here is that various aspects of this analysis are unmotivated, in particular details of the constituency and the presence of the functional category D. It will be argued that the standard GB/MP analysis is the consequence of a theory in which grammatical functions are represented in terms of constituent structure, and that an approach in which grammatical functions are modeled as a distinct dimension of linguistic structure is better able to account for both the grammatical functions and the constituency. On the other hand, we will support the hypothesis that the structure of the action nominal in Hebrew includes both verbal and nominal projections. In lexicalist terms, the Hebrew action nominal is what Bresnan (1997) calls a “mixed category,” and we will show how Bresnan’s theory accounts for the mixed properties of the action nominal without derivations and empty categories. We will also argue that the lexicalist implementation of the NP-over-VP analysis of mixed categories is superior to the derivational implementation in the analysis of Hebrew action nominals.

2. The Hebrew Noun Phrase

2.1. The Problem

The N-to-D movement structure shown above (3) is supposed to account for various facts about noun phrases, such as the relative order of nouns and adjectives, the special morphology of construct state, and the inheritance of the definiteness of the possessive noun phrase by the construct state head. These are illustrated in the following.

(4) a. ha- gina ha- metupax- at
the- garden(F) the- cared.for- FSG
‘the tended garden’

b. gina metupax- at
garden(F) cared.for- FSG
‘a tended garden’

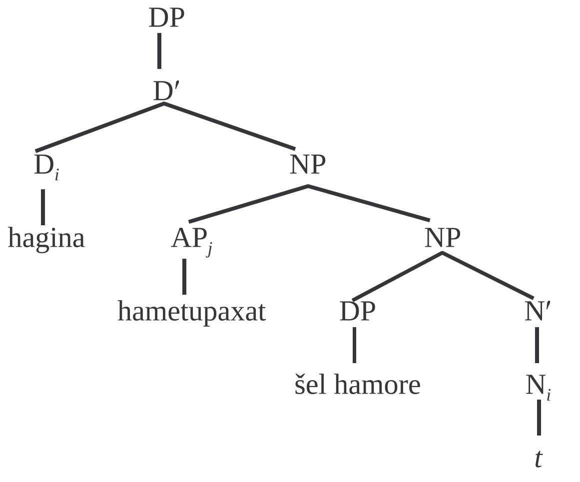
(5) a. ha- gina ha- metupax- at šel ha- more
the- garden the- cared.for- FSG of the- teacher(M)

b. ginat ha- more ha- metupax- at
garden(F).CONSTR the- teacher(M) the- cared.for- FSG

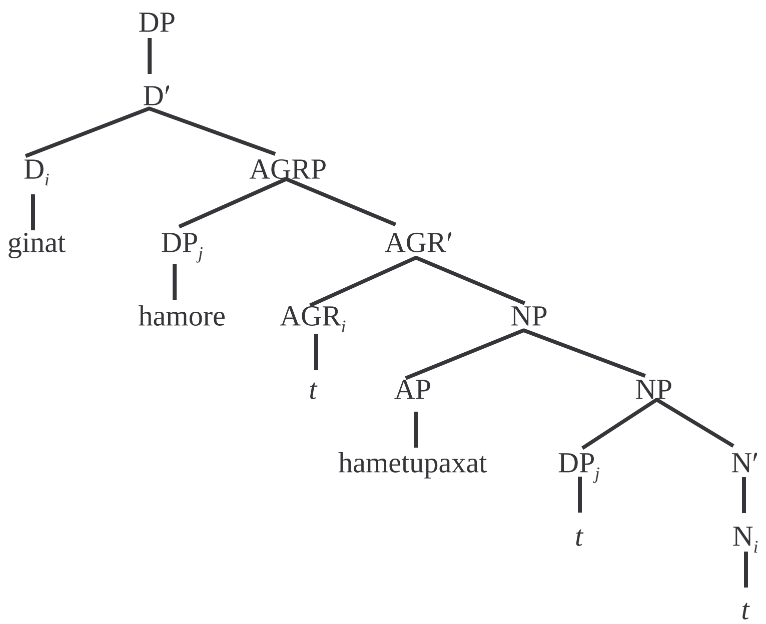
‘the teacher’s tended garden’

To make this somewhat more concrete, consider the structures assigned by the theory of Siloni (1997) to (5a,b):

(6) a.



b.



These structures, and similar ones in other analyses, are motivated by several considerations. In the first place, the possessor must occupy the D-structure position [SPEC, NP], the structural position universally associated with possessors ("subjects" of noun phrases). This specifier position of possessors is motivated even further in the case of nominals with two arguments, like action nominals, in which the possessor can bind the other (complement) position. On the assumption that binding is governed by the structural relation of c-command, this provides evidence for the higher structural position of possessors. Second, the AP adjunct must be in an adjoined position; again, this is motivated on universal grounds. Finally, positions must be hypothesized as the targets of movement in order to derive the surface word order, which does not correspond to the D-structure. The functional category D and other functional categories (such as AGR) are motivated on those grounds. On the assumption that agreement is the reflex of a SPEC-head relation, the movement to higher positions also supports various elements of the morphology.

The above argumentation for the constituent structure of the Hebrew noun phrase is theory-internal. It is based on an unconstrained theory of categories and structure, one in which categories need not be justified on lexical grounds and constituency need not be argued for. It is therefore instructive to test the resulting constituency against a more constrained theory, such as those assumed in unification/constraint-based lexicalist theories.

The first problem is the positing of a category Determiner. As observed by Wintner (2000), there is no evidence in Hebrew for such a category. The Hebrew "definite article" *ha-* is a prefix, not a word (Engelhardt 1998). It provides what in Hebrew is an inflectional feature of definiteness, one which triggers agreement on modifying adjectives. The only reason for positing D is, as mentioned above, to provide a landing site for the noun.

Second, as argued by Sadler (2000) for similar analyses of Welsh, the constituency itself is inconsistent with traditional constituency tests, like coordination.

(7) a.

ha-	gina	ha-	metupaxat	ve-	ha-	bayit	ha-	yafe	šel	ha-	more
the-	garden	the-	cared.for	and-	the-	house	the-	beautiful	of	the-	teacher

'the teacher's tended garden and beautiful house'

b.

*ginat	ha-	more	ha-	metupaxat	ve-	ha-	talmid
garden.CONSTR(F)	the-	teacher(M)	the-	cared.for.F	and-	the-	student(M)

ha-	muznaxat
the-	neglected.F

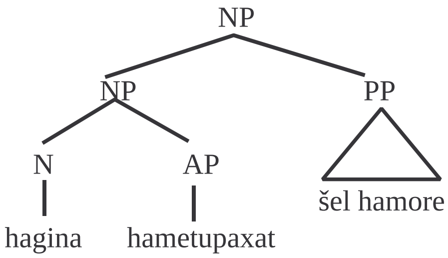
'the teacher's tended garden and the student's neglected one'

c. *ginat ha- more ve- beyt ha- talmid
garden.CONSTR the- teacher and- house.CONSTR the- student
ha- xadaš- im
the- new- PL
‘the new garden of the teacher and house of the student’

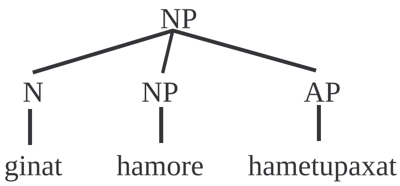
Contrary to the structures assumed in the transformational literature, (7a) shows that the noun and adjective form a constituent which excludes the *šel* phrase and (7b) shows that the possessive NP and adjective do not form a constituent. On the other hand, contrary to what one might expect, the construct-state nominal and possessive NP do not form a constituent that excludes the adjective (7c).

The constituent structures motivated by the above considerations are significantly different from the ones proposed in transformationalist accounts of Hebrew noun phrases.[1]

(8) a.



b.



The question is how to reconcile the apparently conflicting evidence concerning the structure of Hebrew NPs, so that the observations of both approaches can be accounted for.

2.2. Analysis

The theoretical framework to be assumed here is Lexical-Functional Grammar (LFG: Bresnan, ed 1982, Bresnan 2001, Falk 2001). In LFG, constituency and grammatical functions are treated as distinct dimensions of syntactic structure, coexisting in the overall linguistic representation. Parallel to the constituent structure (c-structure), which models the constituency relations, and thus the distribution, of overtly occurring elements, there is a functional structure (f-structure) which models the grammatical functions. First approximations at the f-structures of (5a,b) are as follows:

(9) a.

$$\left[\begin{array}{ll}
\text{PRED} & \text{‘garden } \langle(\uparrow \text{ POSS})\rangle\text{’} \\
\text{GEND} & \text{F} \\
\text{NUM} & \text{SG} \\
\text{DEF} & + \\
\text{POSS} & \left[\begin{array}{ll}
\text{CASE} & \text{POSS} \\
\text{PRED} & \text{‘teacher’} \\
\text{GEND} & \text{M} \\
\text{NUM} & \text{SG} \\
\text{DEF} & +
\end{array}\right] \\
\text{ADJ} & \{[\text{PRED ‘cared-for’}]\}
\end{array}\right]$$

[1]Similar, though not identical, structures are proposed by Dobrovie-Sorin (2001).

b.

$$\left[\begin{array}{ll}
\textsc{pred} & \text{‘garden } \langle(\uparrow \textsc{poss})\rangle\text{’} \\
\textsc{gend} & \textsc{f} \\
\textsc{num} & \textsc{sg} \\
\textsc{def} & + \\
\textsc{poss} & \left[\begin{array}{ll} \textsc{pred} & \text{‘teacher’} \\ \textsc{gend} & \textsc{m} \\ \textsc{num} & \textsc{sg} \\ \textsc{def} & + \end{array}\right] \\
\textsc{adj} & \{[\textsc{pred}\ \text{‘cared-for’}]\}
\end{array}\right]$$

The f-structures in (9) correctly represent the grammatical functions as identical in the two structures. The functional analysis matches closely the D-structure positions hypothesized in the transformational approach. The only difference is that in LFG it is not assumed that grammatical functions are uniformly encoded in constituent structure. The abandonment of this assumption is justified independently of the Hebrew facts by data in nonconfigurational languages (Austin and Bresnan 1996, Nordlinger 1998, Bresnan 2001, Falk 2001). Instead, the mapping between c-structure and f-structure is determined on a language-specific basis, and is formalized as functional equations which are annotated to c-structure positions in phrase structure rules and to lexical entries. The following phrase structure rules provide the basis of the constituency-function mapping in Hebrew NPs:

(10) a.

$$\begin{array}{lcc}
\text{NP} \rightarrow & \text{NP} & \text{PP} \\
 & \uparrow=\downarrow & (\uparrow(\downarrow \textsc{case}))=\downarrow
\end{array}$$

b.

$$\begin{array}{lccc}
\text{NP} \rightarrow & \text{N} & \text{NP} & \text{AP*} \\
 & \uparrow=\downarrow & (\uparrow \textsc{poss})=\downarrow & \downarrow \in (\uparrow \textsc{adj})
\end{array}$$

That is to say, the immediately postnominal NP position is assigned the function POSS, (we will modify this later) and PPs adjoined to NP have grammatical functions defined by the Case properties of the prepositions. At the lexical level, we assume the following lexical entries.

(11) a. *hamore* N
$(\uparrow \textsc{pred})$ = ‘teacher’
$(\uparrow \textsc{gend}) = \textsc{m}$
$(\uparrow \textsc{num}) = \textsc{sg}$
$(\uparrow \textsc{def}) = +$

b. *šel* P
$(\uparrow \textsc{case}) = \textsc{poss}$

c. *hametupaxat* A
$(\uparrow \textsc{pred})$ = ‘cared-for’
$((\textsc{adj}\ \uparrow)\ \textsc{gend}) = \textsc{f}$
$((\textsc{adj}\ \uparrow)\ \textsc{num}) = \textsc{sg}$
$((\textsc{adj}\ \uparrow)\ \textsc{def}) = +$

The heart of the analysis is the treatment of the various realizations of the head of the NP, *gina* ‘garden’ in the example. We assume that non-action nouns optionally take a POSS argument,[2] variously interpreted in terms of different kinds of possession (alienable, inalienable, agent responsible for a result) depending on the semantics of the noun. We thus have the following two lexical entries for the word *gina*.

[2] This essentially follows the approach of Bresnan (2001), where the addition of POSS to the argument structure is stated in terms of an a-structure augmentation template.

(12) a. *gina* N ($\uparrow$ PRED) = ‘garden’
($\uparrow$ NUM) = SG
($\uparrow$ GEND) = F

b. *gina* N ($\uparrow$ PRED) = ‘garden ⟨($\uparrow$ POSS)⟩’
($\uparrow$ NUM) = SG
($\uparrow$ GEND) = F

Nouns can be affixed with the definite prefix *ha-*, which adds a definiteness feature.[3]

(13) a. *hagina* N ($\uparrow$ PRED) = ‘garden’
($\uparrow$ NUM) = SG
($\uparrow$ GEND) = F
($\uparrow$ DEF) = +

b. *hagina* N ($\uparrow$ PRED) = ‘garden ⟨($\uparrow$ POSS)⟩’
($\uparrow$ NUM) = SG
($\uparrow$ GEND) = F
($\uparrow$ DEF) = +

The construct state is a morphophonological variant of the ordinary form of the noun. The form is usually predictable: the two most common alternations in the singular are that in feminine nouns ending in *-a* the construct usually ends in *-(a)t*, sometimes with the reduction of an internal vowel (‘garden’ *gina* construct *ginat*, ‘lip, language’ *safa* construct *sfat*, ‘family’ *mišpaxa* construct *mišpaxat*), and nouns with dipthongs simplify the diphthong (‘house’ *bayit* construct *beyt*, ‘olive’ *zayit* construct *zeyt*). The plural suffix *-im* is changed into *-ey*. In addition to its use in syntax, and perhaps more centrally, the construct is used morphologically as the bound variant of a word, similarly to English forms like *destruct* and *recept*. It is found primarily in compounds (14a), although it is sometimes used with affixes as well (14b).

(14) a. *bayit* ‘house’ + *sefer* ‘book’ → *beyt sefer* ‘school’
safa ‘lip, language’ + *em* ‘mother’ → *sfat em* ‘mother tongue’
miflaga ‘political party’ + *avoda* ‘work’ → *mifleget (ha)avoda* ‘(the) Labor Party’

b. *bayit* ‘house’ + *-i* adjective → *beyti* ‘domestic’
safa ‘lip, language’ + *-on* → *sfaton* ‘lipstick’
šana ‘year’ + *-on* → *šnaton* ‘(annual) course catalog’
zayit ‘olive’ + *-im* plural → *zeytim* ‘olives’

In its syntactic use, the construct has a bound “feel” as well. Some following NP is obligatory, either a POSS (1, 2) or an adjunct similar to English pre-nominal NP/DP adjuncts (15).

(15) *sfarim* ‘books’ + *Harry Potter* → *sifrey Harry Potter* ‘Harry Potter books’
gina ‘garden’ + *vradim* ‘roses’ → *ginat vradim* ‘rose garden’
bgadim ‘clothing’ + *yeladim* ‘children’ → *bigdey yeladim* ‘children’s clothing’

This bound-morpheme-like quality also manifests itself in the fact that the head noun inherits its definiteness from the NP, and the fact that POSS is realized as a plain NP rather than a PP headed by *šel*.

[3]We assume that nouns with no definiteness feature get the default value ‘–’ in the f-structure.

When faced with this kind of collection of facts, the question is how they are to be related to each other. The derivational approach generally accounts for the difference between regular and construct state nouns by attributing different Case-marking properties to construct state nominals and free state nominals (Ritter 1988, Siloni 1997). The abstract determiner is endowed with special Case features in the construct state, thus allowing the possessor nominal to be expressed without morphologically overt Case. The availability of "structural" genitive Case in the construct also controls the surface position of the possessor nominal, and given the appropriate feature-checking mechanism can be used to achieve the passing of the definiteness feature.

What is not explained by the Case-based approach is the morphophonological form, or the relation between the syntactic and morphological uses of the construct form. Wintner (2000) comes a step closer to these aspects in an HPSG-based analysis, hypothesizing a "dependency" attribute DEP for the construct, and linking its value to an immediately postnominal NP.

(16) The relation between absolute and construct forms (HPSG, Wintner 2000)

$$\left[\begin{array}{l} \textbf{word} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM}\,|\,\text{LOC}\,|\,\text{CAT} \left[\begin{array}{l} \textbf{cat} \\ \text{HEAD} \left[\begin{array}{l} \textbf{noun} \\ \text{DEF } - \end{array}\right] \\ \text{COMPS } \boxed{3} \oplus (\text{PP[of]}) \oplus \boxed{5} \end{array}\right] \end{array}\right] \mapsto$$

$$\left[\begin{array}{l} \textbf{word} \\ \text{PHON } \textit{phon\_reduce } (\boxed{1}) \\ \text{SYNSEM}\,|\,\text{LOC}\,|\,\text{CAT} \left[\begin{array}{l} \textbf{cat} \\ \text{HEAD} \left[\begin{array}{l} \textbf{noun} \\ \text{DEF } \boxed{2} \end{array}\right] \\ \text{COMPS } \boxed{3} \oplus \boxed{4} \left[\begin{array}{l} \textbf{synsem} \\ \text{CAT}|\text{HEAD} \left[\begin{array}{l} \textbf{nominal} \\ \text{DEF } \boxed{2} \end{array}\right] \end{array}\right] \oplus \boxed{5} \end{array}\right] \\ \text{DEP } \boxed{4} \end{array}\right]$$

We will follow Wintner's basic idea here. We take the basic property of the construct form to be its bound-morpheme-like quality, which we express in terms of an attribute similar to Wintner's DEP. However, Wintner's name for the attribute is counterintuitive, since the value of the attribute is the dominant element rather than the dependent one. We will call the attribute DOM. In addition to its grammatical function (POSS or ADJ), the postnominal NP position also fills the value of the head's DOM attribute. Wintner ignores the question of exactly how this attribute fits into the overall structure, including it neither in PHON nor in SYNSEM. Ultimately, this seems to be a morphological property marking bound forms of stems; we will tentatively include it in f-structure. Nothing hinges on this, and it probably should be part of some morphological projection.[4] We update the phrase structure rule for NP as follows.

---

[4]The analysis in the text, involving the attribute DOM, is undoubtedly an overly simplified implementation, but it will do for present purposes. There seem to be several aspects involved here. First, the postnominal NP has some grammatical function, the exact nature of which needs further investigation; perhaps a better understanding of the relation between POSS and SUBJ will help. Second, the head noun is a morphologically bound form; we can tentatively hypothesize a word-structure projection (ω) and semi-formally express this as a lexical property of the construct form: (ω(↑) MORPHTYPE)=BOUND. Finally, the postnominal NP is what the construct form is bound to.

(17) $\text{NP} \rightarrow \underset{\uparrow = \downarrow}{\text{N}} \quad \underset{\substack{(\uparrow \text{DOM}) = \downarrow) \\ \left\{ \substack{(\uparrow \text{POSS}) = \downarrow \\ \downarrow \in (\uparrow \text{ADJ})} \right\}}}{\text{NP}} \quad \underset{\downarrow \in (\uparrow \text{ADJ})}{\text{AP*}}$

We assume that construct state nouns are lexically marked to require the DOM attribute; by default, nouns not explicitly marked to require the DOM attribute forbid it.

(18) construct nouns: ($\uparrow$ DOM)
free nouns: ¬ ($\uparrow$ DOM) (by default)

We hypothesize a lexical rule under which nominals which require a DOM inherit definiteness from it. This appears to be a manifestation of the morphological head-like quality of DOM.

(19) **Definiteness Dependency**
($\uparrow$ DOM) $\Rightarrow$ ($\uparrow$ DEF) = ($\uparrow$ DOM DEF)

We note in passing that this account of definiteness dependency is probably oversimplified. In the first place, the duplication of a feature of a syntactic dependent on the syntactic head is unexpected. Second, similar phenomena are apparent in languages where the morphologically dominant/dependent relation cannot be argued for. For example, in Welsh (Sadler 2000) an adjective may intervene between the head and the POSS and there is no morphophonological reflex of the "construct state" construction on the head noun. Despite the total lack of evidence for the kind of morphological relation that obtains in Hebrew, Welsh exhibits the same definiteness inheritance as Hebrew.

(20) llun rhyfedd y ferch
picture strange the girl
'the strange picture of the girl'

In addition, as observed by Dobrovie-Sorin (2001), definiteness effects related to possessors can be observed in languages like English and Rumanian. Note the following examples from Rumanian.[5]

(21) a. casa vecinului
house.the neighbor.the
'the neighbor's house'

b. *o casă vecinului
a house neighbor.the
'a house of the neighbor'

c. o casă a vecinului
a house of neighbor.the
'a house of the neighbor'

These cross-linguistic facts need to be examined in more detail; Dobrovie-Sorin proposes a semantic account of definiteness inheritance. Of course, the existence of semantic motivation does not rule out a lexical specification of the kind we are positing here. It is possible that a better understanding of the grammatical functions involved in possessor constructions might provide a better analysis of how these

[5]Rumanian also requires adjacency between the head and the possessor in the "construct."

languages are similar and differ.

The lexical entry for (the possessed version of) *ginat* is:

(22) *ginat* N $(\uparrow \text{PRED}) = \text{`garden} \langle(\uparrow \text{POSS})\rangle\text{'}$
$(\uparrow \text{NUM}) = \text{SG}$
$(\uparrow \text{GEND}) = \text{F}$
$(\uparrow \text{DOM})$
$(\uparrow \text{DEF}) = (\uparrow \text{DOM DEF})$

Note that since the construct form has an equation specifying a DEF value, it cannot be prefixed with the definite *ha*-. The f-structures of *ginat hamore hametupaxat* (5b) and *bigdey yeladim* 'children's clothes' are as follows.[6]

(23) a.

$$\left[\begin{array}{ll} \text{PRED} & \text{`garden } \langle(\uparrow \text{POSS})\rangle\text{'} \\ \text{GEND} & \text{F} \\ \text{NUM} & \text{SG} \\ \text{DEF} & + \\ \text{POSS} & \left[\begin{array}{ll} \text{PRED} & \text{`teacher'} \\ \text{GEND} & \text{M} \\ \text{NUM} & \text{SG} \\ \text{DEF} & + \end{array}\right] \\ \text{DOM} & \\ \text{ADJ} & \{[\text{PRED `cared-for'}]\} \end{array}\right]$$

b.

$$\left[\begin{array}{ll} \text{PRED} & \text{`clothing'} \\ \text{GEND} & \text{M} \\ \text{NUM} & \text{PL} \\ \text{DEF} & - \\ \text{ADJ} & \left\{\left[\begin{array}{ll} \text{PRED} & \text{`child'} \\ \text{GEND} & \text{M} \\ \text{NUM} & \text{PL} \\ \text{DEF} & - \end{array}\right]\right\} \\ \text{DOM} & \end{array}\right]$$

We can extend this analysis to include the following noun phrase types.

(24) a.

ginat-	o	ha-	metupax-	at
garden-	his	the-	cared.for-	FSG

'his tended garden'

b.

ginat-	o	ha-	metupax-	at	šel	ha-	more
garden-	his	the-	cared.for-	FSG	of	the-	teacher

'the teacher's tended garden'

The suffix -*o* is a POSS agreement suffix (Engelhardt 1998). As an agreement morpheme, it can only cross-reference arguments, not adjuncts; this is a restriction noted by Engelhardt. As with other agreement affixes, a POSS agreement affix can optionally function as an incorporated pronoun (Bresnan 2001 and references cited there). In LFG, this means that it has an optional [PRED 'PRO'] feature. Unlike the

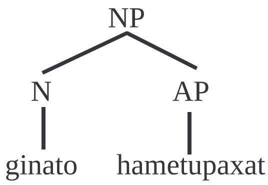
[6]The curved lines in the f-structures indicate that the same f-structure element is the value of more than one attribute (i.e. has more than one function).

construct form from which it is derived, the form with the agreement suffix is not a morphologically bound form; it does not take a DOM element. As a result, when the agreement suffix is not pronominal, the POSS is realized as a *šel* phrase rather than a bare NP.[7] Finally, as observed by Engelhardt (1998), the suffixed form is inherently definite.

(25) *ginato* N $(\uparrow \text{PRED}) = \text{‘garden} \langle(\uparrow \text{POSS})\rangle\text{’}$
$(\uparrow \text{NUM}) = \text{SG}$
$(\uparrow \text{GEND}) = \text{F}$
$(\uparrow \text{DEF}) = +$
$\neg (\uparrow \text{DOM})$
$(\uparrow \text{POSS PERS}) = 3$
$(\uparrow \text{POSS NUM}) = \text{SG}$
$(\uparrow \text{POSS GEND}) = \text{M}$
$((\uparrow \text{POSS PRED}) = \text{‘PRO’})$

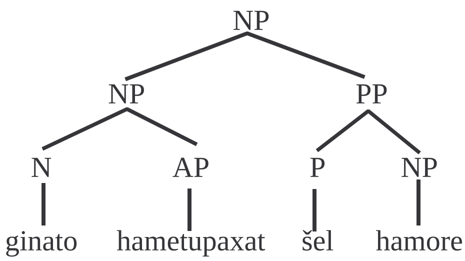
The c- and f-structures for the NPs in (24) are:

(26) a.



$$\left[\begin{array}{ll} \text{PRED} & \text{‘garden} \langle(\uparrow \text{POSS})\rangle\text{’} \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{F} \\ \text{DEF} & + \\ \text{POSS} & \left[\begin{array}{ll} \text{PRED} & \text{‘PRO’} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{M} \end{array}\right] \\ \text{ADJ} & \{[\text{PRED} \quad \text{‘cared-for’}]\} \end{array}\right]$$

b.



$$\left[\begin{array}{ll} \text{PRED} & \text{‘garden} \langle(\uparrow \text{POSS})\rangle\text{’} \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{F} \\ \text{DEF} & + \\ \text{POSS} & \left[\begin{array}{ll} \text{CASE} & \text{POSS} \\ \text{PRED} & \text{‘teacher’} \\ \text{DEF} & + \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{M} \end{array}\right] \\ \text{ADJ} & \{[\text{PRED} \quad \text{‘cared-for’}]\} \end{array}\right]$$

[7]Although stated differently, this is essentially Engelhardt’s analysis. She shows convincingly that the usual GB/MP analysis under which the suffix is a pronominal clitic which absorbs the abstract genitive Case associated with the construct is untenable, and proceeds to argue for agreement status, as assumed here. She attributes the inability of the suffixed form to appear in the construct to a requirement that the head of a construct must be unmarked for definiteness. She does not show how this intuitive idea can be formalized in the MP framework within which her analysis is developed, nor is it clear how we could formalize it in LFG. However, taking the lack of inherent definiteness as a symptom of the bound-morpheme character of construct-state nouns, we can reconstitute Engelhardt’s insight as we have done here.

3. Action Nominals

3.1. Overview

Abstracting away from the different methods of realizing a POSS, there are basically two ways for a Hebrew action nominal to express its arguments. In one, the subject argument of the corresponding verb is realized as a POSS and the object argument as an accusative-marked phrase. In the other, it is the object argument that is projected into the syntax as POSS, with the subject argument surfacing optionally as a *by* phrase.

(27)	a.	sgirat	ha-	mankal	et	ha-	misrad
		closure.CONSTR	the-	director	ACC	the-	office
	b.	sgirat	ha-	misrad	(alyedey	ha-	mankal)
		closure.CONSTR	the-	office	(by	the-	director)

‘the closure of the office by the director’

We will refer to these as the “accusative” and “non-accusative” forms of the action nominal. The non-accusative form is the more common version.

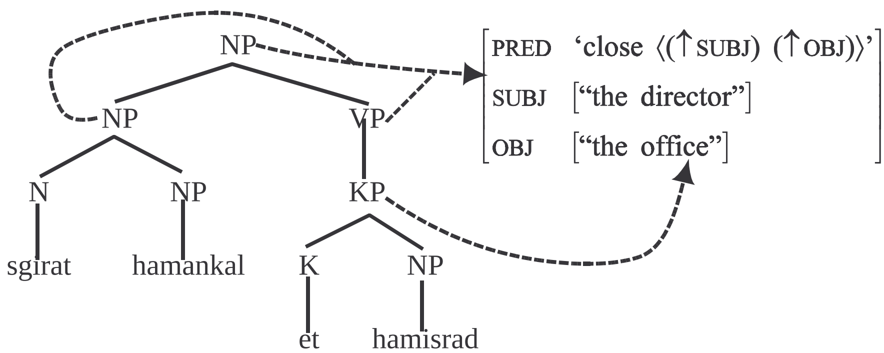
Despite the higher naturalness of the non-accusative form of the action nominal, it is the accusative form that has attracted much attention in the literature. Since accusative Case is otherwise attested only in the objects of verbs, it has invited an analysis in which the structure of the NP includes an embedded VP. In a derivational framework, the most straightforward way to implement this is to insert the nominalized verb in V position and for it to undergo head-movement to a higher N. Such analyses have been argued for by, inter alia, Hazout (1995) and Engelhardt (1998). On the other hand, Siloni (1997) argues against this kind of analysis. We will argue here for a nonderivational version of the NP-over-VP analysis of Hebrew action nominals.

3.2. Hebrew Action Nominals as a Mixed Category

3.2.1. The Analysis: Accusative Nominals

The LFG theory of mixed categories is based on the concept of head-sharing, the LFG equivalent of head-movement. A construction which would be analyzed as head movement in a derivational framework (say, V-to-I movement) can be thought of in terms of two phrases (an IP and a VP) sharing a head. In a head-sharing construction, the shared head is located in the highest of the head-sharing phrases: in I rather than V, for example. A construction like the Hebrew action nominal can be similarly analyzed, as shown in the following c-structure and partial f-structure.[8]

(28)



The noun *sgirat* is the head of both the root NP (of which it is the $\bar{X}$ head), and of the headless VP embedded in it. The head sharing is a result of the fact that the NP and VP map to the same f-structure

[8]I am assuming that the accusative particle *et* belongs to the category K (Case), which is a functional head. Nothing hinges on this.

element, as shown by the arrows indicating the mapping. We assume that the VP is adjoined to NP.

(29)
$$\text{NP} \rightarrow \begin{array}{cc} \text{NP} & \text{VP} \\ \uparrow=\downarrow & \uparrow=\downarrow \end{array}$$

The relative acceptability of coordinating N+POSS in this construction suggests that this structure is correct. (An explanation is still needed for why informants tend to stop short of identifying it as fully grammatical; one possibility is that the relatively marked status of the accusative form of the nominal makes it harder to process, and the coordination adds to the processing load.)[9]

(30)

?ibud	ha-	mumxim	ve-	hašmadat	ha-	politikaim	et
processing	the-	experts	and-	destruction.CONSTR	the-	politicians	ACC

pitkey hahacbaa	be-	xodeš	november
the.ballots	in-	month	November

‘the experts’ processing of the ballots and the politicians’ destruction of the ballots in November’

Mixed categories result from some cases of category-changing morphology, on a language-specific basis. In addition to Hebrew (and the related construction in Arabic), Bresnan (1997) cites examples from Italian, Kikuyu, German, Japanese, and Dagaare. (In these examples, the head is bolded and labeled with its lexical category.)

(31) a. **Italian infinitive** (NP-over-VP)

$[_{\text{DP}}$ il	$[_{\text{NP}}$ suo	continuo	**mormare**$_{\text{N}}$	$[_{\text{VP}}$ [parole	dolci] $_{\text{VP}}]$ $_{\text{NP}}]$ $_{\text{DP}}]$
the	his/her	continual	whisper.INF	words	sweet

‘his continual whispering of soft words’

b. **Kikuyu agentive nominalization** (NP-over-VP)

$[_{\text{DP}}$ $[_{\text{NP}}$ **mũ-**	**thĩĩnj-**	**i**$_{\text{N}}$	$[_{\text{VP}}$ [mbũri] $_{\text{VP}}]$ $_{\text{NP}}]$	ũyũ	$_{\text{DP}}]$
1-	slaughter-	NMNL	10.goat	1.DEM	

‘this slaughterer of goats’

c. **German adjectival participle** (AP-over-VP)

ein	$[_{\text{AP}}$ $[_{\text{VP}}$ [mehrere	Sprachen] $_{\text{VP}}]$	**sprechender**$_{\text{A}}$	$_{\text{AP}}]$	Mann
a	several	languages	speaking.NOM.MSG		man

‘a man speaking several languages’

d. **Japanese verbalized nominalization** (VP-over-NP)

$[_{\text{S}}$ Taroo	ga	$[_{\text{VP}}$ $[_{\text{NP}}$ [kinmedaru	no] $_{\text{NP}}]$	**morai-**	**ta-**	**sa-**	**no**$_{\text{V}}$ $_{\text{VP}}]$ $_{\text{S}}]$
Taro	NOM	goldmedal	GEN	receive-	want-	NMNL-	COP

‘Taro’s desire to receive a gold medal’

e. **Dagaare action nominal** (NP-over-VP)

$[_{\text{DP}}$ a	$[_{\text{NP}}$ Dɛre	$[_{\text{VP}}$ [ga-	ma]	[wiewie] $_{\text{VP}}]$	velaa	**sɔr-**	**oo**$_{\text{N}}$ $_{\text{NP}}]$ $_{\text{DP}}]$
the	Dɛre	book-	PL	quickly	good	read-	NMNL

‘the nice way of Dɛre’s reading books quickly’

They result from a mixed argument structure; in the case of a noun/verb mixture, an argument structure

[9]Note that this is not Right Node Raising, as the “shared” material is not a single constituent under anybody’s analysis.

with verbal and nominal characteristics (Bresnan 1997; Bresnan and Mugane 2000). The verbal argument structure requires that its c-structure correspondent be a VP, and the nominal argument structure requires that its c-structure correspondent be an NP.[10]

(32) a. $(\uparrow \text{PRED}) = \text{'close} \langle \langle x, y \rangle_v \rangle_n$

b. *v:* $\text{VP} \in \lambda (\phi^{-1} (\uparrow))$

n: $\text{NP} \in \lambda (\phi^{-1} (\uparrow))$

As a result, both nominal and verbal projections appear in the c-structure. The word itself is a noun, and thus occupies the structural position of the $\bar{X}$ head of the NP. The theory of head-sharing stipulates that the head occupy the head position of the highest projection, so the NP dominates the VP. As for the appearance of the subject of the action nominal in positions normally reserved for the POSS function, the answer may lie in the relation between the POSS and SUBJ functions; for concreteness, we can follow the proposal in Bresnan (2001) for English gerunds, and assume that Hebrew action nominals include the following specification:

(33) $(\uparrow \text{POSS}) = (\uparrow \text{SUBJ})$

The full f-structure of (28) is:

(34)
$$\begin{bmatrix} \text{PRED} & \text{'close} \langle (\uparrow\text{SUBJ})\ (\uparrow\text{OBJ}) \rangle\text{'} \\ \text{GEND} & \text{F} \\ \text{NUM} & \text{SG} \\ \text{DEF} & + \\ \text{DOM} & \begin{bmatrix} \text{PRED} & \text{'director'} \\ \text{GEND} & \text{M} \\ \text{NUM} & \text{SG} \\ \text{DEF} & + \end{bmatrix} \\ \text{POSS} & \\ \text{SUBJ} & \\ \text{OBJ} & \begin{bmatrix} \text{CASE} & \text{ACC} \\ \text{PRED} & \text{'office'} \\ \text{GEND} & \text{M} \\ \text{NUM} & \text{SG} \\ \text{DEF} & + \end{bmatrix} \end{bmatrix}$$

We thus account for the properties of the accusative form of the Hebrew action nominal.

3.2.2. Non-accusative Action Nominals

The mixed-category analysis of the accusative version of action nominals raises questions concerning the analysis of the non-accusative form. One analysis, based primarily on the optional appearance of the *by* phrase, and also on the object argument appearing as a "subject", is that the non-accusative form is a passive (Rosén 1977, Engelhardt 1998). However, as discussed by Siloni (1997), there are serious problems with such an analysis, such as the lack of relation between which verbs undergo passivization and which action nominals can appear in the non-accusative version.

We will add an additional consideration which argues against a passive analysis. Of the two forms of the action nominal in Hebrew, the non-accusative form is by far the more natural. The accusative form has been described as being not as common (Rosén 1977), "an occasional formal usage" (Glinert 1989), and "stylistically highly marked"(Ritter 1988). There are even speakers of Hebrew who do not accept the

[10]λ is the category labelling function.

accusative version. This contrasts sharply with the relation between active and passive in Hebrew: the passive is a highly marked construction which is not used much in natural Hebrew speech.

We will argue for a different analysis of the non-accusative version, based on the LFG theory of argument mapping. Argument mapping in LFG is mediated by argument structure (a-structure), a level of representation in which argument positions are classified by a system of distinctive features for grammatical functions. The theory of mapping, called Lexical Mapping Theory (LMT), posits the following system of features:

(35) a. restricted
[+*r*]: OBL$_\theta$, OBJ$_\theta$ (or OBJ2)
[–*r*]: SUBJ, OBJ

b. objective
[+*o*]: OBJ, OBJ$_\theta$
[–*o*]: SUBJ, OBL$_\theta$

Arguments with Theme-like and Patient-like thematic roles are classified as [–*r*] and those with other thematic roles (including Agent) are classified as [–*o*]. In the mapping to f-structure, remaining features are filled in; if possible, the thematically most prominent argument becomes SUBJ. Two arguments cannot map to the same grammatical function. As a result, in a typical transitive Agent-Patient verb, the Agent becomes SUBJ and the Patient, which is inherently classified as [–*r*], has to become OBJ.

As discussed by Laczkó (2000), the feature [+*o*] is not available to the arguments of nominals. This affects the possibilities for argument mapping in nominals. Assuming that the function POSS is the nominal equivalent of SUBJ, and thus [–*r*, –*o*], we find the following.

(36) thematic roles Agent Theme/Patient

a-structure ⟨[–*o*] , [–*r*]⟩

GFS *⟨[–*r*] ?⟩ ⟨[+*r*] [–*o*]⟩
POSS OBL$_\theta$ POSS

A verb-like mapping, in which the Agent is mapped to the SUBJ-like function POSS, leaves no grammatical mapping for the Patient. A “passive” mapping, in which the Patient maps to POSS, results in an oblique realization of the Agent; without any additional stipulations, this is the mapping in a nominal. In some languages there are additional stipulations: in Modern Greek an unergative argument cannot be realized in nominals (Markantonatou 1995) and in English a [+*r*] mapping is allowed for the Patient, thus permitting the Agent to be POSS (Laczkó 2000). However, in the simplest case, instantiated in Hungarian according to Laczkó, the lexical form of the nominal of a transitive verb will be one in which the Agent is an oblique (often optional, since oblique arguments of nominals are usually optional) and the Patient is POSS.

(37) ‘close ⟨∅/(↑ OBL$_{Agent}$) (↑ POSS)⟩

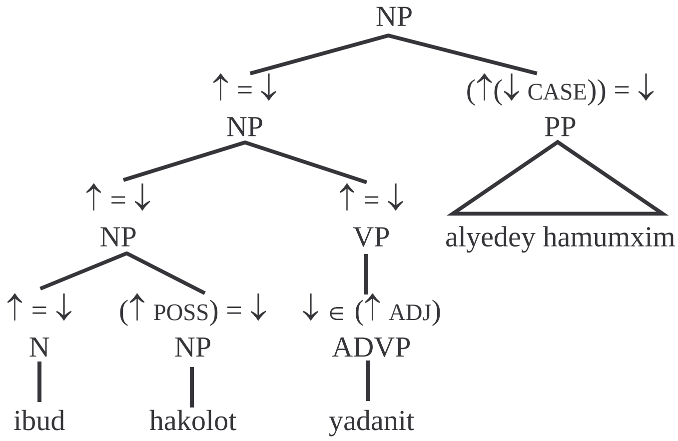
The lexical form in (37) is exactly what is needed to generate the non-accusative version of the action nominal in Hebrew. The simplest analysis of the non-accusative action nominal in Hebrew is thus that it involves a nominal mapping of arguments, while the accusative version involves a verbal mapping.

Even when it maps its arguments as a noun, the action nominal is a mixed category. Modification by both adjectives and adverbs is possible.

(38)	a.	ibud	ha-	kolot	yadanit	alyedey	ha-	mumxim	
		processing	the-	votes	manually	by	the-	experts	
	b.	ibud	ha-	kolot	ha-	yadani	alyedey	ha-	mumxim
		processing	the-	votes	the-	manual	by	the-	experts

‘the manual processing of the votes by the experts’

c. c-structure of (a) with partial functional annotations



The proposed analysis of the non-accusative form is natural under a lexicalist implementation of the mixed category analysis but not under a derivational implementation. From the lexicalist perspective, there is something natural about something with mixed nominal/verbal argument structure being able to map its arguments either as a noun or as a verb. Given the fact that the nominals themselves are lexically nouns, it is also not unexpected that the nominal mapping of arguments will be the less marked one. This explains the intuition of unnaturalness that many Hebrew speakers attribute to the accusative form. On the other hand, it is hard to see how the derivational approach could adopt this analysis. The derivational approach treats the VP as a projection of the (nominalized) verb, with its projected (internal) arguments realized within the VP. While the nominalization is accounted for by raising the verb to N, there is no reason for the arguments to move and become nominal arguments. There is no way to capture the idea that in the non-accusative form of the nominalizations the arguments are realized like the arguments of nouns, nor is it clear how the greater naturalness of the non-accusative form could be explained.

3.3. Alternatives

The analysis we have presented is the lexicalist equivalent of what has become the conventional analysis of Hebrew action nominals. However, this kind of analysis has been argued against, both for Hebrew and universally. We will discuss Siloni’s (1997) argument that Hebrew action nominals are purely nouns, as well as Malouf’s (1998) lexicalist (HPSG) analysis of mixed categories as intermediate between (in this case) noun and verb instead of having distinct nominal and verbal structural sections.

Siloni (1997) argues against a mixed-category analysis for Hebrew action nominals, arguing that they are pure nouns. At the conceptual level, she questions the possibility of mixed categories. However, since they can be shown to exist in other languages they cannot be ruled out in principle. She also argues, on empirical grounds, that action nominals do not take adverbs and that the accusative Case on the object is “inherent” accusative Case.

Siloni’s claim that action nominals do not take adverbs is based on contrasts such as the following.

(39) a. mexikat ha- maxšev et ha- kvacim bi- mhirut
erasing.CONSTR the- computer ACC the- files in- speed
b. *mexikat ha- maxšev et ha- kvacim maher
erasing.CONSTR the- computer ACC the- files quickly
'the computer's quick erasing of the files'

As Siloni correctly observes, forms like *bimhirut*, which are often cited as adverbs, may be better analyzed as PPs (note the gloss). She shows that the adverb *maher* cannot cooccur with action nominals, and argues that this inability is evidence against any analysis involving an embedded VP. However, in the examples in (2) and (38a) we have seen cases of other adverbs, *yadanit* 'manually' and *zmanit* 'temporarily', which can occur in action nominal phrases. We repeat (2a) and (38a) here.

(40) a. ibud ha- mumxim et kolot yadanit
processing the- experts ACC votes manually
'the experts' manual processing of the votes'
b. ibud ha- kolot yadanit alyedey ha- mumxim
processing the- votes manually by the- experts
'the manual processing of the votes by the experts'

The most direct way to express the fact that some adverbs (those formed regularly from adjectives) can occur in action nominal phrases while others (with irregular morphology) cannot is to hypothesize that there is a VP in the structure of the action nominal, but that adverbs like *maher* are lexically marked not to appear in phrases which are co-heads of NPs.

(41) $\mathrm{NP} \notin \lambda\,(\phi^{-1}\,(\textsc{adj}\uparrow))$

Siloni provides several arguments that the accusative Case which appears on the object of an action nominal is not the normal accusative Case, and therefore provides no evidence for an analysis in which the action nominal has verbal properties. The burden of proof is clearly on the kind of an analysis which Siloni argues for: as noted by Hazout (1995), the accusative Case which surfaces on the objects of action nominals seems to have all the properties of accusative Case. Most salient among Siloni's arguments are the ungrammaticality of indefinite objects and the ungrammaticality of pronominal objects. In Hebrew, as in many other languages, accusative Case only appears on OBJs when they are definite; indefinite OBJs are not overtly marked with Case. Siloni argues that the accusative Case in nominals is "inherent" accusative Case, and thus an object on which accusative cannot surface is of necessity ungrammatical (42a). However, as she herself notes, heavier NPs make the construction significantly more acceptable (42b).

(42) a. *ibud mumxim pitkey.hacbaa
processing experts ballots
'experts' processing of votes'

b. ?ibud mumxim mi- florida pitkey.hacbaa šel kšišim yehudim
processing experts from- Florida ballots of senior.citizens Jewish
'processing by experts from Florida of the ballots of Jewish senior citizens'

This casts doubt on the "inherent Case" analysis of accusative in nominals. A more plausible analysis would be to attribute the reduced acceptability of (42) to a difficulty in parsing what is already a highly marked construction.

As for the ungrammaticality of pronominal objects, Siloni adds the observation that dative pronouns are also excluded from action nominals. This point is expanded on by Engelhardt (1998), who notes that even oblique pronouns are excluded. Engelhardt proposes, plausibly, that pronouns need to be adjacent the head. They are not adjacent to the head in the action nominal because the POSS/SUBJ intervenes.

Finally, again as noted by Engelhardt, the fact that the appearance of an overt accusative Case particle in nominals is governed by definiteness, just as it is with verbs, suggests strongly that the accusative Case is the usual grammatical accusative found on the objects of verbs.

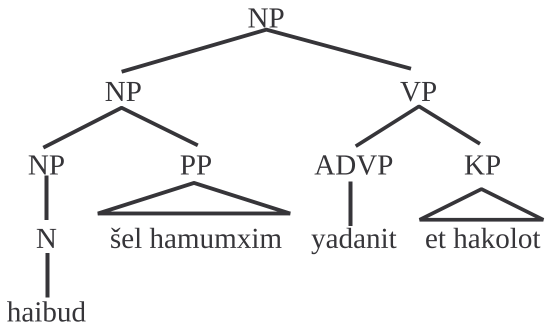
Malouf (1998) proposes an analysis of mixed categories which involves not distinct categorial projections but rather a single category having partial properties of both. In his HPSG analysis of English gerunds, he proposes that the HEAD value *gerund* is a subtype of both *noun* (other subtypes of which are *common noun* and *proper noun*) and *relational* (other subtypes of which are *verb* and *adjective*. He claims that since the external distribution of NPs is based on the type *noun*, mixed categories like gerunds will have the same distribution as NPs. Similarly, since adverbs modify relational elements and adjectives only modify common nouns, gerunds are modified only by adverbs. The verb-like complements of gerunds are attributed in his analysis to the lexical rule deriving gerunds from verbs, which retains the COMPS structure.

There is a fundamental difference between a c-structural two-category approach to mixed categories and Malouf's intermediate category approach. A structurally mixed account predicts that mixed categories should exhibit what Malouf calls "phrasal coherence," whereby the nominal and verbal aspects of the mixed category occupy distinct regions of the c-structure. Hebrew action nominals exhibit phrasal coherence.[11]

(43) a. ?ha- ibud šel ha- mumxim yadanit et ha- kolot
the- processing of the- experts manually ACC the- votes
b. **ha- ibud yadanit šel ha- mumxim et ha- kolot
the- processing manually of the- experts ACC the- votes
'the experts' manual processing of the votes'

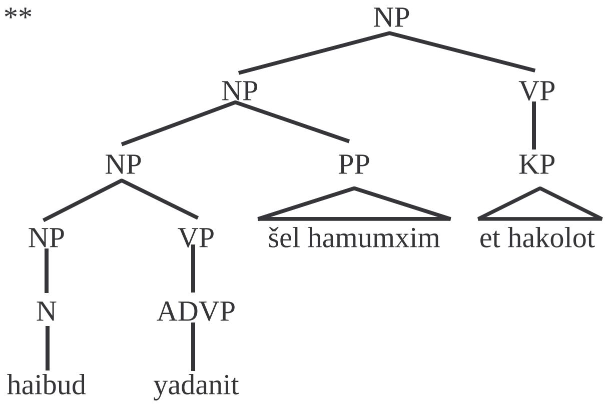
The c-structures of these NPs are as follows.

(44) a.



[11]Speakers of Hebrew are generally not very happy about the preverbal position of the adverb, but for those who do accept it the non-phrasally-coherent version is ungrammatical, and for those who do not the non-phrasally-coherent version is much worse.

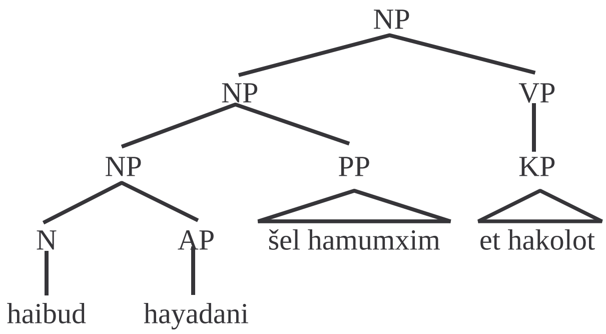
b. **



In the ungrammatical tree, the VP parts and NP parts are interspersed, violating phrasal coherence. On the other hand, Malouf's approach suggests that noun/verb mixtures should be modifiable by adverbs but not adjectives, on the grounds that adjectives only modify common nouns. All things being equal, this should apply to Hebrew action nominals as well.[12] Yet, both adjectival and adverbial modification are possible. On the other hand, given the phrase structure rules we have hypothesized for Hebrew, such a structure is predicted to be grammatical.[13]

(36) a. ?ha- ibud ha- yadani šel ha- mumxim et ha- kolot
the- processing the- manual of the- experts ACC the- votes
'the experts' manual processing of the votes'

b.



c. ha- ibud ha- yadani šel ha- kolot alyedey ha- mumxim
the- processing the- manual of the- votes by the- experts
'the manual processing of the votes by the experts'

This pattern of allowing both adverbs and adjectives seems to be somewhat unusual; it is the opposite of what Bresnan and Mugane (2000) report for Kikuyu, for example, where mixed-category agent nominalizations take adverbs and not adjectives. However, the c-structure-based theory of mixed categories is more capable of accommodating such properties than Malouf's theory.

---

[12]All things are not necessarily equal, since Malouf's theory allows different feature hierarchies in different languages.

[13]Siloni describes a nominal in which both arguments are expressed (in either the accusative or non-accusative form) as "somewhat clumsy and marginal" (p. 51), a description which matches my informants' judgments more for the accusative form than the non-accusative form. Since we have argued that the non-accusative form is "more nominal" than the accusative form, this clumsiness may be a result of the stronger verb-like properties of the accusative version. There may even be speakers who reject an adjective in the accusative form completely; however, the non-accusative form clearly allows both adjectives and adverbs.

It also should be noted, as observed by Siloni, that action nominals are completely and unambiguously nominal in their morphological properties. For example, they take the nominal negative prefix *i-*, which never appears on verbs. This is to be expected in a structurally mixed analysis, where the word itself has to be identified with a specific category, but not from an intermediate-category analysis of the kind Malouf argues for.

4. Conclusion

We have shown that the parallel architecture of LFG allows us to express the properties of Hebrew NPs in a constrained theory of constituent structure: one in which categories not lexically motivated cannot be assumed to exist and in which constituent structure expresses distributional properties rather than functional ones.

We have also shown that deverbal action nominals have NP-over-VP structures, as proposed in some of the literature. However, a lexicalist mixed-category implementation of this analysis along the lines of Bresnan (1997) has the advantage of allowing the nominal to project its arguments into the syntax either as a verb or as a noun. This provides a superior account of the two patterns of argument realization that are found. It also provides a natural explanation of the less marked status of the non-accusative version of the Hebrew action nominal.

References

Austin, Peter, and Joan Bresnan (1996) "Non-configurationality in Australian Aboriginal Languages." *Natural Language and Linguistic Theory* 14: 215–268.

Bresnan, Joan, ed. (1982) *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: MIT Press.

Bresnan, Joan (1997) "Mixed Categories as Head-sharing Constructions." in Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG97 Conference, University of California, San Diego*. On-line: CSLI Publications. `http://csli-publications.stanford.edu/LFG/2/lfg97.html`.

Bresnan, Joan (2001) *Lexical-Functional Syntax.* Oxford: Blackwell.

Bresnan, Joan, and John Mugane (2001) "Mixed Categories in Gĩkũyũ." presented at LFG00, University of California, Berkeley.

Dobrovie-Sorin, Carmen (2001) "(In)definiteness Spread: From Rumanian Genitives to Hebrew Construct State Nominals."

Engelhardt, Miriam (1998) *The Syntax of Nominalized Properties.* Doctoral dissertation, The Hebrew University of Jerusalem

Falk, Yehuda N. (2001) *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax.* Stanford, Calif.: CSLI Publications.

Glinert, Lewis (1989) *The Grammar of Modern Hebrew.* Cambridge: Cambridge University Press.

Hazout, Ilan (1995) "Action Nominalizations and the Lexicalist Hypothesis." *Natural Language and Linguistic Theory* 13: 355–404.

Laczkó, Tibor (2000) "Derived Nominals, Possessors, and Lexical Mapping Theory in Hungarian DPs." in Miriam Butt and Tracy Holloway King, eds., *Argument Realization* Stanford, Calif.: CSLI Publications. 189–227.

Malouf, Robert P. (1998) *Mixed Categories in the Hierarchical Lexicon.* Doctoral dissertation, Stanford University

Markantonatou, Stella (1995) "Modern Greek Deverbal Nominals: An LMT Approach." *Journal of Linguistics* 31: 267–299.

Nordlinger, Rachel (1998) *Constructive Case: Evidence from Australian Languages.* Stanford, Calif.: CSLI Publications.

Ritter, Elizabeth (1988) “A Head-Movement Approach to Construct-State Noun Phrases.” *Linguistics* 26: 909–929.
Rosén, Haiim (1977) *Good Hebrew: Studies in Syntax (Ivrit Tova: Iyunim Betaxbir)*, Third Edition, Corrected And Expanded. Jerusalem: Kiryat Sefer.
Sadler, Louisa (2000) “Noun Phrase Structure in Welsh.” in Miriam Butt and Tracy Holloway King, eds., *Argument Realization* Stanford, Calif.: CSLI Publications. 73–109.
Siloni, Tal (1997) *Noun Phrases and Nominalizations: The Syntax of DPs.* Dordrecht: Kluwer.
Wintner, Shuly (2000) “Definiteness in the Hebrew Noun Phrase.” *Journal of Linguistics* 36: 319–363.

GlueTag
Linear Logic based Semantics for LTAG
– and what it teaches us about LFG and LTAG –

Anette Frank

Language Technology Group
DFKI GmbH
Saarbrücken, Germany
frank@dfki.de

Josef van Genabith

Computer Applications
Dublin City University
Dublin, Ireland
josef@compapp.dcu.ie

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu

GlueTag *
Linear Logic based Semantics for LTAG
– and what it teaches us about LFG and LTAG –

Abstract

We review existing appoaches to semantics construction in LTAG (Lexicalised Tree Adjoining Grammar) based on the notion of *derivation (tree)s*. We argue that derivation structures in LTAG are not appropriate to guide semantic composition, due to a non-isomorphism, in LTAG, between the syntactic operation of adjunction on the one hand, and the semantic operations of complementation and modification, on the other.

Linear Logic based "glue semantics", as developed within the LFG framework (cf. Dalrymple (1999)), allows for flexible coupling of syntactic and semantic structure. We investigate application of glue semantics to LTAG syntax, using as underlying structure the *derived tree*, which is more appropriate for principle-based semantics construction. We show how Linear Logic based semantics construction helps to bridge the non-isomorphism between syntactic and semantic operations in LTAG. The glue approach captures non-tree local dependencies in control and modification structures, and extends to the treatment of scope ambiguity with quantified NPs and VP modifiers. Finally, glue semantics applies successfully to the adjunction-based analysis of long-distance dependencies in LTAG, which differs significantly from the f-structure based analysis in LFG.

1 Introduction

In this paper we review existing appoaches to semantics construction in LTAG (Lexicalised Tree Adjoining Grammar) based on the notion of *derivation (tree)s*. We argue that LTAG derivation trees are not appropriate to guide semantic composition, due to a non-isomorphism, in LTAG, between the syntactic operation of adjunction on the one hand, and the semantic operations of complementation and modification, on the other.

Linear Logic based "glue semantics", by now the classical approach to semantics construction within the LFG framework (cf. Dalrymple (1999)) allows for flexible coupling of syntactic and semantic structure. We investigate application of glue semantics to LTAG syntax,[1] using as underlying structure the *derived tree*, which seems more appropriate for principle-based semantics construction. We show how Linear Logic based semantics construction helps to bridge the non-isomorphism between syntactic and semantic operations in LTAG. Glue semantics captures non-tree local dependencies in control and modification structures, and extends to the treatment of scope ambiguity with quantified NPs and VP modifiers. Finally, glue semantics applies successfully to the adjunction-based analysis of long-distance dependencies in LTAG, which differs significantly from the f-structure based analysis in LFG in terms of functional uncertainty.

On a more general perspective, the exercise is instructive in that it elucidates the role that f-structure plays in LFG syntax and semantics, and helps clarify the similarities and differences between the two frameworks.

The paper is structured as follows. In Section 2 we review basic assumptions of the LFG and LTAG frameworks to set the stage for our investigations. Section 3 examines previous approaches to semantics construction in LTAG based on derivation (tree)s, namely Shieber and Schabes (1990), Schabes and Shieber (1994), Joshi and Vijay-Shanker (1999) and Kallmeyer and Joshi (1999). In

*We are grateful for valuable comments from the audiences of the LFG01 conference and the University of Konstanz, in particular Ron Kaplan, Josef Bayer and Ellen Brandner. Thanks go also to Dick Crouch and Mary Dalrymple for comments on earlier versions of this paper. Some interesting observations could not be given full justice in this paper, but provided important feedback for the overall conception of this work, which we hope to extend in future research. This research was partially funded by a BMBF grant to the DFKI project WHITEBOARD (FKZ: 01 IW 002).

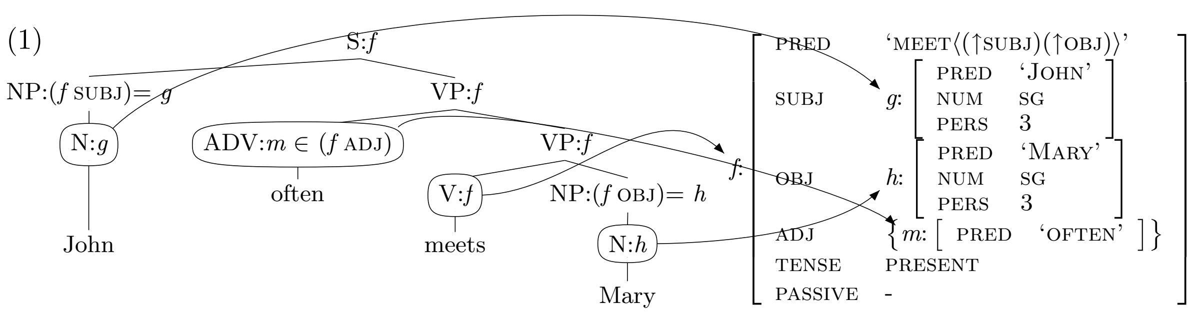
[1] Hepple (1999) sketches LL-based semantics for D-Trees, to overcome problems faced by categorial semantics in the analysis of quantification. Muskens (2001) develops a description-based syntax-semantics interface for LTAG, yet with extension to tree descriptions as used in D-Trees. We briefly discuss these related approaches in Section 4.7.

Section 4 we design LL-based semantics construction for LTAG on the basis of derived trees. In Section 4.1 we design labelling principles for LTAG elementary and derived trees as an interface to LL-based glue semantics. In Section 4.2 these principles are extended and refined to account for non-tree local dependencies and scope constraints, exemplified by modification structures. Section 4.3 summarises the specific assumptions introduced for glue semantics from LTAG derived trees. Section 4.4 shows that glue semantics successfully bridges the non-isomorphism between adjunction in syntax and corresponding operations in semantics. In Sections 4.5 and 4.6 we consider control and long-distance constructions which, in their syntactic analysis, differ considerably from the corresponding analyses in LFG. We show that LL-based semantics construction for LTAG straightforwardly extends to these more intricate cases. It is especially in the context of these constructions that differences and similarities between the two syntactic frameworks emerge most clearly. This is the topic of Section 5. Section 6 concludes.

2 Basic tenets of LFG and LTAG syntax and semantics

In LFG syntactic structure is represented in terms of two levels of syntactic description: c-structure and f-structure (1). Context-free PS rules with f-descriptions and lexical entries define the functional correspondence between c- and f-structure. Subcategorisation and long-distance dependencies are represented in f-structure, via functional descriptions. The correspondence function between c- and f-structure also accounts for word order variation. LFG semantics is driven by Linear Logic based meaning construction from f-structure, which allows for flexible coupling of syntax and compositional semantics. Lexical entries are associated with so-called meaning constructors. These consist of a "glue part", expressions in Linear Logic which refer to f-structure nodes, and a meaning part. The (instantiated) meaning constructors contributed by lexical entries are assembled as premises to a Linear Logic meaning derivation, based on the glue part. Following the Curry-Howard isomorphism, a meaning is computed, in parallel, on the meaning side. We assume familiarity with the glue semantics approach (see Dalrymple (1999), Dalrymple (2001) for more detail).

(1)



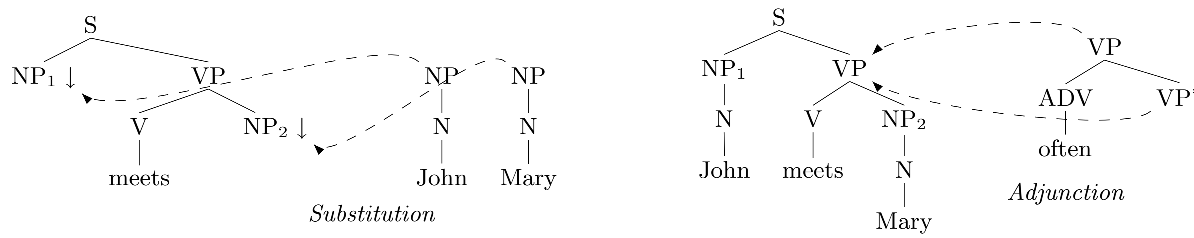
Lexical entries with associated meaning constructors			Instantiated meaning constructors
John	N	$(\uparrow \textsc{pred})=$ ‘JOHN’ $john : \uparrow_\sigma$	$john : g_\sigma$ $mary : h_\sigma$
Mary	N	$(\uparrow \textsc{pred})=$ ‘MARY’ $mary : \uparrow_\sigma$	$\lambda y,x.meet(x,y)$: $h_\sigma \multimap (g_\sigma \multimap f_\sigma)$ $\lambda P,x.often(P(x)) : (g_\sigma \multimap f_\sigma) \multimap (g_\sigma \multimap f_\sigma)$
meets	V	$(\uparrow \textsc{pred})=$ ‘MEET’ $\lambda y,x.meet(x,y)$: $(\uparrow \textsc{obj})_\sigma \multimap ((\uparrow \textsc{subj})_\sigma \multimap \uparrow_\sigma)$	
often	ADV	$(\uparrow \textsc{pred})=$ ‘OFTEN’ $\lambda P,x.often(P(x))$: $((\textsc{adj} \in \uparrow)\ \textsc{subj})_\sigma \multimap (\textsc{adj} \in \uparrow)) \multimap ((\textsc{adj} \in \uparrow)\ \textsc{subj})_\sigma \multimap (\textsc{adj} \in \uparrow))$	

Meaning derivation

$$\dfrac{\dfrac{\dfrac{\lambda y,x.meet(x,y) : h \multimap (g \multimap f) \quad mary : h}{\lambda x.meet(x,mary) : g \multimap f} \quad \lambda P,x.often(P(x)) : (g \multimap f) \multimap (g \multimap f)}{\lambda x.often(meet(x,mary)) : g \multimap f} \quad john : g}{often(meet(john,mary)) : f}$$

An LTAG grammar (Joshi 1987) consists of a set of lexicalised elementary trees (etrees), which are composed by two operations: substitution and adjunction (2). Elementary trees encode lexical syntactic properties: subcategorisation, agreement,[2] and syntactic variation in terms of tree families. The syntactic representation consists of the constituent tree (*derived tree*) built by substitution and adjunction of elementary trees (see (2), (3.a)), and a *derivation tree* (3.b), which records the dependencies between elementary trees as established by substitution and adjunction operations in parsing. LTAG semantics is traditionally based on the structure of *derivation trees*.

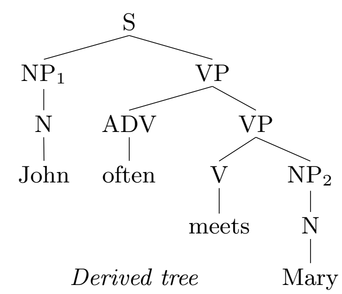
(2) Elementary trees, substitution and adjunction



Substitution *Adjunction*

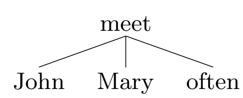
(3) Derived trees and derivation trees

(3.a)



Derived tree

(3.b)

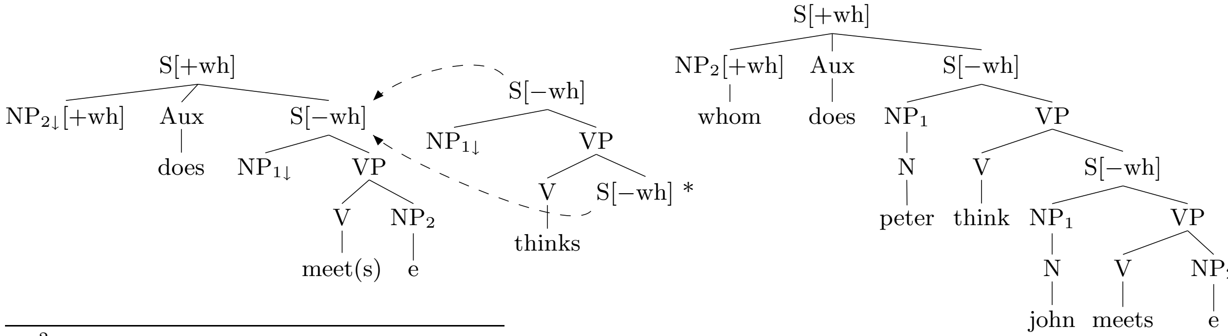


Derivation Tree

Extended Domains of Locality and Adjunction A central feature of LTAG syntax is *strict lexicalisation* combined with the concept of *Extended Domains of Locality*. Besides being *strictly lexicalised*, elementary trees *localise* all subcategorised arguments of the lexical head, representing them as substitution or foot nodes of the elementary tree. The joint assumptions of strict lexicalisation and localisation of arguments lead to *adjunction* as a major operation in syntactic composition. This is already evident in (2). Due to localisation of subject and object NPs in the etree *meets*, the derived tree for *John often meets Mary* can only be obtained by "folding in" the auxiliary tree for *often* into the etree of *meets* by use of the adjunction operation.

Besides optional (or recursive) modification structures, as in (2), localisation of arguments plays a central role in the analysis of long distance dependencies. In (4) the etree for *meets* locally encodes a fronted object wh-phrase. In order to derive sentence (4), the etree for *thinks* must again be "folded into" the etree of *meets* via adjunction. Thus, joint with the concept of localisation of arguments in strictly lexicalised elementary trees, the operation of adjunction naturally leads to the concept of *extended domains of locality*.

(4) *Whom does Peter think John meets?*



[2]Subject to feature constraints on nodes, not shown here.

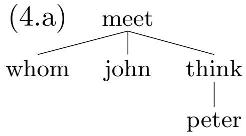
Derivation trees are not dependency trees Derivation trees record the relations between elementary trees as established by substitution and adjunction operations in tree composition, and are traditionally used as the basis for semantics construction in LTAG.

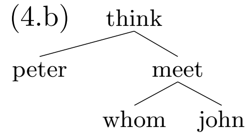
Derivation trees do not in general correspond to well-formed dependency structures. This was observed by Rambow et al. (1995), and is illustrated below. In the derivation tree (4.a) for sentence (4) the dependence of *think* upon *meet* is in fact inverted, as evidenced by the correct dependencies displayed in (4.b). One could argue that dependencies established by adjunction could be specially marked to account for such inverted dependencies, but more complex cases prove that this cannot, in general, lead to a well-formed dependency tree. (5.a) displays the derivation tree for (5).[3] Since *claim* and *seem* independently adjoin to the S and VP nodes of the etree *adore*, the derivation tree cannot represent the dependence of *seem* upon *claim*, as given in the correct dependency tree (5.b).

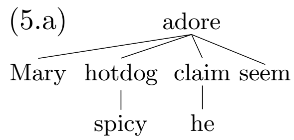
Note further that due to the principle of localisation of arguments, the operation of adjunction applies both to modifiers in (3) and to complementation structures such as sentence embedding verbs in (4).[4] As we shall see, this constitutes an additional complication for principle-based semantics construction from derivation trees.

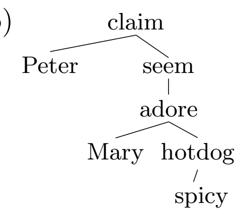
(4) Whom does Peter thinks John meets?

(5) Spicy hotdogs he claims Mary seems to adore.

(4.a)


(4.b)


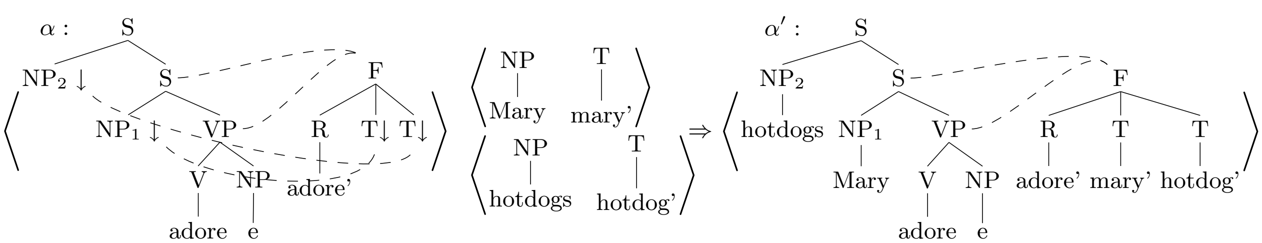
(5.a)


(5.b)


3 Semantics in LTAG

3.1 Shieber and Schabes: Semantics construction with Synchronous TAG

Shieber and Schabes (1990) associate LTAG syntax with a semantic representation in a synchronous TAG extension, where the grammar components are pairs of syntactic and semantic trees. The semantic representation, a tree-like logical form, is built in parallel with the syntactic derivation, making use of a specification of links between nodes in the paired tree components. On substitution of a tree t_1 into a substitution node n_1 in the syntactic tree, a parallel substitution takes place of the paired semantic tree t_2 into the node n_2 that n_1 is linked to. After substitution, the link being used is removed. This is illustrated for substitution of the trees for *Mary* and *hotdogs* into α below.

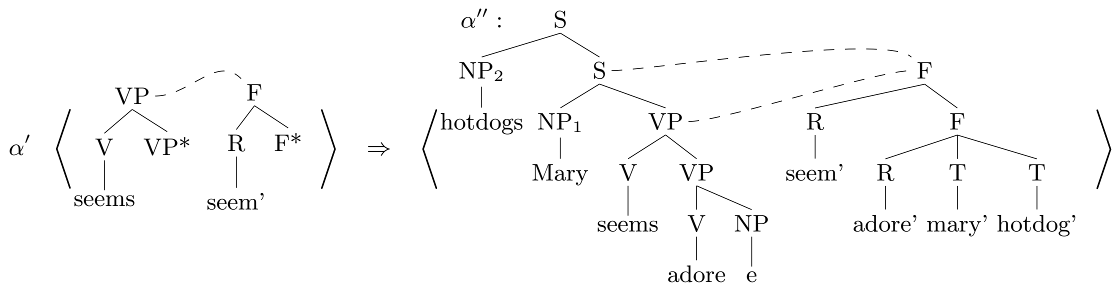


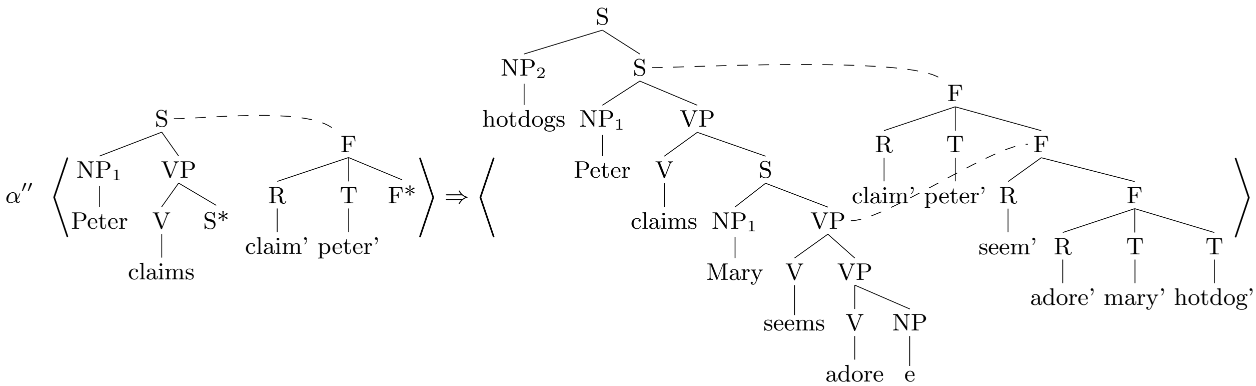
Crucial in this approach is the specification of links between paired syntactic and semantic trees, since they determine the attachment sites for the parallel semantic operations. Note in particular the link between the *tree internal* S and VP nodes of *adore* to the single root of the associated semantic tree. It is due to these links that the more complex cases of non-isomorphic derivations vs. dependencies as in (5) can be accounted for. The syntactic tree for *seems* can adjoin to the tree internal VP node of α', triggering a corresponding operation on the linked root node F in the paired semantic tree, which results in the correct scoping of *seem'* over *adore'*. Second, the syntactic tree

[3] The example is slightly changed from Rambow et al. (1995).

[4] As well as control constructions, see below.

for *claim* adjoins to the internal S node of the resulting syntactic tree in α'', leading to adjunction of its associated semantic tree to the root formula F with relation *seem*. The reader may verify that the same result is obtained for alternative derivations with *claim* being adjoined before *seem*.





However, as pointed out in Shieber and Schabes (1990), the order of derivations *can* have an effect on the semantic representation in that *different orders* of substitution of quantified NPs yields alternative scopings. Parsing must therefore compute all possible syntactic derivation histories, explicitly or implicitly (cf. Schabes and Shieber (1994)) in order to capture ambiguities that are essentially semantic. This is not only problematic conceptually, but in fact leads to spurious analyses in cases like (5) where the order of derivations is not distinctive for semantic interpretation.[5] Finally – in view of the following discussion – it is important to note that in this approach it is the linking of tree internal nodes in (elementary and derived) trees that accounts for cases of non dependency-like derivations, as in (5). This, however, characterises the approach as a hybrid one, in that semantics construction is essentially built on the structure of *derived trees*, while accounting for scope ambiguities in terms of *derivation histories* as determined by syntactic analysis.

3.2 Compositional Semantics from Derivation Trees

Joshi and Vijay-Shanker (1999) propose compositional semantics construction from derivation trees, focussing primarily on predicate-argument relations. Elementary trees are associated with tripartite semantic representations. The first part specifies the main variable of the predication, the second part states the predicate with argument variables, the third part associates variables with argument nodes in the elementary trees.[6] [7]

[5] Schabes and Shieber (1994) distinguish *modifier-type* from *predicative* auxiliary trees such as sentence embedding verbs (*say, claim*), where the foot node corresponds to an argument of the anchor. In contrast to the generally assumed notion of *standard derivations*, which excludes multiple adjunction to single nodes, Schabes and Shieber (1994) propose the notion of *extended derivations*, licensing multiple adjunction to single nodes. Allowing *extended derivations* for multiple adjunction of *modifier-type* auxiliary trees can yield alternative scopes in semantics construction, due to alternative derivation histories. For *predicative* auxiliary trees, however, *standard derivation* – i.e. the constraint against multiple adjunction at single nodes – is preserved. Scoping ambiguities are therefore correctly prohibited with cascaded sentence embeddings driven by adjunction.

[6] This association is not made explicit, but could be formalised by stating pairs of variables and the node addresses of the corresponding arguments in the elementary tree.

[7] For reasons of space we can only illustrate some selected entries (see continuation next page)

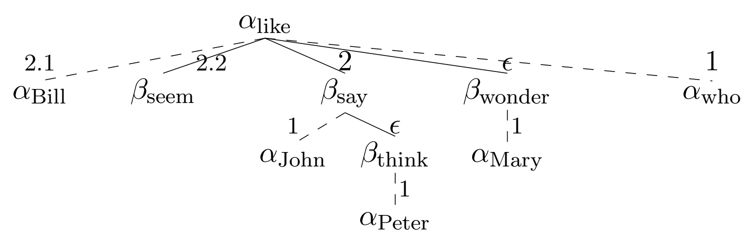
For substitution of NP arguments the binding of variables in the associated semantic representations is straightforward. For adjunction of predicative auxiliary trees (*say, think, wonder*), however, the derivation structure does not mirror semantic embedding: while in (8) *think* takes scope over *say*, this dependency is inverted in the derivation tree (8.a). Joshi and Vijay-Shanker (1999) propose a special treatment for predicative auxiliary trees, in that the adjunction node is basically processed as if it were a substitution node, during semantic composition. This allows for correct embedding of *say* by *think*, as well as *like* by *wonder* in the semantic representation for (8).

Yet, as in (5), multiple adjunction of predicative auxiliary trees to *distinct nodes* of a single elementary tree (here *wonder* and *seem* into *like*) leads to additional complication. While the relative scope of *think* and *say* can be correctly determined – in that the trees stand in a direct adjunction relation – the relative scope of *wonder* and *seem* cannot be determined through variable bindings at adjunction nodes: They adjoin to distinct nodes in the verb's elementary tree. In order to derive the correct relative scopes in such configurations Joshi and Vijay-Shanker (1999) impose an ordering constraint for processing multiple (predicative) adjunctions into etrees, in a bottom-up manner: since *seem* adjoins to a lower node in the etree of *like* than *wonder* (node 2.2 vs. ϵ in (8.a)), it is processed first in semantic composition, thereby taking narrow scope relative to *wonder*.

It seems conceptually problematic to resort to specially designed ordering constraints for the traversal of derivation trees in semantic composition, especially in view of language specific constituent structure properties which might not correspond to the structure of semantic composition. The analysis does also not explicitly deal with scope ambiguities induced by NP quantification.

(8) Mary wondered who Peter thought John said Bill seemed to like.

(8.a)



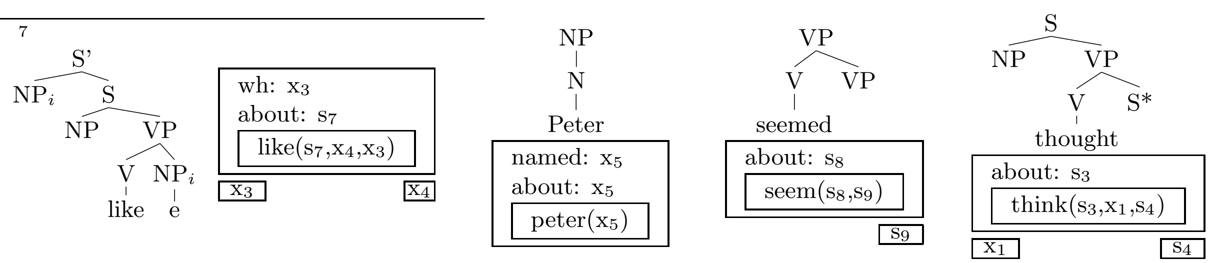
Kallmeyer and Joshi: Underspecified Semantics with MRS Kallmeyer and Joshi (1999) develop an account for underspecified semantics construction from LTAG derivation trees using MRS semantics. Very close to the architecture proposed in Joshi and Vijay-Shanker (1999) they associate flat semantic representations with elementary trees, now adopting the framework of Minimal Recursion Semantics to deal with scope underspecification. Semantic composition is again determined by the structure of derivation trees. The paper provides an underspecified analysis of quantification which accounts for quantifier scope ambiguities. It then focusses on examples of adjunct scope as in (9), where – given the assumption of *standard derivation* – *allegedly* must adjoin to *usually*, and is therefore restricted to take wide scope. Yet, given the assumption of *standard derivations*, scope ambiguities as in (10) can only be derived in terms of distinct derivations, i.e. *distinct* derivation trees. It is not clear in which way a *single underspecified* representation can be constructed from distinct derivation trees.

(9) Pat allegedly usually drives a cadillac.

(10) John intentionally knocked twice.

---

[7]



The paper is not really explicit about the distinction between semantic composition operations for adjunction versus substitution, in particular concerning the distinction between modifier and predicative auxiliary trees. We suppose that cascaded sentence embeddings can be handled along the lines of Joshi and Vijay-Shanker (1999)'s approach, by special conditions for variable binding on adjunction of predicative auxiliary trees. However, we see similar problems, in Kallmeyer and Joshi's account, to determine the correct embedding structure for multiple auxiliary trees adjoining to *distinct nodes* in a single elementary tree, as discussed for examples (5) and (8) above.

3.3 Discussion

We detailed the characteristics of LTAG syntax, in particular the structure of derivation trees, and the complexities that arise for semantics construction on the basis of LTAG derivation trees. We conclude that LTAG derivation trees do not provide an appropriate structure for principle-based semantics construction. The non-isomorphism between adjunction in syntax and modification in semantics introduces considerable complexity in semantics construction from derivation trees. The principle of extended domains of locality, in conjunction with the adjunction operation, yields semantically inappropriate dependencies in *derivation trees*, as these are imposed by purely syntactic operations in the composition of strictly lexicalised elementary trees. We therefore set out to investigate semantics construction in LTAG on the basis of the *derived tree*, which we consider more appropriate to guide principle-based meaning composition. We apply the framework of Linear Logic based glue semantics, which allows for considerable flexibility in the coupling of syntactic and semantic structure, while still remaining compositional in meaning construction.

4 Glue Semantics for LTAG

In applying glue semantics to LTAG we (i) define semantics on the basis of *derived* trees, which seems more appropriate for principle-based semantics construction. (ii), the loose coupling of syntactic and semantic structures with glue allows us to bridge the gap imposed by the aforementioned non-isomorphism in LTAG. (iii), we show that the glue approach captures non-tree local dependencies in modifier and control constructions. Finally, (iv) we propose a glue-based analysis of long-distance constructions, which in LTAG are driven by tree adjunction – as opposed to the f-structure based analysis in LFG with functional uncertainty.

4.1 Meaning constructors for LTAG initial elementary trees

To drive LL-based semantics construction, we need to associate meaning constructors with elementary trees, the lexical units of an LTAG grammar. As an interface to glue semantics we define principles for labelling nodes in elementary and derived trees with variables. These variables are referred to in the glue part of the associated lexical meaning constructors, and guide meaning composition.

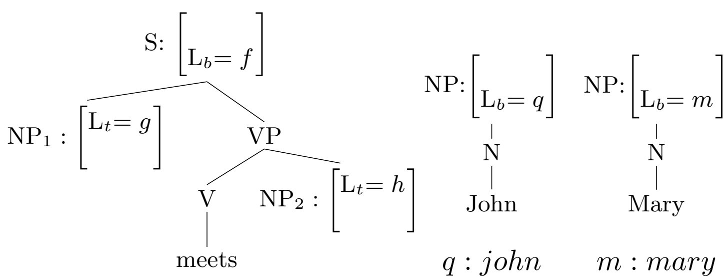
Tree Labelling Principle I, to be stepwise refined along the way, labels argument and root nodes in LTAG initial trees with atomic features L_t and L_b, which we will call *upper and lower labels*:

Tree Labelling Principle I
Assign variables $f, g, h \in VAR$ to top/bottom labels L_t/L_b of nodes n in LTAG initial etrees α

- Root nodes root(α) : $L_b(\text{root}(\alpha)) = x$, x a new variable from VAR
- Argument nodes $\text{arg}_i(\alpha)$: $L_t(\text{arg}_i(\alpha)) = x$, x a new variable from VAR

(11.a) displays sample etrees with associated meaning constructors. For *John meets Mary* we obtain the labelled derived tree (11.b). On substitution, feature bundles on substitution nodes and inserted root nodes are unioned, the resulting nodes display both upper and lower labels L_t and L_b.

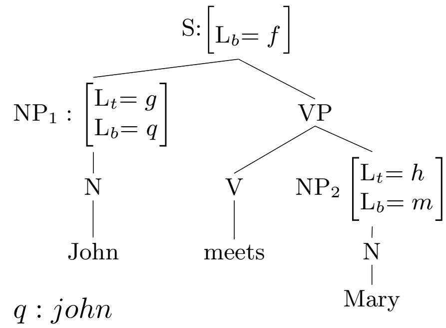
(11.a) Labelled elementary trees with associated meaning constructors



$q : john$ $\quad m : mary$

$h \multimap (g \multimap f) : \lambda y, x.meet(x, y)$

(11.b) Derived tree with assembled meaning constructors



$q : john$
$m : mary$
$h \multimap (g \multimap f) : \lambda y, x.meet(x, y)$

The meaning constructors of the etrees used in tree composition are assembled (11.b), but in their present form do not yield a successful proof in meaning derivation, since the variables in the glue parts are not connected. The missing equalities between variables are determined by tree composition, along the following lines.

Variable Equations in Tree Composition

- **Substitution:** when substituting β into α at node n_α, add equation: $\mathrm{L}_t(n_\alpha) = \mathrm{L}_b(\mathrm{root}(\beta))$
- **Adjunction:** see below

Using this information about variable equations, we could either resolve the equation system globally, or else trigger systematic variable substitutions in the set of assembled meaning constructors. We choose the latter option here, by adopting the following convention: for all substitution nodes n, and all assembled meaning constructors mcs, we replace (all occurrences of) the lower label variable $\mathrm{L}_b(n) = t$ by n's upper label variable $\mathrm{L}_t(n) = t'$. In (11.b), this triggers the substitutions $q \rightarrow g$ and $m \rightarrow h$.

Variable Substitution in (glue part of) meaning constructors *mcs*

- $\forall n \forall mcs.((L_b(n) = t \wedge L_t(n) = t' \wedge mcs[t]) \rightarrow mcs[t \rightarrow t'])$

LL-derivations for meaning construction With this in place we obtain the set of meaning constructors in (12) which yields a successful proof of the meaning associated with the tree's root variable, based on the Curry-Howard isomorphism.

(12) $g : john$
$h : mary$
$h \multimap (g \multimap f) : \lambda y, x.meet(x, y)$

$$\dfrac{g : john \qquad \dfrac{h \multimap (g \multimap f) : \lambda y, x.meet(x, y) \quad h : mary}{g \multimap f : \lambda x.meet(x, mary)}}{f : meet(john, mary)}$$

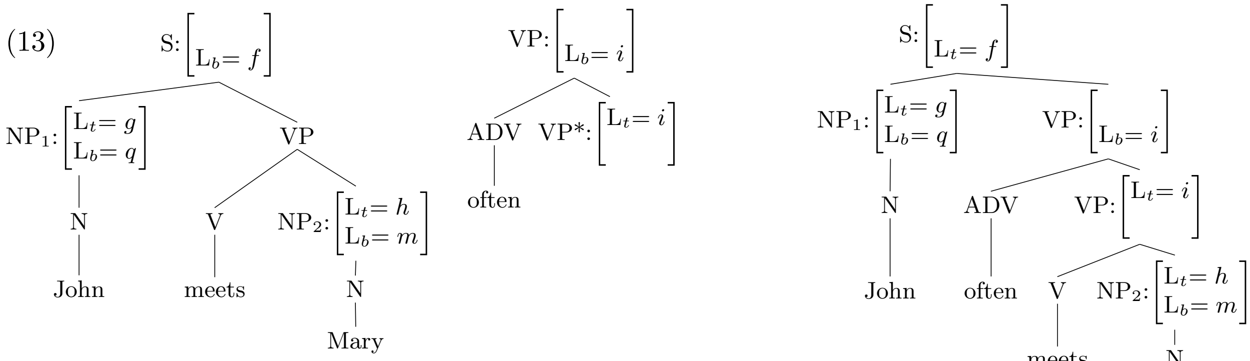
4.2 Non-tree local dependencies I: VP modification

Up to now we were only looking at constructions involving tree-local dependencies identified by argument substitution nodes. We now turn to constructions involving non-tree local dependencies, i.e. dependencies which are not identified by tree-local nodes in elementary trees. Constructions that fall into this class are modifiers and control constructions.

Let us first consider a VP modifying adverb like *often*. In LTAG, it is represented as an auxiliary tree that adjoins to VP, as displayed in (2). We extend our tree labelling principle to *modifier-type* auxiliary trees, assigning identical bottom and top labels to root and foot nodes: $\mathrm{L}_b(\mathrm{root}(\beta)) = \mathrm{L}_t(\mathrm{foot}(\beta)) = x$, x a fresh variable from the set of variables VAR.

For *often*, we obtain a labelled tree β as in (13). The meaning constructor for *often*, as a VP modifying adverbial, should consume and produce a VP meaning, which is characterised by consuming the subject's glue variable $\mathrm{L}_t(\mathrm{NP}_{1\alpha})$ of the tree α that β adjoins to, to produce the glue variable $\mathrm{L}_b(root(\alpha))$ of α's root node, as sketched below. But neither of these is local to the auxiliary tree β (*often*), and can therefore not be referred to in its associated meaning constructor.

(13)



$h \multimap (g \multimap f) : \lambda y, x.meet(x, y)$

$\lambda P, x.often(P(x))$:
$(\mathrm{L}_t(\mathrm{NP}_{1\alpha}) \multimap \mathrm{L}_b(root(\alpha))) \multimap (\mathrm{L}_t(\mathrm{NP}_{1\alpha}) \multimap \mathrm{L}_b(root(\alpha)))$

4.2.1 Labelling tree internal nodes: head projections

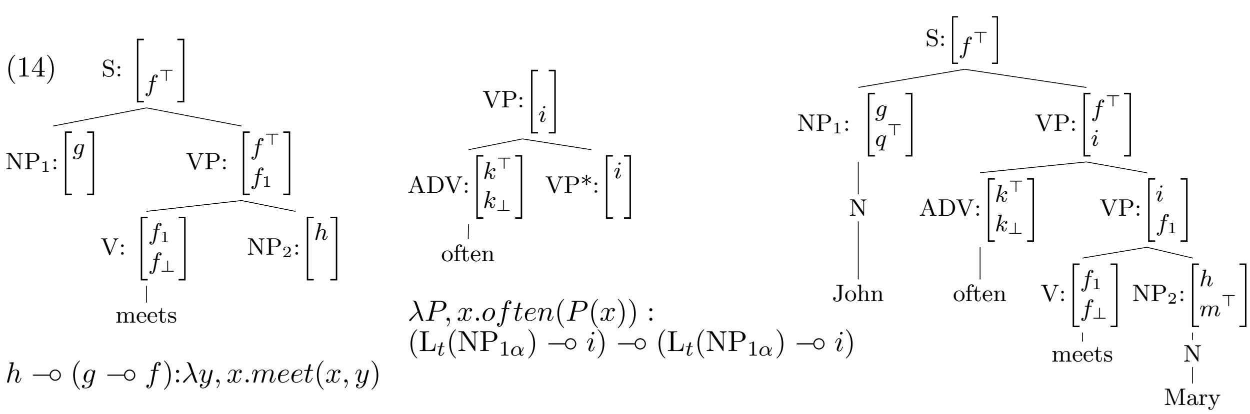
To capture such non-tree local dependencies, we revise our Tree Labelling Principle in two ways: (i) instead of root nodes, we label the lexical anchor (head) node with some variable $x_\perp$. (ii) this anchor node label is projected to all non-labelled tree internal nodes, introducing a chain of variables with intermediate projection labels $x_\perp, x_1, x_2, \ldots, x^\top$, as seen in (14).[8] Yet, given LTAG's concept of lexicalised elementary trees, with direct encoding of subcategorised argments, we will keep these projection variables distinct, triggering variable substitutions only locally, i.e. at local adjunction nodes, as opposed to unification of f-structure nodes in head projection chains in LFG.

Tree Labelling Principle II

- Anchor nodes: $\mathrm{L}_b(anchor(\alpha)) = x_\perp$, x new variable from VAR
- Argument nodes $\arg_N(\alpha)$: $\mathrm{L}_t(\arg_N(\alpha)) = x$, x new variable from VAR
- Modifier-type auxiliary trees β: $\mathrm{L}_b(\mathrm{root}(\beta)) = \mathrm{L}_t(\mathrm{foot}(\beta)) = x$, x new variable from VAR
- Projecting anchor node variable to all non-labelled tree internal nodes, introducing intermediate projection variables $\mathrm{x}_\perp$ x_1 x_2 ... $\mathrm{x}^\top$

On adjunction of some auxiliary tree β to a node n in α, n's label features are split: the top label feature L_t of n is assigned as the top label feature L_t of $root(\beta)$, and the bottom label feature L_b of n is assigned as the bottom label feature L_b of β's foot node. We obtain the derived tree in (14).

(14)



$\lambda P, x.often(P(x))$:
$(\mathrm{L}_t(\mathrm{NP}_{1\alpha}) \multimap i) \multimap (\mathrm{L}_t(\mathrm{NP}_{1\alpha}) \multimap i)$

$h \multimap (g \multimap f)$:$\lambda y, x.meet(x, y)$

[8]In the following we omit the feature names L_t and L_b, to avoid confusion with the upper and lower bounds of projected labels $\mathrm{x}_\perp$ x_1 x_2 ... $\mathrm{x}^\top$.

We now extend the conditions for equating variables in tree composition to the case of adjunction, and generalise the conditions for variable substitutions in meaning constructors to account for both substitution and adjunction. With these extensions and the labelling of tree internal nodes by projected anchor variables, the variable i in the meaning constructor for *often* will – after variable substitution – successfully refer to the sentence's root node variable $f^{\top}$.[9]

Variable Equations in Tree Composition

- **Substitution:** when substituting β into α at node n_α, add equation: $\mathrm{L}_t(n_\alpha)$=$\mathrm{L}_b(\mathrm{root}(\beta))$
- **Adjunction:** when adjoining β to α at node n_α, add equation: $\mathrm{L}_t(n_\alpha) = \mathrm{L}_b(\mathrm{root}(\beta))$

Variable Substitutions in meaning constructors mcs **using set of equations** EQ **(final)**

- $\forall n$ ($(\mathrm{L}_t(n) = x$ and $x = y \in EQ) \rightarrow \forall mcs : mcs[y \rightarrow x]$)

4.2.2 Labelling arguments as arguments of lexical heads

With these changes it is still not possible to refer to the non-tree local variable for the subject $\mathrm{L}_t(\mathrm{NP}_{1\alpha})$ in the meaning constructor of *often.* We therefore further revise Tree Labelling Principle (II) by encoding argument nodes in elementary trees as *arguments of their lexical head*, using the local tree's anchor label and a grammatical function identifier. Rather than using identifiers like NP_1, NP_2, etc., which in LTAG encode grammatical functions, we make use of grammatical function labels *subj*, *obj*, *comp*, etc., similar to those used in LFG.[10] So, if $f_{\perp}$ is the label of the lexical anchor, the subject node NP_1 will be labelled $f_{\perp}$:subj, the object node NP_2 $f_{\perp}$:obj.

As a result, the labelled trees look more LFG-like, but do not introduce additional linguistic assumptions into LTAG: Note that indices on NPs, such as NP_1, NP_2 do in fact encode grammatical functions in LTAG. This is brought out by looking at pairs of trees in the passive relation change.[11] Lexical heads are identified as primary anchors of elementary lexicalised trees, and head projection lines from these anchors emerge naturally as the complement of the set of nodes which are marked as argument or paired root/foot nodes of the local tree. Finally, LTAG's principle of extended domains of locality requires all arguments of a lexical item to be encoded within its elementary tree. Encoding arguments as *arguments of their head* is thus in line with LTAG's basic assumptions.

Tree Labelling Principle III (clause (ii)) (revised from (II))

- Argument nodes: $\mathrm{L}_t(\mathrm{arg}_N(\alpha)) = \mathrm{L}_b(\mathrm{anchor}(\alpha))$:$\mathrm{GF}_N$,
  where GF_N is the grammatical function corresponding to arg_N

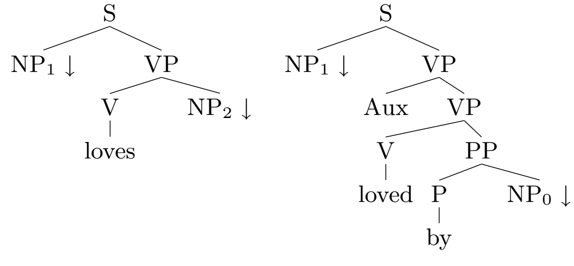
Elementary trees and meaning constructors for *John often sees Mary* are now revised according to the new conventions (see (15)). The meaning constructor for *often* refers to the non-tree local subject NP variable in terms of the local variable i: (i:subj $\multimap$ i) $\multimap$ (i:subj $\multimap$ i).

In tree composition we establish variable equations, which trigger variable substitutions in the assembled set of meaning constructors (16.a). The resulting set of (resolved) meaning constructors is (16.b). The glue formula contributed by *often* does now refer to the verb's subject node label as $f^{\top}$:subj and to the root node label as $f^{\top}$.

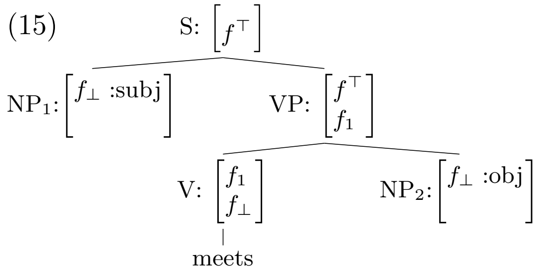
[9] We will display the analysis in (15) and (16), after a further revision to our Tree Labelling Principle.

[10] See also extensions in FTAG, which makes use of grammatical function features like SUBJ, OBJ, etc.

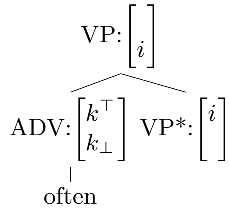
[11] The subject NP in the active tree is marked NP_1, the object NP_2. In the passive, the logical object is marked as the passive subject, again by index 1. The demoted logical subject, if present, is labelled NP_0.



(15)



$f_\perp$:obj $\multimap$ ($f_\perp$:subj $\multimap$ $f^\top$) : $\lambda y, x.meet(x, y)$

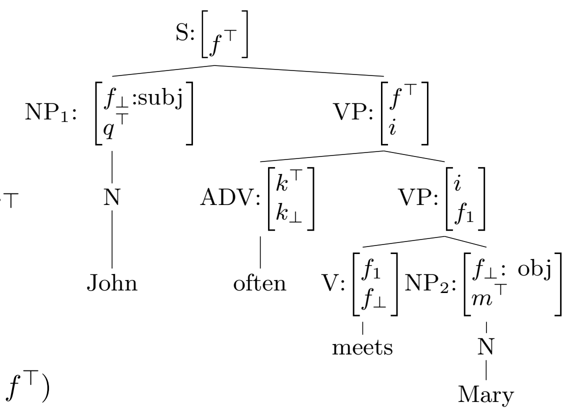


$\lambda P, x.often(P(x))$:(i:subj $\multimap$ i) $\multimap$ (i:subj $\multimap$ i)

(16.a)
$john : q^\top$
$mary : m^\top$
$\lambda y, x.meet(x, y)$: $f_\perp$:obj $\multimap$ ($f_\perp$:subj $\multimap$ $f^\top$)
$\lambda P, x.often(P(x))$: (i :subj $\multimap$ i) $\multimap$ (i :subj $\multimap$ i)

Substitutions: $q^\top \rightarrow f_\perp$:subj, $m^\top \rightarrow f_\perp$:obj, $i \rightarrow f^\top$

(16.b)
$john$: $f_\perp$:subj
$mary$: $f_\perp$:obj
$\lambda y, x.meet(x, y)$: $f_\perp$:obj $\multimap$ ($f_\perp$:subj $\multimap$ $f^\top$)
$\lambda P, x.often(P(x))$: ($f^\top$:subj $\multimap$ $f^\top$) $\multimap$ ($f^\top$:subj $\multimap$ $f^\top$)



However, the set of meaning premises (16.b) does not yield a successful meaning derivation. The resource $f_\perp$:obj (from *Mary*) can be consumed by $f_\perp$:obj $\multimap$ ($f_\perp$:subj $\multimap$ $f^\top$), to produce the resource $f_\perp$:subj $\multimap$ $f^\top$ corresponding to the VP meaning $\lambda x.meet(x, mary)$. But this latter resource cannot be consumed by the meaning constructor for *often*, which expects a VP constructor ($f^\top$:subj$\multimap f^\top$). This latter glue formula resulted from variable substitutions in the meaning constructor of *often*, which can only refer to its local variable i in root and foot node. In tree composition we established equality of this variable with the connecting adjunction node's top label $f^\top$.

4.2.3 Labelling arguments as arguments of local head projection

The LFG instructed reader will now suggest that we give up the distinction between head projection variables $x^\top \dots x_\perp$, by triggering global variable substitutions on head projection labels. However, we want to introduce as little additional assumptions in LTAG-based semantics construction as needed. We will show that we can continue to restrict ourselves to *local* variable substitutions at adjunction nodes, by a weaker extension of our labelling principle, which will encode the attachment of arguments as *attachments to the respective level of the local head projection*. And as we shall see later, in the discussion of long-distance dependencies, it is by this move – as opposed to global variable substitutions in head projections – that we can correctly define scope constraints in long-distance constructions, given LTAG's principle of *extended domains of locality*.

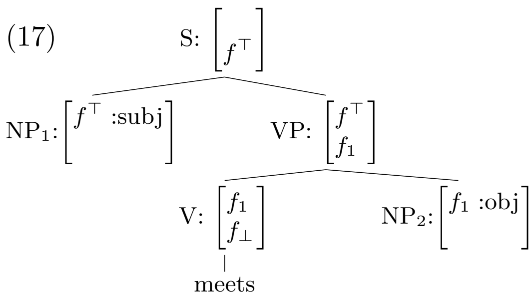
Our final version of the Tree Labelling Principle does now encode argument nodes $\text{arg}_N(\alpha)$ as *arguments of their local head projection*, by referring to the projection label of the argument's mother node (a tree internal node). That is, the projection index y in labels x_y:GF refers to the projection variable of the argument's mother (head projection) node.

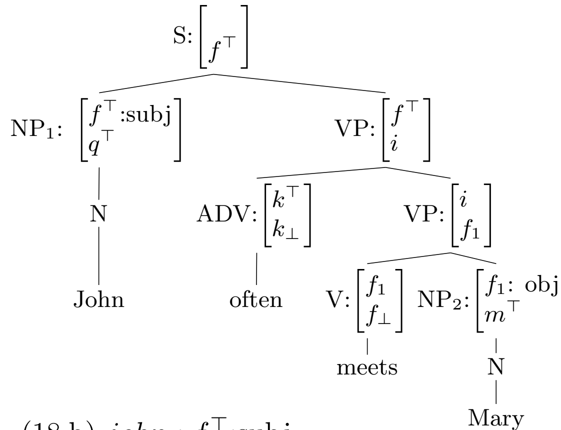
Tree Labelling Principle (Final Version)

- Anchor nodes: $\text{L}_b(anchor(\alpha)) = x_\perp$, x new variable from VAR
- Modifier-type auxiliary trees: $\text{L}_b(\text{root}(\beta)) = \text{L}_t(\text{foot}(\beta)) = x$, x new variable from VAR
- Projecting anchor node variable to all non-labelled tree internal nodes, introducing intermediate projection variables $\text{x}_\perp$ x_1 x_2 ... $\text{x}^\top$
- Argument nodes $\text{arg}_N(\alpha)$: $\text{L}_t(\text{arg}_N(\alpha)) = \text{L}_b(\text{mother}(\text{arg}_N(\alpha)))$: GF_N

The labelling of elementary trees differs only slightly from the previous version, the subject NP of *meets* being labelled $f^{\top}$:subj as before, since it attaches to the highest projection of the elementary tree, whereas the object NP attaches to projection level f_1, and is thus labelled f_1:obj. With variable substitutions $q^{\top} \rightarrow f^{\top}$:subj, $m^{\top} \rightarrow f_1$:obj, and $i \rightarrow f^{\top}$ the premises to meaning construction in (18.b) yield a successful proof in meaning derivation.

(17)





$f_1 \text{ :obj} \multimap (f_{\top} \text{ :subj} \multimap f^{\top}) : \lambda y, x.meet(x, y)$

(18.a) $john : q^{\top}$
$mary : m^{\top}$
$\lambda y, x.meet(x, y){:}f_1\text{:obj} \multimap (f_{\top}\text{:subj} \multimap f^{\top})$
$\lambda P, x.often(P(x))$:
$(i \text{ :subj} \multimap i) \multimap (i \text{ :subj} \multimap i)$

(18.b) $john : f^{\top}$:subj
$mary : f_1$:obj
$\lambda y, x.meet(x, y){:}f_1\text{:obj} \multimap (f^{\top}\text{:subj} \multimap f^{\top})$
$\lambda P, x.often(P(x))$:
$(f^{\top}\text{:subj} \multimap f^{\top}) \multimap (f^{\top}\text{:subj} \multimap f^{\top})$

$$\dfrac{\dfrac{\dfrac{\lambda y, x.meet(x,y){:}f_1\text{:obj} \multimap (f^{\top}\text{:subj} \multimap f^{\top}) \quad mary{:}f_1\text{:obj}}{\lambda x.meet(x, mary){:}f^{\top}\text{:subj} \multimap f^{\top}} \quad \lambda P, x.often(P(x)){:}(f^{\top}\text{:subj} \multimap f^{\top}) \multimap (f^{\top}\text{:subj} \multimap f^{\top})}{\lambda x.often(meet(x, mary)){:}f^{\top}\text{:subj} \multimap f^{\top}} \quad john{:}f^{\top}\text{:subj}}{often(meet(john, mary)){:}f^{\top}}$$

4.2.4 Deriving scope ambiguities

With our Tree Labelling Principle in place, we will now illustrate that LTAG semantics construction based on derived trees accounts for scope ambiguities induced by modifiers and NP quantifiers.

By identifying root and foot labels of modifiers, and due to local variable substitutions in tree composition, we account for the scoping behaviour of modifiers to take scope over other modifiers within their clausal projection. In particular, proper labelling conditions of argument nodes ensures that the meaning constructor's non-local variable stays local to the clause nucleus.[12]

In (19) we consider a case of modifier scope ambiguity, with one of the adverbs adjoining to S, the other to VP.[13] After substitutions, the set of premises allows for derivation of alternative meanings, by either first consuming the meaning constructor for *twice*, and then *intentionally*, or vice versa. That is, from the single derived tree we obtain ambiguous semantic analyses, with alternative modifier scopes: *intentionally*(*john*, *twice*(*call*(*john*, *mary*))), and *twice*(*intentionally*(*john*, *call*(*john*, *mary*))).

[12]See Section 4.4 for the analysis of predicative auxiliary trees.

[13]Example (19) brings us to the special case where modifier adjunction applies to the root node, which doesn't specify an upper label. For this case we need to refine the conditions for adjunction in tree composition by adopting clause (i), which provides an upper label for modifier adjunction to root nodes. Clause (ii) can then apply as before. Clause (ii) will also cover predicative auxiliary trees (see below).

(i) On adjunction of a modifier auxiliary tree β to a node n in α, if n does not specify an upper label, we instantiate an upper label, assigning it the value of n's lower label.

(ii) On adjunction of some auxiliary tree β to a node n in α, n's label features are split: the top label feature L_t of n, if instantiated, is assigned as the top label feature L_t of $root(\beta)$, and the bottom label feature L_b of n is assigned as the bottom label feature L_b of β's foot node.

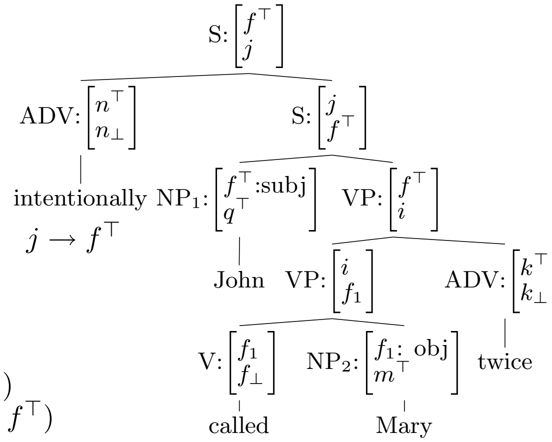
We further assume that for meaning derivation it is the lower label of the sentence's root node that constitutes the target of the proof in meaning derivation (resp. the variable it is substituted with).

(19)

$john : q^\top$
$mary : m^\top$
$\lambda y, x.call(x, y)$: f_1:obj $\multimap (f^\top$:subj $\multimap f^\top)$
$\lambda P, x.intent(x, P(x))$: $(j$:subj $\multimap j) \multimap (j$:subj $\multimap j)$
$\lambda P, x.twice(P(x))$: $(i$:subj $\multimap i) \multimap (i$:subj $\multimap i)$

Substitutions: $q^\top \rightarrow f^\top$:subj, $m^\top \rightarrow f_1$:obj, $i \rightarrow f^\top$, $j \rightarrow f^\top$

$john : f^\top$:subj
$mary : f_1$:obj
$\lambda y, x.call(x, y)$: f_1:obj $\multimap (f^\top$:subj $\multimap f^\top)$
$\lambda P, x.twice(P(x))$: $(f^\top$:subj $\multimap f^\top) \multimap (f^\top$:subj $\multimap f^\top)$
$\lambda P, x.intent(x, P(x))$: $(f^\top$:subj $\multimap f^\top) \multimap (f^\top$:subj $\multimap f^\top)$



Yet, for left and right adjoining VP modifiers as in *John intentionally called Mary twice*, assuming *standard derivation*, we obtain alternative derived trees with identical, ambiguous meanings. We therefore adopt Schabes and Shieber (1994)'s *extended derivations* for modifier-type auxiliary trees, allowing multiple adjunction to single nodes. Assuming further that simultaneous left and right adjunction to a single node produces a shared adjunction root node, as described in (Schabes and Waters 1995), we obtain a *single* derived tree, which yields the same ambiguity as in (19).

Finally, we need to account for scope restrictions. In Crouch and van Genabith (1999) scope restrictions are defined by *scope constaints*, which restrict the order of derivations in meaning construction, to yield corresponding scope meanings. Scope constraints refer to variables in glue expressions. But modifiers attaching to the same clause exhibit *identical* glue variables (after equality resolution) – they cannot be distinguished in the glue part. This problem could be solved by allowing scope constraints to refer to *glue : meaning pairs*, i.e. by exploiting the Curry-Howard isomorphism. In *John called Mary intentionally twice* the governing modifier could introduce a constraint *intentionally* $\prec$ *twice* – an instruction to *consume/apply* the *glue : meaning* pair associated with *intentionally* before the one associated with *twice*, corresponding to wide scope of *twice*.

Scope ambiguities with NP quantifiers are accounted for in the definition of the associated meaning constructors, in line with standard glue semantics (cf. Dalrymple (1999)). In (20) we display the entries for quantified pronominal NPs. For *Everyone meets someone*, we obtain – after substitutions – a set of instantiated meaning constructors which allows for alternative meaning derivations, corresponding to alternative quantifier scopes.

(20)

NP: $\begin{bmatrix} g^\top \end{bmatrix}$
everyone
$\lambda P.\forall x(person(x) \rightarrow P(x))$:
$(g^\top \multimap X) \multimap X$

NP $\begin{bmatrix} h^\top \end{bmatrix}$
someone
$\lambda P.\exists x(person(x) \wedge P(x))$:
$(h^\top \multimap X) \multimap X$

Instantiated mcs for *Everyone meets someone*
$\lambda P.\forall x(person(x) \rightarrow P(x))$:$(f^\top$:subj $\multimap X) \multimap X$
$\lambda P.\exists x(person(x) \wedge P(x))$:$(f_1$:obj $\multimap X) \multimap X$
$\lambda y, x.meet(x, y)$:f_1:obj $\multimap f^\top$:subj $\multimap f^\top$

4.3 Taking stock

Let us review the basic assumptions in our design of LL-based semantics construction for LTAG. As in glue semantics applied to LFG, semantics construction is lexicon driven, here by associating meaning constructors with lexicalised elementary trees. While glue semantics in LFG is based on f-structure – variables in the glue part of meaning constructors refer to instantiated f-structure nodes (or their semantic σ–projections) – glue semantics for LTAG is based on *labelled* derived trees. The principle for labelling elementary trees was introduced stepwise, in order to introduce as little additions to the LTAG framework as needed, and to motivate specific conceptual moves and machinery. Labelling of trees with features assigning upper and lower variables is a necessary extension to LTAG if semantics construction is based on *derived* as opposed to *derivation trees*, and we have presented evidence for serious shortcomings of the latter approach in Section 3. It is especially the labelling of *tree internal nodes by (projected) anchor labels* that extends basic assumptions of LTAG, which in its basic form only “talks about” root, substitution and adjunction nodes in elementary trees. As we have seen, in semantics construction from derived trees, reference to tree internal nodes is crucial for connecting variables in adjunction structures for modification. Further we have seen that non-tree local dependencies as in the case of VP modifying adverbs (and similarly for control verb constructions, see below) can only be captured by (i) labelling arguments *as arguments of their lexical head*, and (ii) association with grammatical function labels (or similar naming conventions). Finally, (iii), the labelling of arguments needs to specify the head projection level where attachment takes place in order to allow for a principled solution to the scoping of adverbs and – as we shall see – wh-elements in long-distance constructions.

In tree composition we establish equations between variables in upper and lower labels, based on two basic principles for substitution and adjunction. Variable substitutions in meaning constructors are uniquely based on these equations. That is, only those variables that are *local to substitution or modifier root nodes* play a role in establishing connections between the isolated variables assigned in elementary trees. Variables of intermediate projection nodes which are not touched by adjunction or substitution operations can be safely ignored in meaning construction.

We consider *head projection labelling* as the most crucial addition to LTAG’s standard assumptions. It is a necessary extension for semantics construction from derived trees, and a crucial assumption to account for the most characteristic and difficult aspects of LTAG syntax and its syntax-semantics interface, namely adjunction and its interplay with semantic composition and the definition of scope. This is evident in the analysis of modifiers and scope ambiguity, and required for long distance dependencies in Section 4.6.

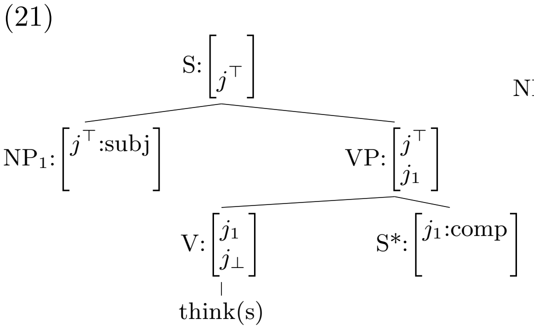
4.4 Bridging the syntax-semantics non-isomorphism in LTAG

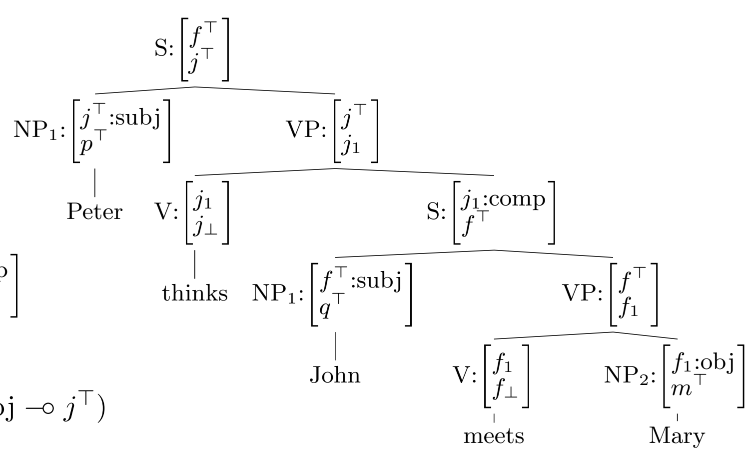
As outlined in Section 2, a main characteristic of LTAG syntax is that – due to the principle of extended domains of locality and the existence of long distance phenomena – sentential embedding is necessarily driven by adjunction in syntax, that is, by the same syntactic operation that applies to optional (and recursive) modifiers. This results in a non-isomorphism between adjunction in syntax on the one hand, and the semantic operations of complementation vs. modification on the other. As we saw in discussion of Joshi and Vijay-Shanker (1999), this generates inverted embedding structures in derivation trees, as opposed to correct embeddings in derived tree structures, and leads to complications in semantics construction from derivation trees.

One would therefore expect that in our approach, where the syntax-semantics interface is built on the derived tree, the problem of this non-isomorphism is automatically circumvented. But this is not the case. The problem arises in our approach as well, yet in a slightly different way.

Let us take a look at a sentence embedding verb like *think* in a simple embedding configuration (21). Sentence embedding verbs (and similarly control verbs) are represented as auxiliary trees, yet differ from *modifier-type* auxiliary trees in that their foot node corresponds to an argument of the

(21)





$\lambda P, x.think(x, P)$: j_1:comp $\multimap$ ($j^\top$:subj $\multimap$ $j^\top$)

lexical head. In LTAG this difference is captured by the distinction between *modifier* and *predicative* auxiliary trees. In our Tree Labelling Principle labelling of root and foot nodes with identical variables was restricted to *modifier auxiliary trees*. The foot node of a *predicative auxiliary tree*, by contrast – as argument node of the lexical anchor – falls under the conditions for argument nodes, and is assigned a composed label, consisting of a projection variable and a grammatical function label.[14] For the *predicative* auxiliary tree *think* we thus obtain a labelled elementary tree as displayed in (21). In contrast to modifier trees, root and foot node labels are distinct.

In tree composition, the bottom label of the foot node is instantiated with the bottom label of the S node to which adjunction applies ($f^\top$ in (21)), in accordance with our principles for assigning top and bottom labels in tree composition.[15]

Although the derived tree does not reflect any crucial difference between adjunction and substitution, we need to adjust the conditions for *variable equations* in tree composition to account for the special aspects of semantic composition with *predicative auxiliary trees*, as opposed to *modifier auxiliary trees*. On adjunction of modifiers root and foot node variables need to be equated with the projection variable of the adjunction node, to allow modifiers to take scope within the maximal projection. This we obtained by equating upper and lower labels at the modifier's root node in the derived tree. For *predicative auxiliary trees*, by contrast, we need to link, i.e. equate, the labels of argument foot node and adjunction node, parallel to the standard case of argument *substitution*.

All other things being equal, then, we account for the non-isomorphism of adjunction in semantics construction through refinement of variable equations in tree composition: We distinguish between adjunction of *modifier* and *predicative* auxiliary trees, the latter being defined along the lines of standard cases of argument substitution. Thus, in variable equation, on adjunction of *predicative auxiliary trees* the foot node plays the role the substitution node n_α plays in substitution, and the adjunction node n_α plays the role of the root node of the inserted tree in substitution.

Variable Equations in Tree Composition (Final Version)

- **Substitution:** When substituting β into α at node n_α, add equation: $\mathrm{L}_t(n_\alpha)=\mathrm{L}_b(\mathrm{root}(\beta))$
- **Adjunction:** When adjoining β to α at node n_α,
  if β is a *modifier auxiliary tree*, add equation: $\mathrm{L}_t(n_\alpha) = \mathrm{L}_b(\mathrm{root}(\beta))$
  if β is a *predicative auxiliary tree*, add equation: $\mathrm{L}_t(\mathrm{foot}(\beta)) = \mathrm{L}_b(n_\alpha)$

With these adjustments we trigger substitution of $f^\top$ by J_1:comp in the set of meaning constructors (22) assembled for (21).[16] After substitution of variables, complex terms such as j_1:comp:subj are treated as atoms in matching producers and consumers in Linear Logic meaning derivation.

(22) Substitutions: $p^\top \to j^\top$:subj, $q^\top \to f^\top$:subj, $m^\top \to f_1$:obj, $f^\top \to \mathrm{J}_1$:comp

$peter : p^\top$	$peter : j^\top$:subj
$john : q^\top$	$john : j_1$:comp:subj
$mary : m^\top$	$mary : f_1$:obj
$\lambda y, x.meet(x,y)$:f_1:obj $\multimap$ ($f^\top$:subj $\multimap$ $f^\top$)	$\lambda y, x.meet(x,y)$:f_1:obj $\multimap$ (j_1:comp:subj $\multimap$ j_1:comp)
$\lambda P, x.think(x,P)$:j_1:comp $\multimap$ ($j^\top$:subj $\multimap$ $j^\top$)	$\lambda P, x.think(x,P)$:j_1:comp $\multimap$ ($j^\top$:subj $\multimap$ $j^\top$)
	$\vdash think(peter, meet(john, mary))$

[14] We assume a mapping from argument foot nodes S with discriminating features (such as [$\pm fin$] etc.) to corresponding grammatical functions COMP and XCOMP.

[15] Since the upper label of the adjunction node is not instantiated, the upper label of the root node is instantiated to $f^\top$ (see the refined definition for adjunction in fn. 13). The target variable in the meaning proof is $\mathrm{L}_b(\mathrm{root}) = j^\top$.

[16] Note that in this example the order in which substitutions are triggered is critical. In general, though, we can avoid specification of some canonical order for substitutions by *repeatedly* applying substitutions to the modified premises in the same, arbitrarily fixed order, until no more modifications are triggered in a complete run. Alternatively, one could adopt a general resolution algorithm, along the lines of Kaplan and Bresnan (1995).

4.5 Non-tree local dependencies II: Control constructions

We now show how LL-based semantics construction on the basis of labelled LTAG trees accounts for non-tree local dependencies in control constructions.

4.5.1 Control constructions in LFG

We first take a quick look at the treatment of control in LFG syntax and LL-based glue semantics. Lexical entries of control verbs specify the relation between controller and embedded subject in terms of functional equations. In the resulting f-structure the matrix SUBJ of *tries* is unified with the SUBJ of the embedded verb, that is, they are instantiated to some unique f-structure node (g in (23)). The meaning constructor associated with *tries*[17] consumes the VP meaning of the embedded verb $(\uparrow\text{XCOMP SUBJ}_\sigma) \multimap (\uparrow\text{XCOMP}_\sigma)$ to produce the VP meaning of *tries*, which then consumes the overt matrix subject to produce the sentence meaning $((\uparrow\text{SUBJ}_\sigma) \multimap \uparrow_\sigma)$. Parallel to linear logic derivation based on the glue formulae the meaning is computed in the associated meaning terms.

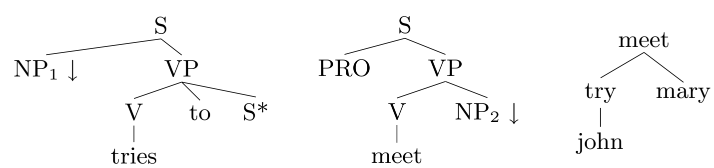
(23) tries V,	Mcs for: *John tries to meet Mary*
$(\uparrow \text{PRED})=$ ‘TRY$\langle (\uparrow \text{SUBJ})\ (\uparrow \text{XCOMP})\rangle$’	$john : g$
$(\uparrow \text{SUBJ})= (\uparrow \text{XCOMP SUBJ})$	$mary : m$
$\lambda P, x.try(x, P)$:	$\lambda y, x.meet(x, y)$:$m \multimap (g \multimap h)$
$((\uparrow\text{XCOMP SUBJ}_\sigma) \multimap (\uparrow\text{XCOMP}_\sigma)) \multimap ((\uparrow\text{SUBJ}_\sigma) \multimap \uparrow_\sigma)$	$\lambda P, x.try(x, P)$:$(g \multimap h) \multimap (g \multimap f)$

$$\dfrac{\dfrac{\dfrac{\lambda y, x.meet(x, y) : m \multimap (g \multimap h) \quad mary : m}{\lambda x.meet(x, mary) : g \multimap h} \quad \lambda P, x.try(x, P) : (g \multimap h) \multimap (g \multimap f)}{\lambda x.try(x, \lambda x.meet(x, mary)) : g \multimap f} \quad john : g}{try(john, \lambda x.meet(x, mary)) : f}$$

4.5.2 Control constructions in LTAG

In LTAG syntax control verbs are encoded as *predicative auxiliary trees*, since they can occur in long-distance dependencies. Infinite verbs embedded under control verbs do not display an overt subject. The corresponding elementary tree therefore encodes a PRO subject node. Due to adjunction, the *derivation tree* for *John tries to meet Mary* incorrectly represents the control verb as dependent on the embedded verb, *meet*. Moreover, since no substitution takes place into the subject argument position of *meet* the derivation tree does not represent *John* as an argument of *meet*.

(24)

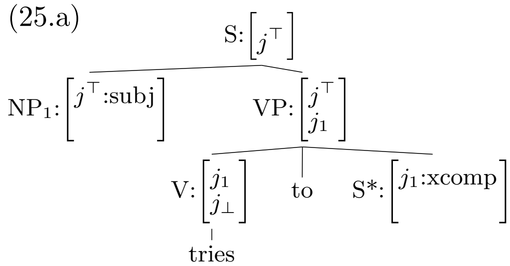


4.5.3 Glue Semantics for LTAG control contructions

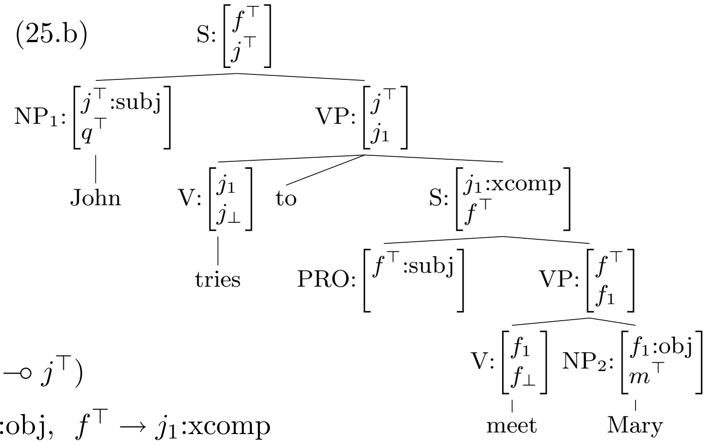
Applying our Tree Labelling Principle to the *predicative* auxiliary tree *tries* (25.a), we label the subject NP and the adjunction foot node S* as subj and xcomp arguments, anchored to the respective head projection variables $j^\top$ and j_1. Labelling and meaning constructor for the embedded infinitive etree *meet* (see (24)) is identical to (17), except for the presence of a subject PRO argument, which is not a substitution node. For *John tries to meet Mary* we obtain the labelled derived tree (25.b). With variable substitutions in the set of assembled meaning constructors, we obtain a successful proof of the sentence meaning, similar to the meaning derivation in the LFG analysis.

[17] We follow the analysis given in Asudeh (2000) and Asudeh (2001).

(25.a)



(25.b)



$\lambda P, x.try(x, P)$:
$(j_1\text{:xcomp:subj} \multimap j_1\text{:xcomp}) \multimap (j^\top\text{:subj} \multimap j^\top)$

Substitutions: $\quad q^\top \rightarrow j^\top\text{:subj}, \;\; m^\top \rightarrow f_1\text{:obj}, \;\; f^\top \rightarrow j_1\text{:xcomp}$

$john : q^\top$
$mary : m^\top$
$\lambda y, x.meet(x, y) : \; f_1\text{:obj} \multimap (f^\top\text{:subj} \multimap f^\top)$
$\lambda P, x.try(x, P)$:
$(j_1\text{:xcomp:subj} \multimap j_1\text{:xcomp}) \multimap (j^\top\text{:subj} \multimap j^\top)$

$john : j^\top\text{:subj}$
$mary : f_1\text{:obj}$
$\lambda y, x.meet(x, y)$:
$f_1\text{:obj} \multimap (j_1\text{:xcomp:subj} \multimap j_1\text{:xcomp})$
$\lambda P, x.try(x, P)$:
$(j_1\text{:xcomp:subj} \multimap j_1\text{:xcomp}) \multimap (j^\top\text{:subj} \multimap j^\top)$

$$\cfrac{\cfrac{\cfrac{\lambda y, x.meet(x,y) : f_1\text{:o} \multimap (j_1\text{:xcmp:s} \multimap j_1\text{:xcmp}) \quad mary : f_1\text{:o}}{\lambda x.meet(x, mary) : j_1\text{:xcmp:s} \multimap j_1\text{:xcmp} \quad \lambda P, x.try(x,P) : (j_1\text{:xcmp:s} \multimap j_1\text{:xcmp}) \multimap (j^\top\text{:s} \multimap j^\top)}}{\lambda x.try(x, \lambda x.meet(x, mary)) : j^\top\text{:s} \multimap j^\top \qquad john : j^\top\text{:s}}}{try(john, \lambda x.meet(x, mary)) : j^\top}$$

On closer inspection, though, an interesting difference emerges between the LFG and LTAG analyses. In LFG coreference of controller and controlled subject is represented in f-structure by mapping them to a single f-structure node, due to the control relation in the lexical entry of the control verb. This is explicit in the instantiated meaning constructor of *tries*. In the LTAG analysis this control relation is not expressed in terms of tree labelling. This is seen in the meaning constructor of *tries*, where embedded and matrix subject are referred to by distinct labels, j_1:xcomp:subj and $j^\top$:subj, both before and after variable substitutions. Despite the missing identification of controller and controlled argument, meaning derivation in the LTAG analysis yields the same sentence meaning.

This fact is interesting in that it points to an important difference between the LFG and LTAG frameworks. LFG's syntactic f-structure representation effectively encodes predicate-argument dependencies. In LTAG syntax – even if augmented with grammatical relation labelling – neither labelled derived trees, nor derivation trees reflect full-fledged predicate-argument structures. While the embedded subject in (25.b) displays an upper label $f^\top$:subj, there is no equivalent to a control relation to establish coreference of controller ($j^\top$:subj) and controlled argument. It is even more compelling, then, that in coupling LTAG syntax with glue semantics construction from labelled derived trees, we obtain a well-formed meaning, i.e., the same meaning we obtain from fully specified dependencies in LFG f-structures.[18] We will come back to this issue in Section 5.

[18]The reader might object here, since the meaning of controlled infinitives is a property, which does not express the control relation in question. This is most evident with transitive control verbs, where the derived meaning does not encode a subject vs. object control reading. For *John promised Mary to leave*, e.g. we obtain the meaning *promise* $(john, mary, \lambda x.leave(x))$. Meaning postulates need to encode which of the matrix arguments satisfies the property.

A slight revision of the meaning of control verbs proposed in Asudeh(2000, 2001) solves this problem. Instead of a property meaning, *tries* could select an open proposition $\lambda P, x.try(x, P(x))$, a verb like *promise* could be assigned the meaning $\lambda P, y, x.promise(x, y, P(x))$. In meaning derivation, then, consumption of the matrix subject will appropriately instantiate the embedded open proposition ($promise(john, mary, leave(john))$ in the above example). Note that an open proposition analysis still accounts for the well-known inference problems with control verbs, if we

4.6 Long-distance dependencies and wh-scope

We finally turn to long-distance constructions, which turned out to be particularly difficult for LTAG semantics construction from derivation trees. Wh-constructions involve the additional complexity of defining an operator semantics for wh-phrases such as *who(m)*, and defining the scope of the wh-operator. As we shall see now, it is the treatment of scope and wh-semantics in long-distance constructions which motivated much of the architecture we developed in the previous sections.

Labelling etrees for long extraction (26) displays an etree for wh-object extraction, with a fronted wh-object NP and a corresponding (unlabelled) trace. Due to object fronting and the additional projection level, argument labelling differs from (17): the object attaches to the highest projection, while the subject NP attaches to the f_2 projection level. Note further that the meaning constructor is specified to produce $f2$ as (the glue equivalent of) the meaning of the clause headed by *meet*, which is the label of the tree internal node that is the target for adjunction for clause embedding auxiliary trees, such as *think*. The assignment of f_2 as the target of sentence meaning reflects – on the semantic side – the split that is triggered – in syntax – by adjunction of sentence embedding trees, which split the basic elementary tree into a long-distance extracted projection level ($f^\top$), and the clause's basic tree trunk, cut at projection level f_2. This is seen in (27): on adjunction of *think* to S[−wh], the foot node's argument label j_1:comp is linked to f_2.[19]

Deriving meanings for long extractions Ignoring the role of the wh-phrase for a moment, by looking at the lower projections in (27), a meaning for f_2 can only be derived by *hypothesising*, or *assuming* a resource $z : f^\top$:obj, z an unbound variable, to fill the object argument position of *meet*. This kind of hypothetical reasoning we observe in the upper part of the meaning derivation (28), which leads to the partial meaning $meet(john, z) : j_1$:comp, z the unbound variable from the assumed premise $[z : f^\top$:obj]. Due to variable equations, j_1:comp ($= f_2$) is consumed by *think*, which after consumption of its subject (j_1:subj) delivers the partial meaning $think(peter, meet(john, z)) : f^\top$. Since this meaning was obtained on the *assumption* of $[z : f^\top$:obj], the assumption has to be discharged in order to yield a valid proof. This is done in the implication introduction step (I) in (28), which yields the partial meaning $\lambda z.think(peter, meet(john, z)) : f^\top$:o $\multimap f^\top$. This basically says that *had we found* a premise $z : f^\top$:o, whatever the meaning of z, it is of this z that Peter thinks that John meets z. The meaning constructor associated with *whom* in (28) is now easily understood: it combines with a (property) meaning which was built on the assumption of an object-NP meaning for *meets* ($f^\top$:o $\multimap f^\top$), and should deliver the meaning of the root node ($f^\top$) to which *whom* attaches, by applying the function $\lambda v.who(v)$ to this partial meaning $\lambda z.think(peter, meet(john, z))$: ($f^\top$:o $\multimap f^\top$) $\multimap f^\top : \lambda v.who(v)$.

Defining wh-scope The difficult aspects in meaning construction for long wh-extractions are the following: (i) the scope of the wh-operator is constrained by the position of the wh-phrase in the *derived* tree. That is, we are not allowed to derive narrow scope of the wh-operator in (27) and (28). (ii) from the local elementary tree of *whom* we cannot access the nodes S[−wh] or S[+wh] of the verb's tree, which are crucial to define these scope constraints. The labelling of the etree *whom* captures the non-tree local variable $f^\top$ of the tree it is substituted in by assigning its root node the label $X^\top$:obj, X a metavariable. The meaning constructor is defined by reference to $X^\top$. In tree composition, substitution triggers the equality $X^\top$:obj $= f^\top$:obj, and at the same time instantiates $X^\top$ to $f^\top$. With global substitutions as given in (27) we obtain a set of premises that yield the

assume quantification to range over a variable P of type property, just as in the classical property analysis.

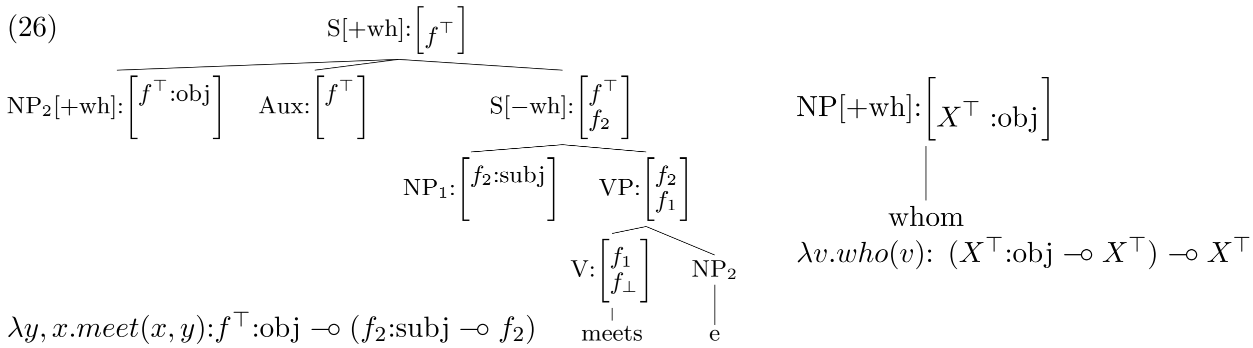
Peter tries to leave.	try(peter,λx.leave(x)(peter))
John tries everything Peter tries.	∀P [try(peter,P(peter))→try(john,P(john))] where P = λx.leave(x)
⊢ John tries to leave.	⊢ try(john,λx.leave(x)(john))
⊬ John tries for Peter to leave.	

With these additional assumptions we can safely argue that Linear Logic based semantics construction is in fact able to bridge the missing control relation in LTAG syntax.

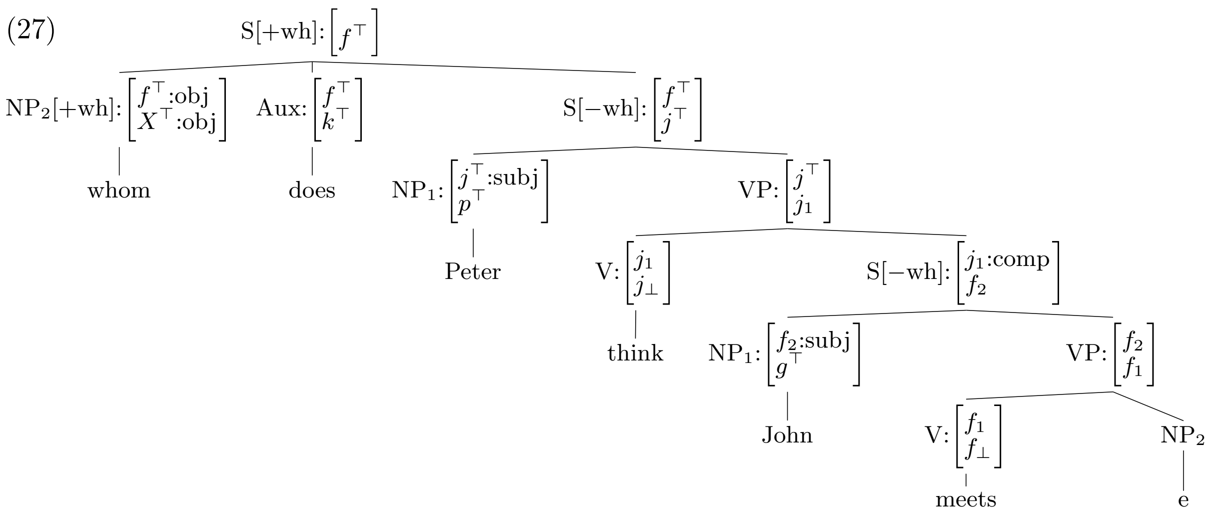
[19] See the variable equation conditions for predicative auxiliary trees (p. 15).

desired meaning in (28), with wide scope of the wh-operator over the clause headed by *think*. Wide scope is enforced by the definition of *whom*, which is constrained to consume $f^{\top}$:obj $\multimap f^{\top}$, where $f^{\top}$ is the clause label of the projection *whom* is attached to. Narrow scope is prohibited due to the labelling of the etree *meets*, which produces the resource f_2, and thus a partial meaning $meet(jn, z) : f_2$, where $f_2 = j_1$:comp by variable substitution. If implication introduction (I) were to apply to this partial meaning, the resulting meaning term $\lambda z.meet(john, z) : f^{\top}$:obj $\multimap j_1$:comp would not provide an adequate resource for *whom* to define narrow scope under *think*.

(26)



(27)

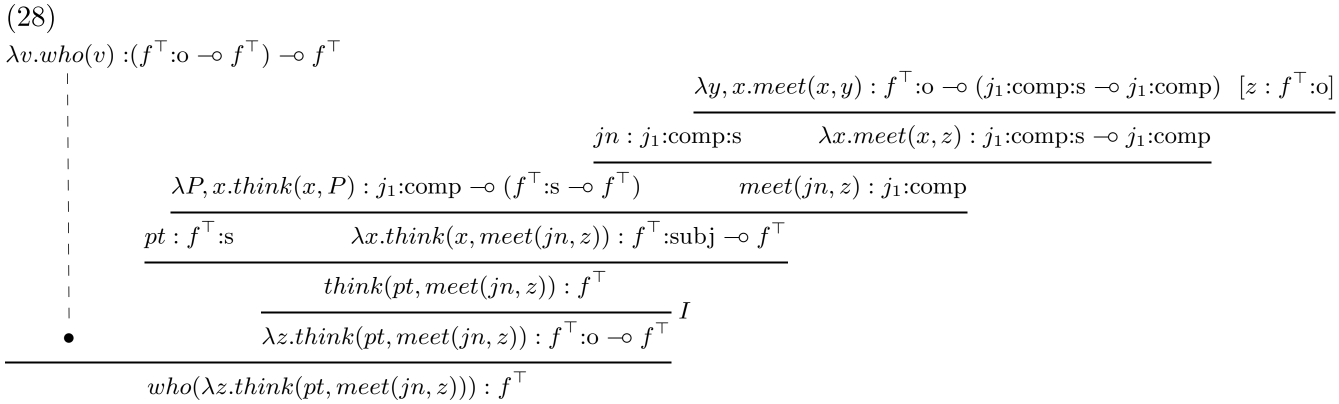


Substitutions:

$g^{\top} \rightarrow f_2$:subj, $\quad f_2 \rightarrow j_1$:comp, $\quad p^{\top} \rightarrow j^{\top}$:subj, $\quad j^{\top} \rightarrow f^{\top}$, $\quad X^{\top}$:obj$\rightarrow f^{\top}$:obj, $\quad X^{\top} \rightarrow f^{\top}$

$\lambda v.who(v) : (X^{\top}\text{:obj} \multimap X^{\top}) \multimap X^{\top}$	$\lambda v.who(v) : (f^{\top}\text{:obj} \multimap f^{\top}) \multimap f^{\top}$
$peter : p^{\top}$	$peter : f^{\top}\text{:subj}$
$john : g^{\top}$	$john : j_1\text{:comp:subj}$
$\lambda y, x.meet(x, y)\text{:}f^{\top}\text{:obj} \multimap (f_2\text{:subj} \multimap f_2)$	$\lambda y, x.meet(x, y)\text{:}f^{\top}\text{:obj} \multimap (j_1\text{:comp:subj} \multimap j_1\text{:comp})$
$\lambda P, x.think(x, P)\text{:}j_1\text{:comp} \multimap (j^{\top}\text{:subj} \multimap j^{\top})$	$\lambda P, x.think(x, P)\text{:}j_1\text{:comp} \multimap (f^{\top}\text{:subj} \multimap f^{\top})$
	$\vdash who(\lambda z.think(peter, meet(john, z)))$

(28)



In sum, by tree internal projection labelling, meaning constructors in extraction contexts can refer to the adjunction node label (f_2) which defines the upper bound of the embedded clause trunk, and thereby correctly constrains the scope of the extracted wh-phrase, which by reference to $X^\top$ ($=f^\top$) is forced to take scope over the projection level to which it attaches in the derived tree.

Note that in the labelled derived tree (27), similar to the case of control constructions, substitution of equated variables does not yield a well-formed dependency structure. The object *whom* labelled ($f^\top$:obj $\multimap f^\top$) $\multimap f^\top$ is not linked to *meets*, which is looking for a premise labelled $f^\top$:obj. So, similar to what we observed in control constructions, it by Linear Logic meaning derivation that the missing dependency link is established, here in terms of hypothetical reasoning. This contrasts again with the architecture of LFG, where extracted phrases are represented in their basic argument position in the f-structure, local to their governing predicate, by the interplay of functional uncertainty equations and completeness and coherence constraints.

Having dealt with long-distance dependencies, we can resume our discussion of semantics construction from derivation trees in Section 3. These approaches were shown to have difficulties in representing the embedding structure of long distance constructions involving adjunction of sentential embedding and raising verbs to *distinct* nodes in the basic verb's elementary tree, as in examples (5) and (8). For reasons of space we cannot go through the examples in detail, but we sketch the relevant aspects here.[20]

For long object topicalisation in (5) and long object extraction in (8) we assume transitive verb trees along the lines of *meet* in (26), without auxiliary insertion at the root node. Raising verbs like *seem* are represented as modifier-type VP auxiliary trees, with a meaning constructor $\lambda P.seem(P)$: $i \multimap i$, where i the variable assigned to root and foot node labels. Variable substitutions in tree composition ensure that in the derived meaning for (5) (*Spicy hotdogs he claims Mary seems to adore*) *claim* takes scope over *seem*, which in turn takes scope over *adore*. Given appropriate meaning constructors for question embedding verbs, our account provides the correct semantics and embedding structure in cases like (8).

4.7 Related approaches: Semantics construction for D-Tree Grammars

Finally we briefly mention two approaches which are related to our work, yet at the same time need to be differentiated. Hepple (1999) proposes categorial-style semantics construction from derived trees of D-Tree Grammar (DTG, Rambow et al. (1995)), to dispense with the more problematic, process-based interpretation model provided by derivation structures. To overcome problems faced in the analysis of NP quantification he then provides a small fragment for glue semantics from DTG derived trees,[21] exploiting the greater flexibility of glue semantics in meaning assembly. Yet, while closely related to LTAG, DTG elementary trees (d-trees) can provide a larger syntactic context than LTAG etrees[22] and therefore lead to a different set-up for tree labelling, as compared to the LTAG framework we were dealing with. Still, we believe that comparison of our LTAG-based approach to (an extended fragment of) Hepple's account in DTG could lead to interesting insights into the special aspects of the respective frameworks in the syntax-semantics interface.

Muskens (2001) develops a very elegant description-based model of syntax and the syntax-semantics interface, which allows syntactic and semantic representations to be highly underspecified. The model is applied to LTAG syntax, yet again coupled with extension to tree descriptions as used in D-Tree Grammar. As in Hepple's and our approach, semantics construction is not guided by derivation structures. Elementary tree descriptions are enriched with semantic expressions in predicate logic, and instructions for meaning composition. The design of the syntax-semantics differs from our approach in using D-Tree Grammar and a different model for semantic composition. Detailed discussion and comparison cannot be provided within the scope of this paper.

[20]The reader is invited to go through the analysis in detail.

[21]The fragment in Hepple (1999) does not cover VP modification, control or long-distance dependencies.

[22]Especially for operators in long-distance dependencies and quantifiers.

5 What we learn about the relation between LTAG and LFG

We defined a Linear Logic-based semantics interface for LTAG syntax which is lexicon driven, and builds on the structure of derived trees. We argued that a major extension to the LTAG framework is *labelling of tree internal nodes by projected anchor variables*. Since we operate on derived trees, tree composition records variable equations which are restricted to variables of *substitution and adjunction nodes*, and allow for coherent meaning composition from assembled meaning constructors.

We noted a striking difference of LTAG tree labelling as opposed to the LFG framework: while in LFG all levels in head projection chains are unfied in terms of $\uparrow=\downarrow$ equations, LTAG tree labelling can restrict itself to equation of variables at substitution and adjunction nodes. This is of course due to the principle of strict lexicalisation and localisation of arguments in elementary trees.

We observed that labelled LTAG derived trees are not merely a mirror image of LFG c-structures annotated with f-descriptions. LTAG labelled trees do not encode control relations, nor are extracted arguments represented as local to their governing predicates. That is, we could conceive of labelled LTAG trees as an impoverished version of annotated c-structure trees in LFG: impoverished in that they lack functional path and uncertainty equations, and only allow for restricted variable equations, at local substitution and adjunction nodes. As a consequence, labelled derived trees (including variable equations) do not in general encode dependency structures, yet they provide semantically valid embedding structures, as opposed to LTAG derivation trees. An interesting observation emerging from this comparison is that this lack is compensated by LL-based semantics construction, which establishes the correct dependencies at the level of meaning representation.

It is especially with respect to long distance constructions that semantics construction from constitutent trees proved to be superior to meaning composition from derivation trees. This we achieved by a major addition to the LTAG framework, in that we label arguments to record the local projection level to which they attach. This addition is critical for the definition of scope constraints, as well as the correct definition of embedding structures, in particular in long extraction contexts with multiple adjunctions into single elementary trees. Moreover, we could show that glue semantics from labeled derived trees accounts for modifier and quantifier scope ambiguities.

Should we conclude from this exercise, then, that the additional power that LFG provides – the projection of f-structure by functional descriptions (in particular functional uncertainty, and other path equations), completeness and coherence constraints in f-structure, and global variable equations of head projection chains – is a luxury, or formal overhead which we could dispense with on the basis of a sparser syntactic formalism? The answer is certainly – No.

It is, first of all, a matter of conceptual clarity and modularity of linguistic representation which calls for an adequate *syntactic* representation of predicate-argument relations, as it is provided by dependency, or f-structure representations. Second, it is a well-known fact that tree composition operations in LTAG syntax are restricted in formal power, and do not extend to languages with special word order properties. Example (29) from Kashmiri, discussed in Rambow et al. (1995), illustrates such a case where wh-words end up in sentence-second position, preceded by a topic from the matrix clause. This type of interleaved combination of elementary trees, called *subsertion* in D-Tree Grammars, cannot be accounted for by adjunction of atomic, elementary lexicalised trees. – A case for (functional uncertainty in) f-structure.

(29)	Rameshan	kyaa$_i$	chu	baasaan	[ki	me	kor	t$_i$]
	Ramesh$_{\text{ERG}}$	what	is	believe$_{\text{Nperf}}$	that	I$_{\text{ERG}}$	do	
	What does Ramesh believe that I did?							

6 Conclusion

In conclusion we argue that LTAG derivation trees are inappropriate for principle-based semantic composition. We developed an account for semantics construction from LTAG *derived trees* by establishing a tree labelling regime as an interface to Linear Logic-based meaning deduction. We showed that glue semantics from LTAG derived trees captures non-tree local dependencies in modification and control constructions, and extends to the adjunction-based analysis of long-distance dependencies, which is, however, formally and empirically more restricted than LFG's analysis in terms of functional uncertainty. Linear Logic-based semantics from *derived trees* successfully bridges LTAG's non-isomorphism between syntactic and semantic operations. It further allows us to derive scope ambiguities induced by modifiers and quantified NPs, and accounts for scope constraints in long-distance extraction contexts.

We established that labelled LTAG derived trees do not correspond to full-fledged dependency structures, as opposed to LFG f-structures, since LTAG tree labelling does not extend to the formal power of f-descriptions in LFG. However, dependencies which cannot be established by local variable substitutions can be appropriately bound at the meaning level, that is, through Linear Logic-based meaning assembly in the glue semantics interface.

References

Asudeh, A. (2000). Functional Identity and Resource-Sensitivity in Control. In Butt, M. and King, T., editors, *Proceedings of the LFG00 Conference*, pages 2–24, University of California, Berkley. CSLI Online Publications, `http://csli-publications.stanford.edu/`.

Asudeh, A. (2001). A Resource-Sensitive Semantics for Equi and Raising. In Beaver, D. and Kaufmann, S., editors, *Proceedings of Semantics Fest 2000*, pages 5–25. CSLI Publications, Stanford, CA.

Crouch, R. and van Genabith, J. (1999). Context Change, Underspecification, and the Structure of Glue Language Derivations. In Dalrymple, M., editor, *Semantics and Syntax in Lexical Functional Grammar*, pages 117–189. MIT Press.

Dalrymple, M., editor (1999). *Semantics and Syntax in Lexical Functional Grammar*. MIT Press.

Dalrymple, M. (2001). *Lexical-Functional Grammar*, volume 34 of *Syntax and Semantics*. Academic Press.

Hepple, M. (1999). A Functional Interpretation Scheme for D-Tree Grammars. In *Proceedings of the Third International Workshop on Computational Semantics*, pages 117–130, KUB, Tilburg.

Joshi, A. (1987). An Introduction to Tree Adjoining Grammar. In Manaster-Ramer, A., editor, *Mathematics of Language*. John Benjamins, Amsterdam.

Joshi, A. and Vijay-Shanker, K. (1999). Compositional Semantics with Lexicalized Tree-Adjoining Grammar (LTAG): How Much Underspecification is Necessary? In Bunt, H. and THijsse, E., editors, *Proceedings of the Third International Workshop on Computational Semantics (IWCS-3)*, pages 131–145, Tilburg.

Kallmeyer, L. and Joshi, A. (1999). Factoring Predicate Argument and Scope Semantics: Underspecified Semantics with LTAG. In Dekker, P., editor, *Proceedings of the 12th Amsterdam Colloquium*, pages 169–174.

Kaplan, R. and Bresnan, J. (1995). A Formal System for Grammatical Representation. In *Formal Issues in Lexical-Functional Grammar*. CSLI Lecture Notes, No.47.

Muskens, R. (2001). Talking about Trees and Truth-conditions. *Journal of Logic, Language, and Information*, 10(4).

Rambow, O., Vijay-Shanker, K., and Weir, D. (1995). D-Tree Grammars. In *Proceedings of ACL-95*, pages 151–158.

Schabes, Y. and Shieber, S. (1994). An Alternative Conception of Tree-Adjoining Derivation. In *Computational Linguistics 20(1)*, pages 91–124.

Schabes, Y. and Waters, R. C. (1995). Tree Insertion Grammar: A cubic-time, parsable formalism that lexicalizes context-free grammar withough changing the trees produced. In *Computational Linguistics 21(4)*, pages 479–513.

Shieber, S. and Schabes, Y. (1990). Synchronous Tree-Adjoining Grammars. In *Proceedings of COLING*, pages 1–6.

Treebank vs. Xbar-based Automatic F-Structure Annotation

Josef van Genabith
Computer Applications
Dublin City University
Dublin, Ireland
josef@compapp.dcu.ie

Anette Frank
Language Technology Group
DFKI GmbH
Saarbrücken, Germany
frank@dfki.de

Andy Way
Computer Applications
Dublin City University
Dublin, Ireland
away@compapp.dcu.ie

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu

Treebank vs. Xbar-based Automatic F-Structure Annotation

Abstract

Manual, large scale (computational) grammar development is time consuming, expensive and requires lots of linguistic expertise. More recently, a number of alternatives based on treebank resources (such as Penn-II, Susanne, AP treebank) have been explored. The idea is to automatically "induce" or rather read off (P)CFG grammars from the parse annotated treebank resources and to use the treebank grammars thus obtained in (probabilistic) parsing or as a starting point for further grammar development. The approach is cheap, fast, automatic, large scale, "data driven" and based on real language resources. Treebank grammars typically involve large sets of lexical tags and non-lexical categories as syntactic information tends to be encoded in monadic category symbols. They feature flat rules (trees) that can "underspecify" attachment possibilities. Treebank grammars do not in general follow Xbar architectural design principles (this is not to say that treebank grammars do not have design principles). As a consequence, treebank grammars tend to have very large CFG rule bases (e.g. Penn-II 17,000 CFG rules for about 1 million words of text) with often only minimally differing rules. Even though treebank grammars are large, they are still incomplete, exhibiting unabated rule accession rates. From a grammar engineering point of view, the size of the rule base poses problems for maintainability, extendability and, if a treebank grammar is to be used as a CF-base in a LFG grammar, for functional (feature-structure) annotations. From the point of view of theoretical linguistics, flat treebank trees and treebank grammars extracted from such trees do not express linguistic generalisations. From the perspective of empirical and corpus linguistics, flat trees are well-motivated as they allow underspecification of subtle and often time consuming attachment decisions. Indeed, it is sometimes doubted whether highly general Xbar schemata usefully scale to "real" language. In previous work we developed methodologies for automatic feature-structure annotation of grammars extracted from treebanks. Automatic annotation of "raw" treebank grammars is difficult as annotation rules often need to identify subsequences in the RHSs of flat treebank rules as they explicitly encode head, complement and modifier relations. Xbar-based CFG rules should substantially facilitate automatic feature-structure annotation of grammar rules. In the present paper we conduct a number of experiments to explore a space of possible grammars based on a small fragment of the AP treebank resource. Starting with the original treebank fragment we automatically extract a CFG . We then apply an automatic structure preserving grammar compaction step which generalises categories in the original treebank fragment and reduces the number of rules extracted, resulting in a generalised treebank fragment and in a compacted grammar ^{c}. The generalised fragment is then manually corrected to catch missed constituents (and the like) resulting in an automatically extracted, compacted and (effectively manually) corrected grammar ^{c,m}. Manual correction proceeds in the "spirit" of treebank grammars (we do not introduce Xbar analyses). We then explore how many of the manual correction steps on treebank trees can be achieved automatically. We develop, implement and test an automatic treebank "grooming" methodology which is applied to the generalised treebank fragment to yield a compacted and automatically corrected grammar ^{c,a}. Grammars ^{c,m} and ^{c,a} are very similar to compiled out "flat" LFG-82 style grammars. We explore regular expression based compaction (both manual and automatic) to relate ^{c,m} to a LFG-82 style grammar design. Finally, we manually recode a subsection of the generalised and manually corrected treebank fragment into "vanilla-flavour" Xbar based trees. From these we extract a compacted, manually corrected, Xbar-based grammar ^{c,m,x}. We evaluate our grammars and methods using standard labelled bracketing measures and according to how well they perform under automatic feature-structure annotation tasks.

1 Introduction

Manual, large scale (computational) grammar development is time consuming, expensive and requires lots of linguistic expertise. More recently, a number of alternatives based on treebank resources (such as Penn-II, Susanne, AP treebank) have been explored. The idea is to automatically "induce" or rather read off (P)CFG grammars from the parse annotated treebank resources and to use the treebank grammars thus obtained in (probabilistic) parsing [Charniak,93] or as a starting point for further grammar development [Krotov et al,98]. The approach is cheap, fast, automatic, large scale, "data driven" and based on real language resources.

Treebank grammars typically involve large sets of lexical tags and non-lexical categories as syntactic information tends to be encoded in monadic category symbols. They feature flat rules (trees) that can "underspecify" attachment possibilities. Treebank grammars do not in general follow Xbar architectural design principles (this is not to say that treebank grammars do not have design principles). As a consequence, treebank grammars tend to have very large CFG rule bases (e.g. Penn-II $>$ 17,000 CFG rules for about 1 million words of text) with often only minimally differing rules. Even though treebank grammars are large, they are still incomplete, exhibiting unabated rule accession rates. From a grammar engineering point of view, the size of the rule base poses problems for maintainability, extendability and, if a treebank grammar is to be used as a CF-base in a LFG grammar, for functional (feature-structure) annotations. From the point of view of theoretical linguistics, flat treebank trees and treebank grammars extracted from such trees do not express linguistic generalisations. From the perspective of empirical and corpus linguistics, flat trees are well-motivated as they allow underspecification of subtle and often time consuming attachment decisions. Indeed, it is sometimes doubted whether highly general Xbar schemata usefully scale to "real" language.

In previous work [Sadler, van Genabith and Way,00] and [Frank,00] developed methodologies for automatic feature-structure annotation of grammars extracted from treebanks (see also [Frank et al,01]). Automatic annotation of "raw" treebank grammars is difficult as annotation rules often need to identify subsequences in the RHSs of flat treebank rules. Xbar-based CFG rules should substantially facilitate automatic feature-structure annotation of grammar rules as they explicitly encode head, complement and modifier relations [Bresnan,01].

In the present paper we report on a number of experiments in which we explore a space of possible grammars based on a small fragment of the AP treebank resource.

Starting with the original treebank fragment we automatically extract a CFG G. We then apply an automatic structure preserving grammar compaction step which generalises categories in the original treebank fragment and reduces the number of rules extracted, resulting in a generalised treebank fragment and in a compacted grammar G . The generalised fragment is then manually corrected to catch missed constituents (and the like) resulting in an automatically extracted, compacted and (effectively manually) corrected grammar G . Manual correction proceeds in the "spirit" of treebank grammars (we do not introduce Xbar analyses). We then explore how many of the manual correction steps on treebank trees can be achieved automatically. We develop, implement and test an automatic treebank "grooming" methodology which is applied to the generalised treebank fragment to yield a compacted and automatically corrected grammar G . Grammars G and G are very similar to compiled out "flat" LFG-82 style grammars. We explore regular expression based compaction (both manual and automatic) to relate G to a LFG-82 style grammar design. Finally, we manually recode a subsection of the generalised and manually corrected treebank fragment into "vanilla-flavour" Xbar-based trees. From these we extract a compacted, manually corrected, Xbar-based grammar G .

We evaluate our grammars and methods using standard labelled bracketing measures and according to how well they perform under automatic feature structure annotation tasks.

2 Automatic F-Structure Annotation

F(eature)-structures [Kaplan and Bresnan,82] are "higher level", abstract syntactic representations. It would be desirable to have treebank resources with both CF-PSG and feature-structure representations

- as training resources for probabilistic LFG/unification/constraint-based grammars
- to develop stand-alone LFG/unification/constraint-based grammar resources

In previous work [Sadler, van Genabith and Way,00] and [Frank,00] developed two methodologies for automatically annotating treebank trees and grammars extracted from such trees with feature-structure information. One method involves the statement of regular expression based feature-structure annotation principles, the other (cascaded) rewriting of tree descriptions in a tree description language.

2.1 Regular Expression Based Annotation

Regular expression based annotation is best explained by way of example. The following is a flat CFG rule extracted from a treebank resource (note that each constituent is of the form `cat:F` where `cat` is a monadic CFG category and `F` is a variable designed to carry feature-structure information):

```
vp:VP > adv:A v0:V0 v0:V1 v0:V2 s:S pp:P
```

In flat treebank rules such as this, RHSs often correspond to sequences of multiple (embedded) constituents. As a consequence, annotation principles need to be partial and underspecified in order to be able to associate subsequences of RHSs with the required feature-structure annotations. We employ regular expression to express underspecified rule LHSs and RHSs and to pick out subsequences in rule RHSs. Annotation principles are of the form `LHS > RHS @ ANN`, where `ANN` is a set (in our implementations a list) of conjunctions or disjunctions of feature-structure equations, such as:

```
vp:VP > * v0:V1 v0:V2 * @ [V1:xcomp=V2,V1:subj=V2:subj].

vp:VP > *(~v0) v0:V0 *  @ [VP=V0].

vp:VP > * v0:V0 s:S *   @ [V0:comp=S].

      :                  :
```

The first principle identifies `v0:V1 v0:V2` subsequences in rule RHSs and states that for two adjacent `v0`s the feature structure associated with the second `v0` provides the value of an `xcomp` feature of the first. Additionally, both share the same value for their `subj` attributes. The second principle states that the leftmost `v0` is the head of a `vp` rule. A particular principle may apply several times to different subsequences in the RHS of a single rule and a number of different principles may apply to any one rule. All principle applications are collected and the rule is annotated accordingly:

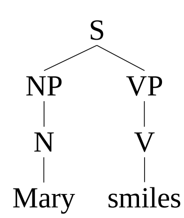
```
vp:VP > adv:A v0:V0 v0:V1 v0:V2 s:S pp:P
      @ [VP=V0, V0:xcomp=V1, V0:subj=V1:subj, V1:xcomp=V2,
         V1:subj=V2:subj, V2:comp=S].
```

2.2 Tree Annotation Using Tree Descriptions

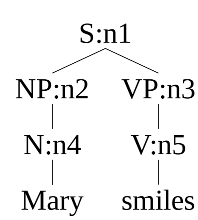
In our second method annotation principles are applied to partial configurations of treebank *trees* using a tree rewriting technique. Again, we illustrate the approach by way of a simple example.

Treebank trees are first encoded as a flat collection (conjunction) of basic tree description predicates (`arc, prec, lex`) in a tree description language. Tree nodes are associated with a unique identifier (`n1,n2,..`).

```
[s,[np,[n,Mary]],
   [vp,[v,smiles]]].
```




```
arc(n1,s,n2,np), arc(n2,np,n4,n)
arc(n1,s,n3,vp), arc(n3,vp,n5,v),
prec(n2,n3),
lex(n4,Mary), lex(n5,smiles)
```

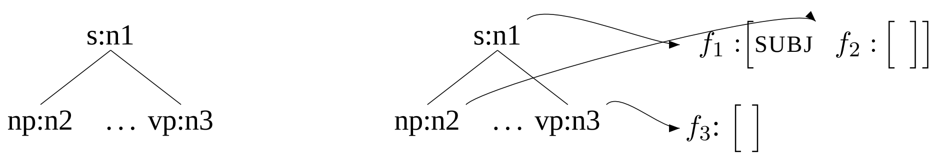


We then state annotation principles in terms of rewriting rules which apply to the flat set of tree descriptions. Rewriting rules are of the form `CD = AD` where `CD` is a constraining tree description and `AD` is an annotation description. The formalism supports a set of switches that optionally allow consumption of constraining tree descriptions during rule application. Annotation principles can thus support both order independent (enriching) and cascaded order dependent (rewriting) annotation.

In the case at hand the constraining tree configuration is provided by a local `s > np vp` subtree. The annotation description introduces new tree description predicates which associate each tree node with a feature-structure via the `phi` projection. The feature-structure associated with the subject `np` is constrained to be the value of the `subj` attribute of the feature-structure associated with the sentence node.

Unlike the regular expression based annotation method which is restricted to (partial) local trees of depth one, the description-based approach can access (partial) non-local tree configurations.

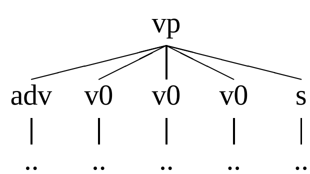
```
+arc(A,s,B,np),
+arc(A,s,C,vp), +prec(B,C)  =    phi(A,FA), phi(B,FB), subj(FA,FB).
```



In the present paper we will use the regular expression based method operating on PS rules for specifying feature-structure annotation principles and the tree description based rewriting approach for our automatic treebank "grooming" methodology presented in Section 7 below.

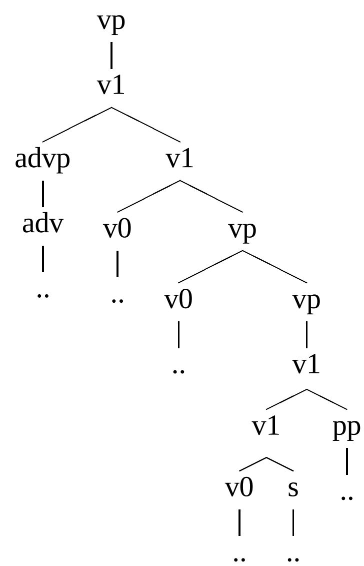
3 Flat vs. Xbar-based Treebank Rules and Annotation

Consider again the flat treebank example rule here displayed as a local tree configuration:



Automatic annotation can be difficult as annotation principles have to identify suitable subsequences on the RHS of the corresponding CFG rule for annotation. This may not always be possible resulting in partial and/or incorrect annotations.

Contrast this with a "vanilla flavour" Xbar recoding of the tree (we will say more in Section 9 below about what exactly we mean by "vanilla flavour" Xbar encoding):



The recoding is exclusively binary and unary branching with each local subtree clearly identifying head, adjunct and complement relations. The Xbar configuration and grammar design principles should guide and facilitate automatic feature-structure annotation to a much greater extent than the flt treebank rule representations which often conflate otherwise embedded subconstituents. The rule base for the recoded tree consists of:

```
vp > v1       v1 > advp v1    v1 > v1 pp
vp > v0 vp    v1 > v0 vp      v1 > v0 s
```

The corresponding annotation principles are:

```
__:X > __:Y      @ [X=Y]
v1:X > v1:Y  _:Z @ [X=Y,X:adjn $ Z]
v1:X >  _:Z v1:Y @ [X=Y,X:adjn $ Z]
v1:X > v0:Y  s:Z @ [X=Y,X:comp Z]
    ...
```

The first principle states that any unary branching rules project their feature-structures. The second and third identify the head in pre- and post-modifier configurations and make the modifiers elements '`$`' of an adjunct set. Finally, a sentential complement provides the value of the complement attribute `comp`. Notice that in the case at hand, f-structure annotation is completely driven by Xbar configuration. Notice further that in contrast to our original annotation principles, for the flat treebank configuration, Xbar-based annotation principles are no longer underspecified. There is no need to identify subsequences of constituents as a separate constituent in RHSs of binary branching Xbar rules.

Intuitively, Xbar-based automatic annotation looks attractive. At this stage we may summarise our preliminary findings in the following table:

Xbar	TreeBank
principle based architecture	underspecification and pragmatism
generalisations	many flat rules
small rule base	large rule base
conducive to automatic annotation	difficult for automatic annotation

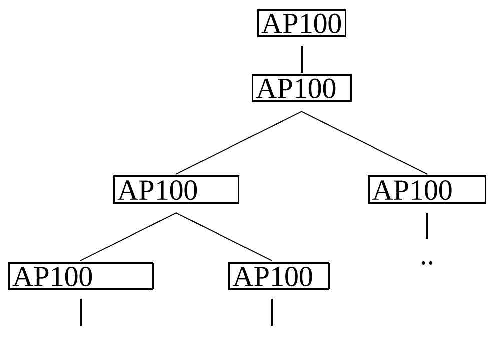
However, there are a number of considerations (often expressed by practitioners working in the fields of empirical and corpus linguistics) that sound a more cautionary note:

- Xbar design principles often force subtle and difficult attachment decisions and sometimes it may not even be clear how these can be resolved. Cases in point are modifier attachment. Binary branching rules force attachment decisions which are not always clear cut in the case of multiple modifiers, pre- and post-modifiers of the same bar one level node or multiple attachment possibilities for a single modifier (e.g. where does the pp in our recoded example rule attach?). In any case, subtle and difficult decisions such as these are too time consuming and thus not practical for treebanking purposes. By contrast, flat treebank representations support (albeit in a somewhat crude fashion) a notion of underspecification not provided by classical Xbar design principles.

- Somewhat more speculatively it is sometimes assumed that Xbar design principles do not really scale up to real language in any useful way.

4 Treebank Grooming

In the present paper we report on a number of experiments to evaluate a "vanilla flavour" Xbar design approach to a fragment of the AP treebank to measure performance under automatic feature structure annotation and to ascertain the feasibility of Xbar inspired approaches to tree banking tasks and resources derived from treebanks. Starting with the original treebank fragment we move the fragment in stages to a resource inspired by Xbar design principles. Some of the moves are automatic and involve novel treebank "grooming" methodologies; others are manual. The different stages (and their interdependencies) are best displayed by way of a family tree:



Our experiments are based on the publicly available subsection AP100 (the first 100 parse-annotated sentences) of the AP treebank [Leech and Garside,91]. From this we automatically extract a grammar G following the method of [Charniak,93].[1] As our starting point we apply an automatic, structure preserving grammar compaction technique [van Genabith, Sadler and Way,99], [Hepple and van Genabith,00] to AP100 and generate AP100 . The technique generalises overspecific categorisation and this can have the effect of collapsing grammar rules. From AP100 we extract a compacted grammar G . We manually annotate G with feature-structure annotations to create a "gold-standard" against which we evaluate automatic feature-structure annotation. We then develop a set of regular expression based feature-structure annotation principles and apply them automatically to an unannotated version of G . We evaluate the results of automatic annotation against the gold standard using precision and recall measures [Sadler, van Genabith and Way,00]. These results provide the base line for our further experiments.

In the next step, we manually correct any "obvious" errors in AP100 . These include missed groupings of constituents, attachment errors, relabelling of certain nodes and changes to the original representation of punctuation. All changes are in the spirit of the original treebank grammar (i.e. we do not introduce Xbar assumptions at this stage). Manual correction transforms AP100 into AP100 . From this we extract

[1] All the grammars in our experiments are automatically extracted from the original or "groomed" parse-annotated treebank resources in this fashion. The basic idea is simple: a recursive procedure traverses treebank trees and for each local tree of depth one records a corresponding CFG rule. CFG rule token counts allow the computation of simple relative frequency based maximum likelihood estimators to compute probabilistic versions of the CFGs extracted.

grammar G . We manually[2] annotate the corrected G to produce a gold standard against which to evaluate automatic feature-structure annotation of G . We then develop a set of regular expression based feature-structure annotation principles and apply them automatically to an unannotated version of G . We evaluate the results of automatic annotation against the gold standard using precision and recall measures. The expected outcome is, of course, that a manually corrected grammar is "cleaner", more systematic and contains fewer "quirks" than the orginal treebank grammar and is thus easier to annotate automatically. We also report labelled bracketing measures to compare the structure of the trees in AP100 to the manually "groomed" AP100 .

In our next experiment, we investigate to which extent the manual transformation operations required to "groom" or "convert" the trees in AP100 into AP100 can in fact be performed automatically. Automatic treebank grooming is important as manual correction is extremely time consuming, error prone and risks missing analysis "bugs" in some of the source trees. We apply automatic treebank conversion techniques, which are based on the description based tree rewiting methodology presented in Section 2.2. above. A similar approach to treebank conversion is presented in [Frank,01]. We apply this technique to AP100 to generate an automatically converted ("groomed") resource AP100 . We compare the manually groomed AP100 with the automatically converted AP100 in terms of standard labelled bracketing measures and precision and recall results on transformation events.

Our next move is based on the observation that grammars G and G extracted from both the manually and the automatically groomed AP100 and AP100 look very much like compiled out (expanded) versions of flat LFG-82 [Kaplan and Bresnan,82] style grammars which involve regular expressions (such as optionality or Kleene-star) in CFG rule RHSs. We conduct two experiments to relate G to a flat LFG-82 style grammar. The first is based on FST machinery: given a LHS we collect the corresponding RHSs in the grammar and union them into a regular expression. We then use FST machinery to compile the regular expression into a FSA and apply minimisation. The resulting net is then retranslated into a regular expression. In the second experiment, for a given LHS we hand code a regular expression based rule RHS based on the attested RHSs in the grammar. We verify that the language defined by the regular expression includes the language of the attested RHSs. Regular expression based rule RHSs can provide compact representations of CFG resources facilitating automatic (and manual) feature-structure annotation. For reasons of space, here we can only briefly outline our experiments. We hope to report on the regular expression based "compactions" elsewhere.

Finally, we manually recode the first 50 parse-annotated sentences from the generalised and manually corrected AP100 using "vanilla flavour" Xbar design principles in AP50 . From this we extract a grammar G . In all the grammars extracted in our experiments, the NP fragments pose the largest, most varied and complicated annotation tasks. We extract the NP fragment from AP50 . We manually construct gold-standard feature-structure annotations for the Xbar-based NP fragment and develop regular expression based feature-structure annotation principles for the NP fragment. We evaluate the results of automatic annotation against the gold standard in terms of precision and recall measures. We also compare CFG rule growth rates for AP100 and AP50 .

5 Structure Preserving Grammar Compaction

Treebanks tend to express considerable amounts of information in monadic category labels.[3] Typically, this results in large lexical tag sets and in a large number of non-lexical categories. The AP treebank features

lexical tags	$>$ 180
non-lexical cats	$>$ 50

[2] Actually, to a large extent the gold standard annotation of is done automatically: the majority of the CFG rules in are identical to rules in . The corresponding annotations from the already available gold standard are moved automatically to the manually corrected and it is only the residue – that needs to be annotated manually to create the gold standard feature-structure annotations for .

[3] In addition, some treebanks allow adornment of CFG nodes in terms of a simple feature decoration mechanism, e.g. Penn-II supports features indicating grammatical functions or coindexation.

This, together with the flat treebank analyses, results in many often only minimally different CFG rule types. From the AP100 fragment we extract 509 rule types,[4] that is more than 5 distinct types per tree. (Automatic) feature-structure annotation repackages information: often - particularly so in the case of lexical entries - it "copies" information originally encoded in monadic category labels into feature-structure annotations (and hence the resulting feature-structure representations). As a consequence we can collapse highly discriminating tags into "supertags". In the original tag set, number information in nominals is encoded in terms of two tags `nn1` and `nn2` for singular and plural common nouns, respectively. Once number information is represented in terms of a feature value pair in a feature-structure description `num=sg` we can collapse the two orginal tags into a supertag `n0`

```
{ nn1,nn2 } => n0
```

Collapsing tags into supertags can have the effect of collapsing grammar rules. Consider the following scenario where `jj` and `jb` are two types of adjectives:

```
n > at1 jb nn1
n > at1 jj nn1
```

Collapsing `{ jj,jb } => adj` has the effect of collapsing the two CFG rules expanding `n` into the single rule

```
 n > at1 adj nn1
```

Structure preserving grammar compaction can be applied automatically to parse-annotated treebank resources and compacted grammars can then be extracted [van Genabith, Sadler and Way,99], [Hepple and van Genabith,00]. In previous work [Sadler, van Genabith and Way, 00] applied structure preserving compaction to AP100 generating AP100 and extracted the corresponding grammars G (from AP100) and G (from AP100) with the following number of rule types:

	rule types
G	509
G	331

In order to evaluate automatic feature-structure annotation in terms of a gold standard, G was first annotated manually. This was followed by the development of regular expression based annotation principles. These annotation principles are then applied automatically to an unannotated version of G .

Automatic annotation is evaluated against the gold standard in terms of precision and recall results. Precision measures how many of the automatically generated annotations are correct (i.e. are in the gold standard); recall measures how many of the target gold standard annotations were in fact generated automatically:

	precision	recall
G	93.4 %	91.6 %

Automatic annotation is more partial than incorrect. The results for the compacted grammar G provide the base line for our further treebank grooming experiments reported below.

[4] As opposed to rule tokens.

6 Manual Correction

In the following sections we move the automatically generalised (with respect to category labels) treebank fragment to an Xbar-based encoding in stages. The first step is a manual correction/clean up of the generalised fragment. The idea is to correct obvious errors or omissions in the generalised trees. The clean up step is performed in the general spirit of the flat treebank design philosophy (i.e. at this stage we do not introduce Xbar assumptions). As a rough guideline for restructuring transformations, we use analyses attested elsewhere in the fragment. We illustrate this step by way of an example. Consider the very first NP in the (generalised) treebank fragment:

```
                                     np
           ___________________________|___________________________
          /               |      |      |     |       |           \
         np              n0      n0    n0     p      np          tgp
   _______|______         |      |      |     |     __|__          |
  /       |      \        |      |      |     |    /     \         |
 n0      n0     pnct   18-foot  jump  shot  with  num    n0       v0
  |       |       |                                |      |        |
  |       |       |                                |      |        |
alvis  rogers  gen-sym                            one  second  remaining
```

The example shows a missed PP post-modifier grouping at the tail of the NP. In addition we also relabel the initial NP daughter as a determiner phrase DETP:

```
                             np
           ___________________|____________________
          /               |      |      |          \
        detp             n0      n0    n0          pp
   _______|______         |      |      |      _____|_____
  /       |      \        |      |      |     /           \
 n0      n0     pnct   18-foot  jump  shot    p           np
  |       |       |                           |     _______|_______
  |       |       |                           |    /       |       \
alvis  rogers  gen-sym                       with  num    n0       tgp
                                                   |      |        |
                                                   |      |        |
                                                  one  second     v0
                                                                   |
                                                                   |
                                                               remaining
```

Notice that it would be possible to further restructure the initial DETP daughter into:

```
           detp
        ___|____
       /        \
      np         gen
    __|_          |
   /     \        ,
  n0      n0
   |       |
alvis  rogers
```

However, here we abstain from this restructuring as this analysis is not attested elsewhere in the treebank fragment.

Manual correction involved a total of 143 transformation events on the 100 trees in the generalised treebank fragment. The major transformation types can be classified in the following typology:

- missed groupings
- single node relabelling
- attachment restructuring
- surrounding punctuation

Due to space limitations here we cannot give a detailed discussion of the transformation types. The first two transformation types were illustrated in the example above. Attachment restructuring mainly concerns certain instances of pre- and post-modifier phrases while surrounding punctuation restructures opening and closing punctuation as daughters of the same mother node.

We then extract a grammar G from the generalised and manually corrected treebank fragment. As expected, the extracted rule set is "cleaner" and "improved" compared to G . To give but a single illustration based on the NP example discussed above, G features a rule

```
np -> np n0 n0 n0 p np tgp
```

derived from the top level NP local tree while in G this rule is replaced by

```
np -> detp n0 n0 n0 pp
```

G features the same number of rule types as G , however, G has more rule tokens (applications of rules) as evidenced by the labelled bracketing counts. Intuitively this means that the manually corrected treebank has more structure:

	# rule types	labelled bracketing
G	509	1318
G	331	1318
G	331	1422

In order to evaluate manual correction in terms of how well the extracted grammar performs under automatic feature-structure annotation, we manually annotate G with feature-structure information. The results of the manual annotation provide a gold-standard for automatic annotation. We then develop a set of regular expression based annotation templates and apply these automatically to an unannotated version of G . We then compare the manually annotated grammar with the automatically annotated grammar in terms of precision and recall results and contrast the results obtained with the corresponding figures for G

	precision	recall
G	93.4 %	91.6 %
G	93.7 %	94.9 %

Precision and recall results for grammar G extracted from the generalised and manually correct treebank fragment are improved. This is as expected. G shows fewer quirks and "odd" rules compared to G. The strongest reflection of this is in the 3.3% increase in recall, meaning that automatic annotation generates more of the gold-standard target annotations compared to G.

7 Automatic Correction

Manual correction or clean up of a treebank resource is very time consuming, expensive, error prone and partial in that it risks missing appropriate corrections. In this section we apply a method for automatic treebank conversion, based on rule-based rewriting of trees encoded in a flat tree description language. The method is a special application of the automatic feature-structure annotation methodology developed in [Frank,00], yet applied to convert the structure of treebank trees. A similar approach to treebank conversion is presented in [Frank,01].

In the first step, trees are automatically translated into a set (a conjunction) of basic tree description predicates of a general tree description language. The tree description language features a number of basic and derived predicates:

Basic tree predicates (`A,B` node identifiers)

`arc(A,LA,B,LB)`	B daughter of A, with category labels LA,LB
`prec(A,B)`	immediate precedence A–B
`lex(A,LA,Lex)`	lexical node A with category label LA

Derived tree predicates (`A,B,..` node identifiers)

`dom(A,B)`	immediate dominance
`dom_x(A,C)`	dominance
`prec_x(A,C)`	precedence
`first_d(A,X)`	X first daughter of A
`last_d(A,Y)`	Y last daughter of A
`c_label(A,CA)`	CA functional label of A
...	possibly more ...

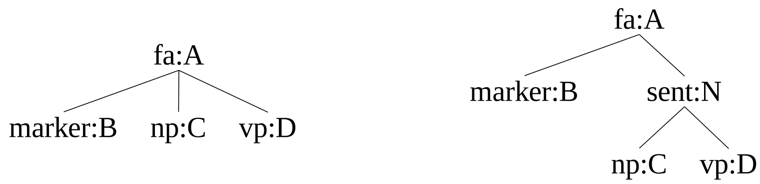
```
where:
dom(A,B) :- arc(A,_,B,_).
dom_x(A,C) :- dom(A,C)   (dom(A,B) ∧ dom_x(B,C)).
prec_x(A,C) :- prec(A,C)   (prec(A,B) ∧ prec_x(B,C)).
```

We then define parameterised tree transformation rules, using predicates (both basic and derived) of the tree description language. The rules apply to partial tree configurations, and rewrite (part of) the tree predicates, to derive a new set, and thus a new tree description. The new tree description is then reconverted to a standard tree representation. Tree transformation rules are of the form `ID :: CD >> TD` where `ID` is a rule identifier, `CD` a constraining tree description and `TD` a set of transformation operations.

Transformation rules are best explained by way of example. The following transformation rule identifies `np vp` sequences preceded by `marker`, preposition or prepositional phrase daughters (`subord_marker`) in `fa`, `fn` or `relcl` constituents (`sub_cl`), and "lowers" the `np vp` sequence below a newly created `sent` node within the original `fa`, `fn` or `relcl` constituent:

```
sub_cl ::  sub_cl(A,CA),
```

```
    arc(A,CA,B,CB), subord_marker(A,CA,CB), prec(B,C),
    arc(A,CA,C,"np"), prec_x(C,D),
    arc(A,CA,D,"vp"), last_d(A,Y)
>>  new_dtr(A,N,"sent"),
    lower_subtree(A,C,Y,N,"sent").
```



We developed a set of 13 transformation rules. Together these transformation rules effect a total of 110 transformation events, compared to 143 manual transformations, with the manual transformation events being a near superset of the automatic transformation events (one of the automatic transformation events is not in the manual transformation events):

	AP100	AP100
# of transf. ev.	143	110

RuleType	# of transf. ev.
sub_cl	23
fn_to_sbar	3
fn_to_sent	26
existential	3
inv_sent	18
surrounding_commas1	11
surrounding_commas2	1
xtraposed_mods	2
detp_gen	8
adv_num_detp	8
unary_sent_np	1
pre_conj_conj	1
flat_coord	5

We evaluate grammar G extracted from the automatically corrected treebank fragment AP100 against the grammar G extracted from the manually corrected treebank fragment AP100 in terms of standard labelled bracketing (LB), labelled precision(LP) and labelled recall (LR) measures from parsing. Compared to the generalised treebank fragment AP100 , both manual and automatic correction introduce more structure, manual correction more so than automatic correction:

	LB
AP100	1318
AP100	1422
AP100	1389

Labelled precision measures how many of the labelled bracketings in the automatically corrected treebank fragment are correct as defined by the manual correction. Labelled recall measures how many of the correct target labelled bracketings as defined by manual correction are generated by automatic correction:

	labelled precision	labelled recall
AP100	97.8 %	95.6 %

A more phenomenon oriented precision and recall measure is shown by the following figures, where we evaluate precision and recall of transformation events (rather than (labelled) bracketing divergences between manual and automatic corrections reflected in the LP and LR results in the table above). Here, 1 out of 110 automatic transformation events is not in the set of manual transformations; 34 manual transformation events are missed.

phenomenon oriented	precision	recall
AP100	99 %	76 %

8 Regular Expression Based Compaction

Grammars G and G look very much like flat, compiled out versions of LFG-82 style grammars [Kaplan and Bresnan,82]. These early LFG grammars allow regular expressions such as Kleene star "`*`" and optionality "`(..)`" in rule RHSs such as the following (here we only show the CFG part):

```
vp -> v (np) (np) (vp) (sent) pp*
```

Regular expressions in rule RHSs allow a compact representation of a multitude of CFG rules and this can greatly aid perspicuity for manual and automatic feature-structure annotation tasks. In our previous work we developed regular expression based feature-structure annotation principles designed to match large numbers of CFG rules extracted from treebank resources. In an inversion of this idea we could try to compact a treebank grammar using regular expressions in rule RHSs. Because of space limitations here we can only outline the basic ideas. We hope to report on this elsewhere. In our work to date we have explored two methodologies: the first is based on finite-state technology (FST) and automatic conversion into a Finite State Automaton (FSA); the second on manual conversion.

The basic idea in FST based automatic conversion is simple: for a given rule LHS, collect all RHSs attested in the grammar.

```
LHS > RHS1
LHS > RHS2
 ..     ..
LHS > RHSn
```

Each RHS constitutes a (trivial) regular expression. For a given LHS union all RHSs into a "disjunctive" regular expression:

```
[RHS1 | RHS2 | .. | RHSn]
```

Use FST tools to convert the regular expression into an automaton. Apply minimisation to the resulting automaton. Convert back to a regular expression. Automatic conversion along these lines (especially of large numbers of rule RHSs) does not always result in "readable" regular expressions. In fact it can be useful to manually split large sets of rule RHSs into "natural" classes prior to automatic conversion.

Manual conversion is based on the following idea. Suppose that a treebank attests the following rules:

```
np > det n0
np > n0 pp
```

these can be "compacted" into

```
np > (det) n0 (pp)
```

Note that the compaction is linguistically well motivated but that the language defined by the regular expression in the compaction is not the same as the language defined by the attested rule RHSs:

```
(det) n0 (pp) =/= det n0 | n0 pp
```

In fact, the compaction covers two unattested possibilities:

```
(det) n0 (pp) = det n0 | n0 pp | n0 | det n0 pp
```

We conducted a manual experiment which compacted G into 20 rules with regular RHSs. We used our regular expression interpreter to verify that for each LHS in the grammar, the language defined by the regular expression based rule RHSs properly includes the language defined by the RHSs attested in the original grammar.

9 Xbar-based Annotation

In our final experiment we transform the first 50 parse-annotated sentences of the generalised and manually corrected version AP100 of the treebank fragment into a "vanilla flavour" Xbar encoding AP50. Our choice of Xbar design principles is pragmatic and similar to other (computational) grammars in the literature: [Pollard and Sag, 94], [Butt et al,99]. It does not constitute deep syntactic theorising! Our design principles are:

1. three projection levels: X0, X1 and XP
2. X1 only present when required (as in adjunction)
3. complements are taken at X0 in one go (this means that we will have a number of flat non-binary branching rules)
4. adjuncts are taken at X1 recursively
5. only X2 (maximal projection) can act as complement/adjunct
6. no ϵ-productions

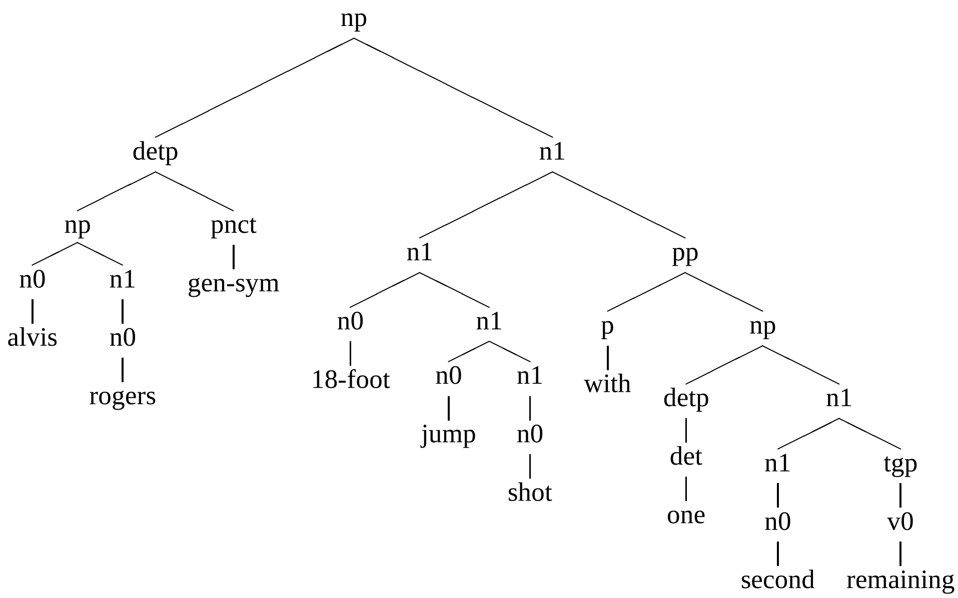
Principle 1 is pretty uncontroversial. Principle 2 (minimal representations: X1 only when needed) will have interesting consequences for the number of CFG rules extracted from the Xbar encoding. Principle 3 aids in functional annotation of subcategorisable grammatical functions (OBJ, OBJ2, XCOMP, COMP, ...). Principles 4 and 5 are again fairly standard assumptions. Principle 4 has interesting consequences for probabilistic versions of CFGs extracted from such encodings. Finally, principle 6 is a reflection of the "surfacy" nature of the CFG component in LFG based grammar architectures.

Taken together principles 1 – 6 constitute fairly "mild" Xbar assumptions. Nevertheless, they are linguistic idealisations which motivate some of our "systematic" vanilla flavour Xbar violations:

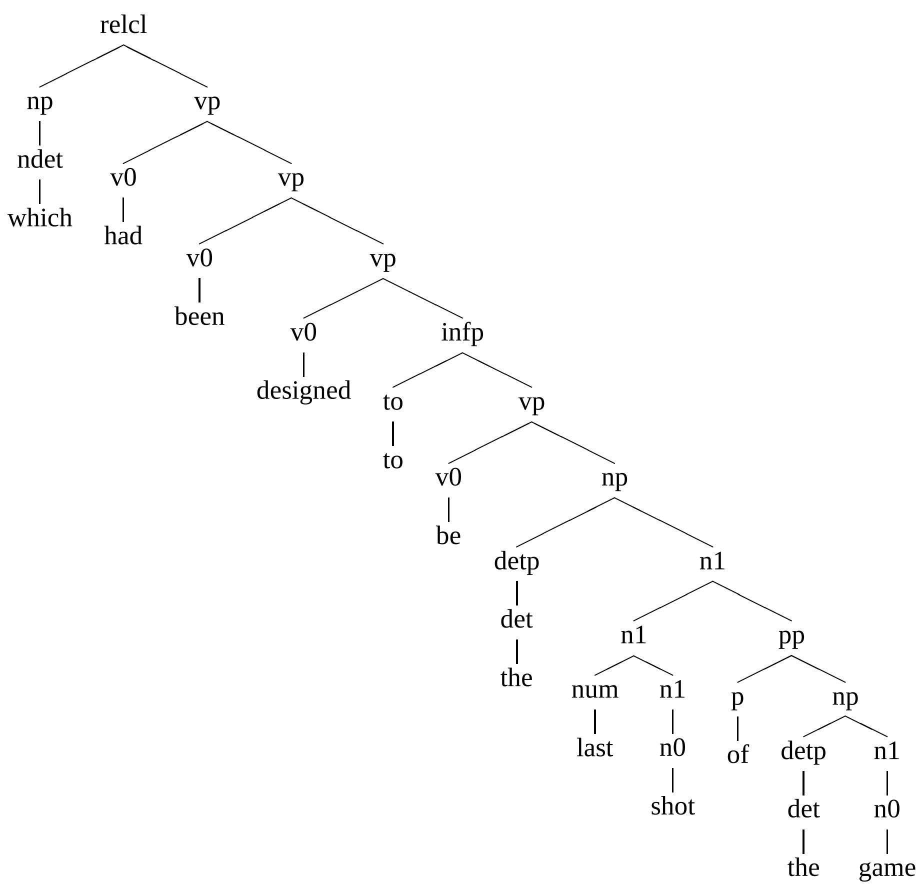
1. inside NP we allow for N0 as modifier of N1 and as a direct modifier daughter of NP (i.e. we do not always introduce N1 in N0 modifier configurations)
2. because of Xbar principle 3 we need to allow adjuncts between complements at X0
3. we keep some (traditional) treebank categories such as SENT, VP, TGP ...
4. we do not provide Xbar accounts of coordination and punctuation

Xbar violation 1 is to account for our treatment of nominal compounds. Violation 2 (of principles 3 and 4) is motivated by the data, i.e. adjuncts occurring between complements. Violation 3 is less a violation than a partial commitment to Xbar-based generalisations. Violation 4 is again a pragmatic decision.

We illustrate both our Xbar design principles and violations with two examples. The first is a recoding of the initial NP from the first sentence in the treebank fragment:



The second is a complex relative clause:



Note that both trees are exclusively binary branching and that together with a few additional assumptions each local subtree identifies head, modifier and complement relations. The additional assumptions include, for example, that in general for English NPs the head projection follows the right branch (given our treatment of nominal compounds this does not follow from the vanilla flavour Xbar encoding). These additional assumptions together with our Xbar design principles and violations should guide automatic feature-structure annotations to a much greater extent than the original flat treebank grammar design.

In order to evaluate the Xbar recoding, we extract grammar G from fragment AP50 . We will give a few general observations on the extracted grammar. Compared to our other grammars the Xbar-based grammar features shorter rule RHSs, fewer rule types but substantially more rule tokens (local trees of depth one) per tree. In the table below we list the number of rule types extracted against increasing corpus size for the generalised and manually corrected fragment AP100 and the Xbar based AP50 :

# sentence	# rules AP50	# rules AP100	difference	difference %
05	47	37	+10	+27.03
10	66	56	+10	+17.86
15	78	78	0	0
20	94	102	- 8	-7.84
25	102	122	-20	-16.39
30	114	135	-21	-15.56
35	133	159	-26	-16.35
40	144	174	-30	-17.24
45	151	184	-33	-17.93
50	161	201	-40	-19.90

As expected, grammar G extracted from AP50 grows more slowly than grammar G extracted from AP100 . Xbar grammars features fewer, more general rules than flat treebank grammars with unabated rule accession rates. However, the effect is less pronounced than could be expected. The reasons for this are that

- according to our Xbar principle 3 which states that X1 is only featured if required, we get duplication of complementation rules at XP and X1 level depending on the presence or absence of modifiers:

  ```
  vp -> v0 np          v1 -> v0 np
  vp -> v0 np np       v1 -> v0 np np
  ```

  Ironically, more parsimonious parse trees yield a larger rule base.

- the fragments considered are too small to see the full effects of Xbar vs. treebank based rule accession rates.

In all our extracted grammars, the NP fragment constitutes the most varied and complicated subsection and thus poses the most difficult section for automatic feature-structure annotation. We isolate the NP section from G and manually annotate it with feature-structure information to create a gold-standard against which to evaluate automatic annotation. We then develop regular expression based feature-structure annotation principles and automatically apply them to an unannotated version of the NP fragment from G . We then compare the results of automatic annotation against the gold standard in terms of precision and recall results.

As most of the CFG rules extracted from AP50 are unary or binary branching (the only exceptions are rules accounting for coordination, punctuation and complementation), most of the regular expressions in the annotation principles are simple unions or contain underspecified category symbols:

```
n1:A > n1:B [adjp:C|advp:C|pp:C|recl:C|tgp:C|vp:C]
     @ [A=B, A:adjn $ C]
v1:A > v1:B __:C
     @ [A=B, A:adjn $ C]
```

The first annotation principle enumerates the possible `n1` post-modifiers in the grammar. The reason we cannot underspecify the modifier category symbol (as in the second `v1` annotation principle) is that the grammar extracted allows for binary branching

```
n1 > n1 n1
```

rules where the head projection follows the right branch. In general, annotation principles for G are no longer underspecified (with the exception of principles dealing with the non-binary branching complementation, coordination and punctuation rules) as there is no need to determine subsequences in flat rule RHS for proper annotation as was the case for flat treebank rules. Automatic annotation should thus be much simpler. The precision and recall results for the NP fragment of G and G (for the first 50 sentences) support this expectation:

	precision	recall
G	93 %	94.4 %
G	100 %	100 %

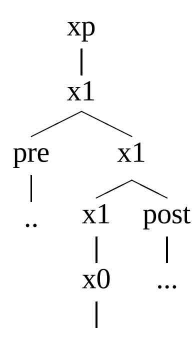
10 Conclusions and Further Work

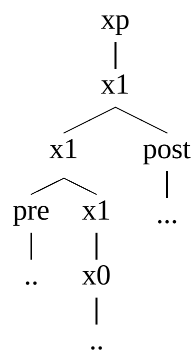
We have presented a number of experiments that move a flat treebank grammar fragment in stages to a vanilla flavour Xbar encoding. We have evaluated the resulting grammars with respect to how they perform in automatic feature-structure annotation tasks. Some of our experiments involved manual recoding of treebank resources. Manual recoding is extremely time consuming and error prone. We have investigated to what extent treebank grooming can be automated and have developed, applied and evaluated automatic treebank conversion procedures. We have sketched regular expression based compaction of rule RHSs extracted from treebank grammars. We hope to be able to expand on this in further work.

The results of our final Xbar encoding can be summarised as follows: vanilla flavour Xbar coding can be scaled but is extremely time and resource intensive. It provides a smaller, cleaner, and more general rule base which makes it easy to spot inconsistencies in the treebank (simply extract the grammar and search for odd rules). Xbar-based trees support (and in fact drive) automatic feature-structure annotation, a result that tallies well with work in theoretical linguistics [Bresnan,01]. Xbar-based trees require fewer and less partial automatic feature-structure annotation principles. However, Xbar design principles force subtle and time consuming scope/attachment decisions and currently do not support a notion of underspecification. This seriously limits its application in large scale treebanking tasks. Small and general rule bases are bought at the price of multiplying (often exploding) analysis possibilities with ensuing load under parsing and generation. What is more, Xbar-based rule sets are ill suited to probabilistic approaches to parsing (and generation) for at least the following two reasons:

First, Xbar-based CFG trees involve large numbers of rule tokens. Under standard PCFG assumptions, the probability contributed by each such rule token introduces an independence assumption into the computation of the probability of the tree and thus weakens the probability model.

Second, the standard Xbar assumption that pre- and post-modification (adjuncts) revolves around a recursive X1 level means that PCFG versions of such grammars in principle cannot distinguish high and low attachment possibilities in the presence of both a pre- and a post-modifier at a local X1 configuration





as the product of the individual rule probabilities involved in both trees are exactly the same.

Bibliography

(Bresnan,01) Joan Bresnan, *Lexical Functional Syntax*, Blackwell Publishers, Oxford, 2001.

(Butt et al,99) Miriam Butt, T.H. King and F. Second, *A grammar writer's cookbook*, Stanford, Calif. : CSLI Publications, 1999.

(Charniak,93) Eugene Charniak, *Statistical language learning*, Cambridge, Mass : MIT Press, 1993.

(Frank,01) Anette Frank, *Treebank Conversion. Converting the NEGRA Treebank to an LTAG Grammar*, in: *Proceedings of the Workshop on Multi-layer Corpus-based Analysis*, Iasi, Romania, 2001, pp. 29–43.

(Frank,00) Anette Frank, *Automatic F-Structure Annotation of Treebank Trees*, in M. Butt and T.H.King (Eds.), Proceedings of the LFG00 Conference, 2000, University of California, Berkeley, CSLI Online Publications, Stanford, CA, 2000, pp. 139-159, http://www-csli.stanford.edu/publications/

(Frank et al,01) Anette Frank, Louisa Sadler, Josef van Genabith und Andy Way, *From Treebank Resources to LFG F-Structures. Automatic F-Structure Annotation of Treebank Trees and CFGs extracted from Treebanks*, to appear in: Abeille, A. (ed): *Treebanks. Building and using syntactically annotated corpora*, Kluwer Academic Publishers, The Netherlands, to appear, 23 pp.

(Hepple and van Genabith,00) Mark Hepple and Josef van Genabith, *Experiments in Structure-Preserving Grammar Compaction*, 1st Meeting on Speech Technology Transfer, Universidad de Sevilla and Universidad de Granada, Seville, Spain, Nov. 6th - 10th, 2000.

(Kaplan and Bresnan,82) Ron Kaplan and Joan Bresnan, *Lexical-Functional Grammar: a Formal System for Grammatical Representation*, in Bresnan, Joan (ed.), *The Mental representation of grammatical relations*, Cambridge, Mass : MIT Press, 1982.

(Krotov et al,98) Krotov, A., M. Hepple, R.Gauzauskas and Y. Wilks, Compacting the Penn Treebank Grammar, in *Proceedings of COLING/ACL'98*, Montreal, Canada, pp.699-703

(Leech and Garside,91) Leech, G. and R. Garside, *Running a Grammar Factory: on the Compilation of Parsed Corpora or Treebanks*, Mouton de Gruyter, Berlin, 1991, pp. 15-32.

(Pollard and Sag,94) Pollard, Carl and Sag, Ivan. *Head-driven phrase structure grammar*, Stanford : Center for the Study of Language and Information; Chicago; Univer, 1994

(Sadler, van Genabith, Way,00) Sadler,L., J. van Genabith and A. Way, *Automatic F-Structure Annotation from the AP Treebank*, in M. Butt and T.H.King (Eds.), Proceedings of the LFG00 Conference, 2000, University of California, Berkeley, CSLI ONline Publications, Stanford, CA, http://www-csli.stanford.edu/publications/

(van Genabith, Sadler and Way,99) Josef van Genabith, Louisa Sadler and Andy Way, *Structure Preserving CF-PSG Compaction, LFG and Treebanks*, Journees ATALA, Corpus annotes pour la syntaxe, Actes, (Proceedings ATALA Workshop - Treebanks), l'Universite Paris 7, France, 18-19 Juin 1999, pp.107-114, also http://talana.linguist.jussieu.fr/treebanks99/

Agentivity and Suffix Selection[1]

Carmen Kelling

University of Konstanz

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications
http://csli-publications.stanford.edu/

1 This is a revised version of the paper presented at LFG2001 in Hong Kong. I am grateful to Miriam Butt, Bruce Mayo, Judith Meinschaefer, Nikolaus Schpak-Dolt, Martine Lorenz-Bourjot, and Christoph Schwarze for discussion of the issues in this paper. Special thanks go to Bruce Mayo and Anne Xavier-Glatz for checking my English.

1 Introduction

In French, as in most languages, the process of deriving nouns from verbal stems is a fairly regular process, and one which has been argued to involve argument structure processes (e.g., Grimshaw 1990, Laczkó 2000, Markatonatou 1995). In particular, there has been much discussion of the licensing of syntactic arguments of the deverbal noun and the effect of the nominalizing suffix on the argument structure of the base verb. However, there has been rather less work on the more subtle issue of which verbal stems are compatible with which nominalizing suffixes and why. One recent paper on this question is Barker 1998 on English *-ee,* as in *attendee*.

In this paper, I would like to show which principles determine the selection of the French suffixes *-age* and *-(e)ment*.

For noun derivation in French, several authors have proposed accounts of suffix selection, claiming that *-age* attaches to transitive verbs and *-(e)ment* is added to intransitive, reflexive or passivized verbs. We shall see that this is not sufficient for an account of these nominalization patterns. Rather, the lexical argument structure properties of the verbal bases must be analyzed more precisely in terms of Dowty's (1991) notion of Proto-Roles.

The paper is structured as follows: First, I will give a descriptive analysis of derivations with *-age* and *-(e)ment*, and a short summary of previous studies of the competition between these two suffixes, showing that not all cases are covered by the traditional analyses (the Transitivity and the Action Hypothesis); second, I will present a proposal within the framework of Lexical Mapping Theory (LMT), which I will call the [– o] Hypothesis and which explains most, but not all of the data, and third, I will present an approach elaborating Dowty's (1991) Proto-Roles with LMT. I call this approach the Agentivity Principle.

However, there are still counterexamples. They will be explained by a diachronic view of the data, and also by blocking mechanisms that play a role in the derivation processes discussed in this paper. The membership in special vocabulary domains may also produce counterexamples. Words that have disappeared from the lexicon, and words that came into the language within the last decades give further evidence for the analysis proposed here.

The conclusion will be that the selection of the French suffixes *-age* and *-(e)ment* is not due to syntactic properties of the base such as transitivity / intransitivity, but to a calculation of Proto-Role property sets.

2 Descriptive analysis

Historically there are a number of suffixes by which French nouns expressing an action or the result of an action have been derived from verbs. The following competing noun-forming suffixes are attested:

(1)	*-ade*	*bousculade* 'hustle'
	-age	*abattage* 'felling'
	-aison	*comparaison* 'comparison'
	-ance	*espérance* 'hope'
	-at	*résultat* 'result'
	-ation	*centralisation* 'centralization'
	-ée	*traversée* 'going through'
	-(*e)ment*	*glissement* 'gliding'
	-erie	*tricherie* 'cheating'
	-is	*cliquetis* 'jingling, clicking'
	-ure	*blessure* 'injury'
	(*-ing*	*feeling* 'feeling') (in loan words)

These suffixes further compete with conversion and 'affixless derivation' as in (2):

(2) *le venir* 'coming'
le vol 'flight'
la nage 'swimming'

However, the productivity of the mentioned suffixes varies widely: above all, -*ation* and its allomorphs is productive today, and also -*age* and -*(e)ment*, the first being a non-native suffix combining almost exclusively with learned stems; the latter are native suffixes choosing native stems as bases. Affixless derivation is productive in familiar speech.

In the following, I will only examine transparent derivations and derivatives without a meaning shift, nominals such as *gouvernement* 'government' will not be considered.

Among the French suffixes deriving nouns, -*age* and -*(e)ment* are particularly interesting in that they yield more than 400 pairs of nominalizations derived from the same verbal base. These pairs were found by searching the FRANTEXT corpus of the *Institut National de la Langue Française* (INaLF).

The examples in (3) illustrate that both suffixes under consideration here derive masculine nouns. Both are productive, and there are derivations from all three

French verb classes, suggesting considerable similarity of the formal morpho-syntactic properties of the base verbs and of the output nouns:

(3)	*-age:*	-er	(class 1):	*lavage* 'washing'
		-ir	(class 2):	*brunissage* 'burnishing'
		-re	(class 3):	*battage* 'beating'
	-(e)ment:	-er	(class 1):	*changement* 'changing'
		-ir	(class 2):	*agrandissement* 'enlargement'
		-re	(class 3):	*battement* 'beating'

3 Previous Approaches

3.1 The Transitivity Hypothesis

Although the question of which kind of verb base combines with which suffix has been the topic of many studies, the distribution of the affixes has not been explained satisfactorily. Dubois (1962, 1999), Trésor (1971ff) and Lüdtke (1978) claim that there is a tendency for *-age* to select transitive verb stems, whereas, *-(e)ment* selects intransitive, reflexive and passivized verb stems, with 'passivized' meaning a so-called 'result passive'. Consider the examples in (4) to (6):

(4)	a.	*battage* 'beating'	<	*battre*$_{tr}$ 'beat, thresh'
	b.	*battement* 'beating (of the heart)'	<	*battre*$_{intr}$ 'beat'
(5)	a.	*étirage* 'stretching (of metal)'	<	*étirer*$_{tr}$ 'stretch'
	b.	*étirement* 'stretching of oneself'	<	*étirer*$_{refl}$ 'stretch'
(6)	a.	*gonflage* 'pumping (of a tire)'	<	*gonfler*$_{tr}$ 'pump up'
	b.	*gonflement* 'expansion (of a tire)'	<	*gonflé*$_{pass}$ 'pumped up'

But there are counterexamples that cannot be explained by these approaches, see (7) and (8). The intransitive verb *fureter* 'search' selects *-age*, and the transitive *essouffler* derives *essoufflement* by adding *-(e)ment.*

(7)	*furetage* 'searching'	<	*fureter*$_{intr}$ 'search'
(8)	*essoufflement* 'losing of one's breath	<	*essouffler*$_{tr}$ 'make lose one's breath'

3.2 The Action Hypothesis

In contrast to Dubois (1962, 1999), Trésor (1971) and Lüdtke (1978), Zwanenburg (1984) and Debaty-Lyca (1986) subsume all suffixes deriving deverbal nouns under one derivation type (Zwanenburg 1984) or one suffix expressing ACTION (Debaty-Luca 1986), thus blurring the undoubtedly existing selection differences. Debaty-Luca assumes that the different suffixes are all allomorphs of a single morpheme ACTION.

3.3 Summary

These and other similar examples indicate that neither the Transitivity Hypothesis, nor even less, the Action Hypothesis can give a satisfactory explanation of suffix selection of -*age* and -*(e)ment*.

However, this first glance at the data gives the impression of a connection between -*age* and agentivity, and I shall try to model this within the LMT framework in section 4.

4 An LMT Approach: the [– o] Hypothesis

Most of the data presented in the previous sections can be elegantly explained by LMT's intermediate level of argument classification, i.e., by the intrinsic argument classification (e.g. Bresnan & Kanerva 1989, Bresnan & Zaenen 1990) for arguments and grammatical functions, using the features [± r] (for restricted and unrestricted) und [± o] (for objective and non objective) for both grammatical functions and thematic roles. The features predict the appropriate mapping of thematic roles onto grammatical functions. (9) shows the features of Grammatical Functions, (10) the intrinsic features of thematic roles.

(9) Grammatical Functions classified by features

Grammatical Functions	Features	
SUBJ	[– r, – o]	r: restricted, o: objective
OBJ	[– r, + o]	
OBJ_θ	[+ r, + o]	
OBL_θ	[+ r, – o]	

(10) Intrinsic features of thematic roles

Thematic Roles	Features	Possible Mappings
agent	[– o]	SUBJ / OBL
theme / patient	[– r]	SUBJ / OBJ
locative	[– o]	SUBJ / OBL

In the following, we will need especially the [– o] feature for the agent role. A combination of these features and a thematic role hierarchy as in (11) results in the appropriate mapping of thematic roles onto grammatical functions:

(11) Thematic Hierarchy:
agent > beneficiary > experiencer / goal > instrument > patient / theme > locative

[– o] Hypothesis on Suffix Selection

The claim here will be that the French suffixes are selected according to the [± o] / [± r] features of the first argument of their verbal base. In the competition between -*age* and -*(e)ment*, the hypothesis predicts that -*age* is chosen whenever we find a [– o] feature for the first argument and that -*(e)ment* is selected in cases where we do not find a [– o] feature for the first argument of the verbal base.

In the following, I give the LMT analyses for some of the verbs listed in (4) to (8). Consider the analysis of *battre* in (12). French transitive *battre* 'beat' has an agent- and a patient-argument. Intrinsically, the first argument has the feature [– o] and is mapped onto the SUBJ function. The patient-argument has the feature [– r] and is mapped onto the OBJ function according to mapping principles. The [– o] Hypothesis predicts the selection of the suffix -*age*, and we get the derivative *battage.*

(12)	BATTRE	<arg1	arg2>		
	'beat'	agent	patient		
		|	|		
Intrinsic		[– o]	[– r]		
		|	|		
GF		SUBJ	OBJ	→	battage

The other transitive verbs can be analyzed in the same way. The intransitive *fureter* 'search' also presents the [– o] feature for the first argument and therefore selects the -*age* derivation. Consider (13):

(13) FURETER <arg1>
'search' agent
|
Intrinsic [– o]
|
GF SUBJ → furetage

Thus, unlike the simpler Transitivity Hypothesis, LMT predicts the correct selection of *-age* even for intransitive verbs.

For the selection of *-(e)ment*, the [– o] Hypothesis also makes the correct predictions for the intransitive and for the passivized cases. For the result passive[2] cases like *est gonflé* in *Le pneu est gonflé* 'The tire is pumped up', we do not have a [– o] feature, and therefore *gonflé* cannot combine with *-age*, but it combines with *-(e)ment*, giving *gonflement*.

(14) EST GONFLÉ <arg1>
'pumped up' theme
|
Intrinsic [– r]
|
GF SUBJ → gonflement

However, for the reflexive and the transitive *-(e)ment*-cases, LMT errs in predicting *-(e)ment*-selection but rather *-age*-selection because of the [– o] feature of the first argument; consider (15): we get é*tirage*, although we should get *étirement*.

2 I assume here an analysis of the result passive as copula plus an adjectival passive here. For a discussion see Bresnan (1982, 2001) and Levin/Rappaport (1986). True verbal passives trigger *-age*-selection, because there is still the [– o] feature of the first argument, whereas in the adjectival passive cases, the [– o] argument is not in the argument structure any more; compare (i) for the verbal passive and (ii) for the adjectival passive:

(i) Le pneu a été gonflé (par Max). → le gonflage du pneu par Max
theme agent
[– r] [– o]

With *gonflement*, there is no [– o] feature, and therefore, the indication of an agent is impossible:

(ii) le gonflement du pneu *par Max

(15)	S'ÉTIRER	< arg1	arg2>		
	'stretch oneself'	agent	patient		
		|	|		
Intrinsic		[– o]	[– r]		
		|	|		
GF		SUBJ	OBJ	→	*étirage
			↓		
			se		

It can be observed nevertheless that when an agentive verb selects *-(e)ment*, the agent is in some sense 'worse' or 'less agentive' than in the cases where *-age* is selected. So we evidently need a finer grained notion of agentivity in order to account for the data examined here. With Dowty's Proto-Role approach we can in fact obtain such a framework.

5 Proto-Roles (Dowty 1991)

In Dowty's (1991) proposal, there are no atomic roles such as AGENT and PATIENT, but he assumes different entailments or properties for prototypical agent and patient respectively. He lists the following AGENT properties:

(16) a. volitional involvement in the event or state
b. sentience (and / or perception)
c. causing an event or change of state in another participant
d. movement (relative to the position of another participant)
e. exists independently of the event named by the verb

Dowty further assumes the Argument Selection Principle in (17):

(17) **Argument Selection Principle** (Dowty 1991: 576)
In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object.

Howewer, Dowty's Argument Selection Principle does not account for the selectional behavior of *-age* and *-(e)ment*: e.g., there is nothing said about argument selection for the passive voice or for reflexive verbs[3].

6 Elaboration of Proto-Roles and LMT

6.1 Previous elaborations of LMT and Proto-Roles

As far as I am aware, previous approaches have not attempted to elaborate LMT with Proto-Roles, which would allow for a gradation of agentivity (Zaenen 1993, Ackerman / Moore 2001). In her 1993 paper, Zaenen already presents a combination of Dowty's Proto-Roles with LMT, replacing the LMT Role Hierarchy by Proto-Role Properties. However, the principles she introduces, serving to ensure mapping of unergative and unaccusative arguments, meet the same problems as standard LMT mapping principles for the data to be analyzed here: they do not distinguish between 'good' and 'worse' agents.
What we need is a selection criterion based on the Proto-Agent Properties proposed by Dowty.

6.2 Agentivity Principle

In order to account for the data presented in this paper, I assume the following Agentivity Principle:

Agentivity Principle
The French suffix *-age* combines with verb stems whose first argument is proto-agentive, whereas the French suffix *-(e)ment* combines with verb stems whose first argument is less proto-agentive.

In section 7, I will present an Analysis of the data using the Agentivity Principle.

7 Analysis of the data using the Agentivity Principle

In the following examples, the first arguments decide the choice of *-age* or *-(e)ment*, and therefore, I will only list the agent-properties of the first arguments.
In the first case in (18), (19) and Figure 1.1 and 1.2 we will deal with the transitive vs. intransitive case (*battre*$_{\text{tr}}$ 'beat, thresh' vs. *battre*$_{\text{intr}}$ 'beat').

3 But see Ackerman / Moore 2001.

The number of proto-agent entailments or properties (a: volition, b: sentience, c: cause, d: movement, e: existence) is important for the choice of *-age* and *-(e)ment* if both can in principle attach to one and the same verbal stem: *-age* is added if many proto-agent properties are present, *-(e)ment* is attached if there are fewer. In a sentence like (18) we have four agent properties; whereas, in (19), there is only one agent property.

(18) Max bat les tapis. > battage
'Max beats the carpets.'

PRED	agent-properties of arg_1	suffix	derivation
battre <SUBJ OBJ>	a, b, c, e	-age	battage

Figure 1.1

(19) Le cœur bat. > battement
'The heart beats.'

PRED	agent-properties of arg_1	suffix	derivation
battre <SUBJ>	e	-(e)ment	battement

Figure 1.2

In the second case, the transitive vs. reflexive distinction: *étirer*$_{tr}$ 'stretch' vs. *s'étirer*$_{refl}$ 'stretch', the number of proto-agent properties is decisive for the choice of *-age* and *-(e)ment* again: for the transitive verb, we have more proto-agent properties than for the reflexivized verb, as seen in (20) and (21) and Figures 2.1 and 2.2. For the first argument in (20), there are four agent-properties, and *-age* is attached. For the first argument in (21) there are only three agent-properties, and accordingly *-(e)ment* is selected.

(20) Max étire le métal. > étirage
'Max stretches the metal.'

PRED	agent-properties of arg_1	suffix	derivation
étirer <SUBJ OBJ>	a, b, c, e	-age	étirage

Figure 2.1

(21) Max s'étire en baillant. > étirement
'Max stretches himself, yawning.'

PRED	agent-properties of arg_1	suffix	derivation
s'étirer <SUBJ>	a, b, e	-(e)ment	étirement

Figure 2.2

Nearly the same holds for the third case, active *gonfler* and the 'result passive' *est gonflé* in (22) and (23). The difference is that the argument in the adjectival passive case has only one agent property, and therefore we get an -*(e)ment* derivative.

(22) Max gonfle le pneu. > gonflage
'Max pumps the tire up.'

PRED	agent-properties of arg_1	suffix	derivation
gonfler <SUBJ OBJ>	a, b, c, e	-age	gonflage

Figure 3.1

(23) Le pneu est gonflé. > gonflement
'The tire is pumped up.'

PRED	agent-properties of arg_1	suffix	derivation
est gonflé <SUBJ>	e	-(e)ment	gonflement

Figure 3.2

The Agentivity Principle also explains the cases in which -*(e)ment* attaches to transitive verbal stems and the cases where -*age* combines with intransitive bases, consider (24), (25) and Figure 4.1 and 4.2: the first argument of *fureter* 'search' has four proto-agent properties, and -*age* is selected. With *essoufler* we have only two proto-agent properties, and we get an -*(e)ment*-derivative.

(24) Max a fureté (dans tous les coins). > furetage
'Max has searched (in all corners).'

PRED	agent-properties of arg_1	suffix	derivation
fureter <SUBJ>	a, b, c, e	-age	furetage

Figure 4.1

(25) Les efforts l'ont essoufflé. > essoufflement
'The efforts made him lose his breath.'

PRED	agent-properties of SUBJ	suffix	derivation
essouffler <SUBJ OBJ>	c, e	-(e)ment	essoufflement

Figure 4.2

8 Apparent counterexamples

Most derivatives with *-age* and *-(e)ment* follow the Agentivity Principle, but there exist counterexamples that can be explained by a diachronic view of the data, blocking mechanisms and the membership in special vocabulary domains.

8.1 Explaining counterexamples from a diachronic view

The *-(e)ment* suffix comes from Latin *-mentum*, deriving verbal nominals, e.g.

(26) FRANGERE 'break' > FRAGMENTUM 'piece broken off, fragment'
ORNARE 'prepare; equip' > ORNAMENTUM 'ornament; equipment'

In contrast to Latin, where *-mentum* derived nouns expressing the result or the instrument of the action conveyed by the base verb as in (26), the French suffix *-(e)ment* has derived action nouns since Old French period.
The suffix *-age* stems from Latin *-aticus*, and unlike *-mentum*, *-(a)ticus* in general did not derive nouns but adjectives; consider (27): AQUATICUS is a derivative of AQUA, and SILVATICUS of SILVA.

(27) AQUA 'water' > AQUATICUS 'of or belonging to water'
SILVA 'wood' > SILVATICUS 'belonging to woodland'

Only very few deverbal adjectives are attested, e.g. (28):

(28) DONARE 'give, present' > DONATICUS 'formally presented'

In French, there are denominal derivatives of adjectives with *-age* until the sixteenth century. Although the first deverbal nouns derived by *-age* are already attested for the thirteenth century, it is more or less productive only since the fourteenth century (examples are those in (29)), and it does not become really productive until the nineteenth century.

(29)	*chauffer* 'heat (up)'	>	*chauffage* 'heating'	13th century
	limer 'file'	>	*limage* 'filing'	14th century

So for -*(e)ment* nominalizations derived from verbs with typical agents, we should expect to find that they were introduced before -*age* is very productive for the derivation of deverbal nouns. In fact we find:

(30)	*abaisser* 'pull down'	>	*abaissement* 'pulling down'	12th century
	effacer 'efface, erase'	>	*effacement* 'effacing, erasing'	13th century

As we will see later in section 9, it is possible that -*age* derivatives can replace -*(e)ment* derivatives, and these replacements are in accordance with the Agentivity Principle. However, not all existing -*(e)ment* nominalizations can disappear, because they are stored in the lexicon and as such stable parts of the French vocabulary.
Evaluating these diachronic facts leads to the assumption that the Agentivity Principle introduced in section 6, has emerged during the recent centuries, and it is very improbable that we will find counterexamples derived since this development.

8.2 Blocking

To complete the picture, I would like to mention that there are blocking rules that may explain further counterexamples, e.g., phonological blocking (in the sense of Wurzel 1988) in (31), where the sequence -*ageage* is avoided:

(31) saccagement 'turning upside down' *saccageage.

8.3 Special vocabulary domains

Membership in a special vocabulary domain may overrule the Agentivity Principle, e.g., in the vocabulary of commerce and finance, -*(e)ment* is selected even in contexts where -*age* would be expected; consider (32):

(32) *intéresser* 'give a share' > *intéressement* 'profit-sharing'

9 Further evidence

9.1 Words that have disappeared

For *age/-(e)ment* pairs, Dubois (1962) investigated which items disappeared from the Larousse Dictionary between 1949 and 1962 (+: still in the lexicon, –: disappeared). If we examine these examples with respect to the agentivity of the first argument of their verbal base, we can see that they follow exactly the Agentivity Principle: if there are many proto-agent properties, the *-age* derivative stays in the lexicon, and if there are fewer, the *-(e)ment*-nominalization survives.

(33)	+ *affichage*	– *affichement*	'posting'
	+ *babillage*	– *babillement*	'babbling'

(34)	– *apparentage*	+ *apparentement*	'grouping (of electoral lists)'
	– *déferlage*	+ *déferlement*	'breaking (of waves)'

9.2 New words

Of course, the Agentivity Principle must also hold for the derivation of new words, and this is the case for the examples in (35) for *-age* and in (36) for *-(e)ment:*

(35) *guillemetage* 'putting in quotes'
spatulage 'typing'

(36) *crainquement, craquement, criquement...* (of steps)
zizillement (of the TV)

10 Conclusion

The degree of agentivity of the first argument of the base verb determines the combinatory possibilities of verbal stem and nominalizing suffixes *-age* and *-(e)ment* in French. This finding allows us to predict all cases including those that are exceptions for the traditional syntactic account.

The correct analysis for the distribution of *-age* and *-(e)ment* can be formulated nicely via an elaboration of LFG's mapping theory with Dowty's (1991) Proto-Role Approach. Lexical Mapping Theory accounts for correct mapping from arguments to grammatical functions, including cases of reflexivization and passivization. Proto-role properties are necessary for the expression of a gradation of agentivity.

References

Ackerman, Farrell (1992): Complex predicates and morpholexical relatedness: Locative alternation in Hungarian. In Ivan Sag and Anna Szabolcsi (eds.) *Lexical Matters*. Stanford, CA: CSLI, 55-83.

Ackerman, Farrell and Moore, John (1999): 'Telic Entity' as a Proto-Property of lexical predicates. In Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG99 Conference.* http://www.csli.stanford.edu/publications.

Ackerman, Farrell and Moore, John (2001): *Dowtyian Proto-Properties and Lexical Mapping Theory*, Paper presented at LFG01. Hong Kong.

Alsina, Alex (1996): *The Role of Argument Strucure in Grammar: Evidence from Romance*. Stanford, CA: CSLI.

Barker, Chris (1998): Episodic *-ee* in English: A thematic role constraint on new word formation. *Language* 74, 4, 695-727.

Bresnan, Joan (1982): The passive in Lexical Theory. In Joan Bresnan (ed.) *The Mental Representation of Grammatical Relations*, Cambridge, Mass: MIT Press, 3-86.

Bresnan, Joan (2001): *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan and Kanerva, Jonni M. (1989): Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20, 1-50.

Bresnan, Joan and Moshi, Lioba (1990): Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21, 147-185.

Bresnan, Joan and Zaenen, Annie (1990): Deep unaccusativity in LFG. In Katarzyna Dziwirek, Patrick Farrell and Errapel Mejías-Bikandi (eds.) *Grammatical Relations: A Cross-Theoretical Perspective*. Stanford, CA: CSLI, 45-57.

Butt, Miriam (1999): *The Development of Linking Theory in LFG*, Überblickshandout, http://www.ling.uni-konstanz.de/pages/home/butt/.

Corbin, Danielle (1987): *Morphologie dérivationnelle et structuration du lexique*, 2 vol. Tübingen: Niemeyer.

Debaty-Luca, Thierry (1986): *Théorie fonctionnelle de la suffixation (Appliquée principalement au français et au wallon du Centre).* Paris: Les Belles Lettres.

Dowty, David R. (1991): Thematic proto-roles and argument selection. *Language* 67, 3, 547-619.

Dubois, Jean (1962): *Etude sur la dérivation suffixale en Français moderne et contemporain*. Paris: Larousse.

Dubois, Jean and Dubois-Charlier, Françoise (1999): *La dérivation suffixale en français.* Paris: Nathan.

Grimshaw, Jane (1990): *Argument Structure*. Cambridge, MA: The MIT Press.

Laczkó, Tibor (2000): Derived nominals, possessors, and Lexical Mapping Theory. In Miriam Butt and Tracy Holloway King (eds.) *Argument Realization*. Stanford, CA: CSLI Publications, 189-227.

Levin, Beth and Rappaport, Malka (1986): The formation of adjectival passives. *Linguistic Inquiry* 17 / 4, 623-661.

Lüdtke, Jens (1978): *Prädikative Nominalisierungen mit Suffixen im Französischen, Katalanischen und Spanischen*. Tübingen: Niemeyer.

Markantonatou, Stella (1995): Modern Greek deverbal nominals: an LMT Approach. *Journal of Linguistics* 31, 267-299.

Meyer-Lübke, Wilhelm (1921, ²1966): *Historische Grammatik der französischen Sprache, Zweiter Teil: Wortbildungslehre*. Heidelberg: Winter. – Zweite, durchgesehene und ergänzte Auflage von Joseph M. Piel. Heidelberg: Winter.

Rappaport, Malka (1983): On the nature of derived nominals. In Lori Levin, Malka Rappaport and Annie Zaenen (eds.) *Papers in Lexical-Functional Grammar*. Bloomington: IULC, 113-142.

Sadler, Louisa and Spencer, Andrew (1998): Morphology and argument structure. In Spencer, Andrew and Zwicky, Arnold M. (eds.) *The Handbook of Morphology*. Oxford: Blackwell, 206-236.

Trésor de la langue française (1971ff). Paris: Éditions du Centre national de la recherche scientifique.

Wurzel, Wolfgang Ulrich (1988): Derivation, Flexion und Blockierung. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41, 2, 179-198.

Zaenen, Annie (1993): Unaccusativity in Dutch: integrating syntax and lexical semantics. In James Pustejovsky (ed.) *Semantics and the Lexicon*. Dordrecht: Kluwer Academic Publishers, 129-161.

Zwanenburg, Wiecher (1984): Word formation and meaning. *Quaderni di semantica,* V / 1, 130-142.

The Polish passive and impersonal in Lexical Mapping Theory*

Anna Kibort
University of Cambridge
ak243@cus.cam.ac.uk

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

1. Introduction: the Polish passive and impersonal

This paper offers an analysis of two constructions in Polish which display systematic alternations thought to occur in the argument structure of the predicate. The first one is a periphrastic passive construction, as in:

(1)
Piotr *był* *bity* *(przez kaprala).*
Peter.MASC.NOM was.MASC beat.PART.MASC.NOM (by corporal)
'Peter was beaten (by the corporal).'

Like other Slavonic languages, Polish has both the personal variety of the periphrastic passive, as in (1), and the impersonal variety of this construction, as in:

(2)
Tutaj *było* *tańczone.*
here was.3SG.NEUT dance.PART.SG.NEUT
'There was dancing here./The dancing was done here.'

Impersonal passives of (unergative) intransitives, like the one in (2), are, however, not unique to Slavonic. Analogous constructions occur in many non-Slavonic languages, including well-researched languages such as German or Latin.

The other Polish construction under consideration is a distinctive impersonal construction ending in *-no/-to*, as in:

(3)
Tutaj *tańczono.*
here dance.PART(?).SG(?).NEUT(?)
'There was dancing here./The dancing was done here.'

The *-no/-to* construction differs from the passive construction in two important respects. First, it applies to a wider class of verbs than the passive. In particular, it unproblematically applies to canonical unaccusative verbs, such as 'remain', 'die', or even 'be' (which will be illustrated in the following section), for example:

(4)
Umierano *z* *wycieńczenia.*
die.PART(?).SG(?).NEUT(?) from exhaustion
'[Those people] died/used to die of exhaustion.'

Second, it does not induce promotion of a structural accusative. In the *-no/-to* construction the structural object of a transitive verb does not advance, and in the resulting sentence a structural accusative occurs without a structural nominative. This can be illustrated by:

(5)
Bito *Piotra.*
beat.PART(?).SG(?).NEUT(?) Peter.MASC.ACC
'Peter was beaten.'

These properties clearly distinguish the *-no/-to* construction from the passive in other languages, and, more importantly, from the passive in Polish exemplified in (1). The *-no/-to* construction thus provides two basic theoretical challenges: how to distinguish this construction from the passive, and how precisely to analyse it.

Since the *-no/-to* construction contains what looks like a passive participle with a SG NEUT ending, it has often been assumed in generative linguistic tradition that it is a (syntactic) variant of the canonical passive – namely, that it is an impersonal passive (e.g. Comrie 1977, Borsley 1988, Franks 1995). Such accounts usually do not take into consideration the existence in Polish, alongside the *-no/-to* construction, of 'true' impersonal passives of intransitives, such as the one in (2), which contain SG NEUT participles.

The misclassification of the *-no/-to* construction as passive leads to theoretical solutions which are forced to compromise some correct descriptive generalizations for which there is extensive evidence. Since the *-no/-to* construction does not feature the advancement of the object to subject position, 'passive plus accusative' constructions have to be allowed into the theory, sacrificing the putatively universal generalization that a passive construction cannot retain a structural accusative object. Also, since the *-no/-to* construction applies unproblematically to typical unaccusative predicates, the theory is forced to allow 'passives of unaccusatives' thus compromising the generalization that passives cannot be formed of 'initially unaccusative' verbs (Perlmutter 1978). Worse yet, analysing the *-no/-to* construction as passive renders theories unable to offer a principled account of the variation within the construction, since there is no principled way to account for the fact that some languages allow 'passives' of unaccusatives alongside passives of unergatives while others strictly prohibit them. It is not even clear that formal accounts have any straightforward way of describing this typological difference. Whether or not to analyse the *-no/-to* construction as distinct from the passive is, therefore, not an argument between two accounts with differing constraints. It is, rather, an argument between an account which provides a coherent notion of the passive, and an account which offers an extended notion of the passive and no principled description of the variation within the construction.

To sum up, the theoretical challenge posed by the *-no/-to* construction is not just to account for its properties mentioned above, but also to explicate its relation to the passive, and to reappraise the analysis of the passive, including the impersonal passive of intransitive, in such a way as to preserve previous, robust generalizations.

It can be argued that the problems posed by the *-no/-to* construction derive from a more general challenge posed by impersonal constructions. No current theory offers an account of impersonals and, as suggested by Blevins (2001:1), this might be due to the fact that most formal frameworks explicitly exclude the possibility of subjectless constructions. This is achieved through subject-legislating constraints such as the 'Extended Projection Principle' of Government and Binding or Minimalism, the 'Final 1 Law' of Relational Grammar, or the 'Subject Condition' of Lexical-Functional Grammar. The challenge is, first of all, to provide a mechanism which would allow a class of constructions which do not, and cannot, have a subject constituent. Moreover, the mechanism should distinguish those constructions which are impersonal and at the same time syntactically subjectless (such as impersonal passives of intransitives) from those which are impersonal but do have a syntactically active covert subject (such as the *-no/-to* impersonal). It is worth pointing out here that, in spite of often serving a similar communicative purpose to personal and impersonal passives, the latter have almost always been considered non-passive in traditional descriptive or specialist literature (c.f., recently, Tommola 1998 on the 'suppressive ambipersonal' in Finno-Ugric).

In Polish descriptive linguistic tradition, for example, the *-no/-to* impersonal has rather unanimously been regarded as '*active* indefinite' (e.g. Wierzbicka 1966, Siewierska 1988). In this paper, I will first look at the distribution of the *-no/-to* verb form and the morphosyntactic properties of the construction and I will present the evidence demonstrating that the Polish *-no/-to* impersonal is indeed not passive. Most importantly, unlike the real impersonal passive, the *-no/-to* construction does not lack the syntactic subject. I will then discuss the model of argument structure underlying the analysis which I propose. Finally, I will offer an analysis of both the passive and the impersonal. Following broadly the current lexicalist analyses, I will argue that the passive is an instance of an alternative (non-default) mapping of grammatical functions of the predicate which does not alter its meaning (i.e. lexical semantics). The *-no/-to* impersonal, on the other hand, results from an operation which 'blocks', or 'holds up' the subject position without affecting either the semantics of the predicate or the assignment of the grammatical functions. Passivization can, therefore, be seen as a function-changing process which demotes an 'inital' subject to an optional oblique, while *-no/-to* impersonalization is a function-preserving process that suppresses the realization of a 'final' subject in a similar way to the one assumed in analyses of analogous constructions in Uralic (cf. Blevins 2001:2).

2. Morphosyntactic properties of the *-no/-to* impersonal[1]

In contrast with the passive, the Polish *-no/-to* construction can never co-occur with an expression of the passive agent. Moreover, contrary to what it may look like superficially, the *-no/-to* form itself does not belong to the inflectional paradigm of modern Polish participles. It is not equivalent to the passive participle marked for SG NEUT (*-ne/-te*) which is used in impersonal passives of intransitives, and it is never used in any contemporary Polish passives. It was historically the nominal neuter passive participle, but it has lost its neuter meaning, acquired an impersonal/indefinite one, and it is now used exclusively in the impersonal/indefinite meaning (Siewierska 1988:270) without any of the passive auxiliaries ('to be', 'to become', etc.). Contemporary Polish passive participles, which require an auxiliary, evolved from the adjectival rather than the nominal declensional system (Dziwirek 1994:184).

The strongest piece of evidence supporting the non-passive status of the *-no/-to* construction is that – being independent from the passive – it can, in fact, interact with it. *-No/-to* impersonalization can be applied to a passivized predicate, as long as the predicate contains a syntactic subject (which, in this case, is a derived 'final' subject) whose surface realization it can suppress. This instance of *-no/-to* impersonalization is, additionally, an example of its applicability to the canonical unaccusative verb 'be':

(6)

Bywano	*bitymi.*
was[HABITUAL].IMPERS	beat.PART.PL.INSTR

'One would be beaten.'

In the remaining part of this section I will summarize the evidence demonstrating that, in contrast with the subjectless periphrastic passive of intransitive, the *-no/-to* impersonal appears to have a syntactically active 'covert' subject which participates in syntactic control and raising, as well as binds reflexive and reflexive possessive pronouns. Sentences exemplifying some of these syntactic properties of the *-no/-to* construction have appeared in Polish linguistic literature since the late 1970s (Neubauer 1979; Dyła 1983; Siewierska 1988; Dziwirek 1994; more recently e.g. Bondaruk & Charzyńska-Wójcik 2000), though they have never so far led to an analysis capturing both the distinctiveness of the *-no/-to* impersonal, and its relation to the passive.

The following are examples of the syntactic phenomena in question. Sentences (7)-(9) illustrate the fact that the subject of infinitival complement clauses and 'participial' (temporal) adjunct clauses is unproblematically omitted under coreference with the non-overt subject of the *-no/-to* construction:

(7)

Chciano	*wyjechać.*
wanted.IMPERS	leave.INF

'There was eagerness to leave.'

(8)

Wychodząc	*z*	*budynku*	*zauważono*	*napisy*	*na ścianach.*
leave.PART$_{\text{CONTEMP}}$	from	building	noticed.IMPERS	inscriptions.NONVIR.ACC	on walls

'On leaving the building they noticed the/some inscriptions on the walls.'

(9)

Zakończywszy	*posiłek*	*rozpoczęto*	*dyskusję.*
finish.PART$_{\text{ANTERIOR}}$	meal.MASC.ACC	began.IMPERS	discussion.FEM.ACC

'Having finished the meal, they began the discussion.'

---

* I gratefully acknowledge the continued financial support of Pembroke College, Cambridge, which has made this research possible.

[1] A detailed discussion comparing the passive and the *-no/-to* impersonal will appear in my paper (in preparation) based on a conference presentation (2000). Parts of the present brief summary of the argumentation which will be provided in that paper have appeared in (2001).

When the non-overt subject of an embedded impersonal clause is raised to the subject position in the main clause, the raising verb (*zdawać się* 'seem') turns up in the impersonal, as in:

(10)
Zdawano się tego nie dostrzegać.
seemed.IMPERS REFL this.GEN NEG notice.INF
'[Those people] seemed not to notice this.'

And, finally, the *-no/-to* construction allows the use of reflexive and reflexive possessive pronouns in cases where they require to be bound by the subject, as in:

(11)
Oglądano się/siebie w lustrze.
looked-at.IMPERS self[REFL].ACC in mirror
'One looked at oneself in the mirror./They looked at themselves in the mirror.'

(12)
Oglądano swoje zbiory.
looked-at.IMPERS own[REFL].ACC collections.NONVIR.ACC
'One looked at one's collection./They looked at their collection.'

Since the *-no/-to* impersonal may overlap with the passive in its communicative effect, it may be considered 'passive in meaning or use', and for this reason it has often been classified as passive in functional accounts (cf. Blevins 2001:5). It is, however, clear, that it is not 'passive in form'. Formally, it "pattern[s] with synthetic verb forms that incorporate a subject argument, except that the suppressed subject of an impersonal receives an indefinite interpretation" (ibid.).

3. Modelling the mapping of arguments to syntactic functions

The core ideas expressed in the two previous sections can be summarized in the following pre-theoretical hypotheses about the passive and the *-no/-to* impersonal. Both constructions are, morphologically, derivational. They result from operations on lexical argument structures of predicates which affect the subject. The passive is an instance of an alternative (non-default) mapping of grammatical functions of the predicate, as a result of which the 'initial' subject is demoted to an optional oblique. It is, therefore, a function-changing operation. The *-no/-to* impersonal suppresses the realization of the 'final' subject. The 'blocking' of the subject position prevents the object from being 'promoted'. *-No/-to* impersonalization is, therefore, a function-preserving operation, i.e. it does not alter the default mapping of the grammatical functions. Neither passivization nor *-no/-to* impersonalization seems to affect the lexical semantics of the predicate: both operations are, therefore, meaning-preserving (cf. Sadler & Spencer 1998, Ackerman & Moore 2001).

3.1. LFG's model of argument structure

In order to capture the above generalizations, we need a model of argument structure which contains minimally the following components: (a) a semantic tier representing the hierarchically organized participants in the event designated by the predicator; (b) a syntactic level that identifies the semantic participants as syntactic dependents of the predicate; and (c) principles of syntactic argument classification that anticipates, or leads to, the assignment of grammatical functions.

The model of argument structure provided by LFG contains all three components enumerated above. It thus provides a useful basis for describing derivational process in lexical argument structures. The outline of the relevant parts of LFG's Lexical Mapping Theory (LMT) given in this and in the following section is based primarily on Bresnan & Zaenen (1990), Zaenen and Engdahl (1994), and Bresnan (2000).

At the semantic level of argument structure, LMT identifies the arguments (θ) of the predicate by their semantic (or, thematic) roles and orders them to a presumably universal hierarchy.

At the syntactic level, LMT provides a partially specified syntactic classification of the arguments via the features [+/– (thematically) restricted] and [+/– objective]. The arguments are associated with the syntactic classification according to the underlying lexical semantics of their thematic roles. The basic principles for determining the choice of syntactic features are as follows:

(13)

patientlike roles	θ [– r]	
secondary patientlike roles	θ [+ o]	
other semantic roles	θ [– o]	

The features [+/– r] and [+/– o] constrain the way in which the arguments are mapped onto grammatical functions and group grammatical functions into natural classes:

(14)

	[– o]	[+ o]
[– r]	SUBJ	OBJ
[+ r]	OBL_θ	OBJ_θ

(where OBL_θ abbreviates multiple oblique functions, and OBJ_θ abbreviates secondary objects). In this way, the principles in (13) enable the mapping from the semantic to the syntactic level for any predicate.

As an example, the 'intrinsic' syntactic classification of arguments in a transitive (i.e. two-place) predicate such as *bić* 'beat' can, therefore, be schematized as in:

(15)

$$\begin{array}{cc} \langle\ x & y\ \rangle \\ [-\,o] & [-\,r] \end{array}$$

where *x* and *y* represent the semantic roles of the participants of the event designated by the predicator *bić* 'beat', which are taken here to be an agent and a patient, respectively. The ordering of the two participants reflects the prominence ranking determined by the thematic hierarchy, where an agent is more prominent than a patient.

3.2. Default mappings of arguments to syntactic functions

The argument structure exemplified in (15) mediates between lexical semantics and surface syntactic structure, as it already contains sufficient syntactic information to enable the mapping of arguments to surface grammatical functions.

The mapping of syntactically pre-classified arguments to grammatical functions obeys the so-called 'Function-Argument Bi-uniqueness' condition (Bresnan 1980; 2000), which requires that each argument in the argument structure is associated with a unique grammatical function, and each grammatical function corresponds to a unique argument. Given that, multiple restricted objects and obliques are, nevertheless, possible because these functions are further individuated by their semantic roles (ibid. 2000:311).

Since the negatively specified features in the diagram in (14) indicate unmarked feature values, the diagram can be read as a markedness hierarchy of grammatical functions, with SUBJ being the least marked syntactic function, and the restricted object (OBJ_θ) being the most marked function. In LMT, the property of markedness of grammatical functions plays a role in determining the mapping of arguments to functions, and the principles according to which the syntactic realizations of the arguments are derived are formulated as follows:

(16) Mapping Principles:

a. Subject roles:

(i) a [– o] argument is mapped onto SUBJ when initial in the argument structure; otherwise:

(ii) a [– r] argument is mapped onto SUBJ.

b. Other roles are mapped onto the lowest (i.e. most marked) compatible function on the markedness hierarchy.

Taking the above as the point of departure, I would like to suggest a reformulation of the above principles to make full use of the markedness hierarchy:

(17) Mapping Principle:

The ordered arguments are mapped onto the highest (i.e. least marked) compatible function on the markedness hierarchy.

This alternative formulation achieves the same mappings for various classes of predicates discussed in the literature (including ditransitives and unaccusatives, for example), but avoids stipulating specific principles where their result is already partially determined by the markedness hierarchy. In this way, it avoids redundancy both in the account of the mapping itself, as well as in the formulation of any conditions or constraints pertaining to the subject. Since it makes redundant the Subject Condition ('Every predicator must have a subject'), it might be theoretically helpful in view of constructions which have posed problems of analysis due to, among other things, their non-standard behaviour with respect to the subject (e.g. impersonal constructions with 'genitive subjects'). It may also enable a reappraisal of causatives (with their multiple agentive arguments) as well as any other types of constructions in which the change in the predicate's meaning (lexical semantics) alters the semantic classification of the predicate's roles, which in turn brings about alterations in the syntactic realization of the arguments.

I argue, therefore, that it is the markedness hierarchy which determines the default mapping of the arguments to surface grammatical functions. The following is an example of the application of the Mapping Principle to the predicate *bić* 'beat' given earlier in (15):

(18)

⟨ *x*	*y* ⟩
[– o]	[– r]
|	|
SUBJ	OBJ

4. Alternative ('re-aligned') mappings of arguments to syntactic functions

The fact that arguments of the predicate receive only partial syntactic classification – that is, they are in fact underspecified for grammatical function – invites a possibility of alternative mappings of the same arguments to syntactic functions. Most importantly, since a [– r] role can be mapped onto an object or a subject, in certain contexts (such as the passive or locative inversion) the grammatical function of the subject can be mapped onto an argument corresponding to a lower role on the thematic hierarchy than it would be otherwise. Another possibility of an alternative mapping is provided by the [– o] feature which allows an argument to be mapped onto a subject or an oblique.

These possibilities have been used in the analyses of passivization and locative inversion – two types of operation occurring in the argument structure which result in systematic alternations in the assignment of grammatical functions and which do not affect the lexical semantics of the predicate. Because of these properties, the operations have been referred to as morphosyntactic (e.g. Ackerman 1990; Sadler & Spencer 1998; Ackerman & Moore 2001), and argued to be "motivated by discourse considerations, in which grammars provide speakers with the means to take different perspectives on truth-functionally equivalent situations" (Ackerman & Moore 2001:3).

4.1. Passivization

Pairs of active and passive predicates are standardly assumed not to differ with respect to their lexical semantics, though their participants display alternative assignments of grammatical functions. Cross-linguistically, in an active transitive sentence the agent nominal is a subject, while the patient or theme nominal is a direct object. In its passive counterpart, however, the patient nominal bears the subject function, while the agent nominal, if it is syntactically expressed, has the grammatical status of an oblique. The fact that passivization involves a change in the mapping of arguments to syntax is now uncontroversial in lexicalist accounts, and – as indicated above – it is explained in LFG by resorting to the syntactic underspecification of the arguments.

Diagram (19) is a repetition of (18) and it represents the default – understood as ‘active’ in the context of the present discussion – assignment of final grammatical functions in the predicate *bić* ‘beat’:

(19)

$$\begin{array}{ll} \langle\ x & y\ \rangle \\ {[-\mathrm{o}]} & {[-\mathrm{r}]} \\ | & | \\ \mathrm{SUBJ} & \mathrm{OBJ} \end{array}$$

while the following diagram (20) represents the alternative assignment of final grammatical functions in the same predicate after passivization:

(20)

$$\begin{array}{lll} \langle\ x & y\ \rangle & \qquad\qquad \textbf{passive} \\ {[-\mathrm{o}]} & {[-\mathrm{r}]} & \\ | & | & \\ \mathrm{OBL} & \mathrm{SUBJ} & \end{array}$$

The diagram in (20) represents only the final result of the application of the passive rule to a predicate – that is, the fact that the arguments of the predicate have received an alternative assignment of grammatical functions, which is possible due to the syntactic underspecification of the arguments. To account for the process which has produced this result, it is possible to put forward two alternative hypotheses for the primary operation at work in passivization. The two options – ‘demotion’ of the logical subject versus ‘promotion’ of the logical object – have been discussed extensively in theoretical, functional and descriptive literature. LFG’s analysis of the passive is essentially a ‘demotional’ account, and it seems to be confirmed as correct by the existence of impersonal passives of intransitives in many languages including Polish (cf. example (2) in the introductory section).

Using only the principles of LMT outlined above, I suggest that the mechanism which is involved in the process of assigning alternative grammatical functions in the passive is the ‘demotion’, or ‘downgrading’, of the highest (underspecified) argument by specifying that it must map onto a ‘restricted’ ([+ r]) grammatical function characteristic of obliques. The remaining argument (patient/theme) is then mapped onto its final function (subject) according to the Mapping Principle:

(21)

$$\begin{array}{lll} \langle\ x & y\ \rangle & \qquad\qquad \textbf{passive} \\ {[-\mathrm{o}]} & {[-\mathrm{r}]} & \\ \mathbf{[+\,r]} & | & \\ | & | & \\ \mathrm{OBL} & \mathrm{SUBJ} & \end{array}$$

and the resulting construction is a personal passive which was illustrated earlier with the Polish example in (1). In the case of an intransitive (i.e. one-place) predicate, where there is no argument present which could be ‘promoted’ to subject position, the operation will result in an impersonal passive which can be represented as:

(22)

⟨ *x* ⟩ **passive**
[– o]
[+ r]
|
OBL

An oblique grammatical function is not obligatorily expressed in the syntax, and it has been observed that most passives, whether personal or impersonal, occur without the oblique agent (cf. Keenan 1985). Furthermore, in some languages the expression of the passive agent does not seem to be as easily acceptable in passives of intransitives, as in passives of transitives. This contrast occurs in Polish as well, where impersonal passives are usually preferred to be agentless, as in example (2) repeated below as (23):

(23)
Tutaj było tańczone (?/ przez uczniów).*
here was.3SG.NEUT dance.PART.SG.NEUT (?/* by students)
'The dancing was done here (?/* by students).'

though the syntactic presence of the passive agent is not always excluded in this type of construction:

(24)
Dzisiaj było już sprzątane – przez sprzątaczki.
today was.3SG.NEUT already clean.PART.SG.NEUT by cleaners
'The cleaning has already been done today – by cleaners.'

In standard LFG accounts of the passive, the mechanism proposed for the re-alignment of the 'demoted' agent participant is, briefly, as follows. The agent argument becomes 'suppressed' and thus prevented from receiving any further syntactic specifications as well as, most importantly, from being mapped onto a syntactic argument. It is allowed to be linked to an 'argument adjunct' such as the *by*-phrase in English, but – since the argument adjunct is "not strictly speaking the expression of the same role as the subject in an active sentence" (Zaenen & Engdahl 1994:193) – the highest [– o] argument is in fact assumed to be syntactically unexpressed. The linking between the original agent argument and the new argument adjunct is established through coindexing the two arguments (or, roles). However, this creates a problem of there being two arguments in the argument structure corresponding to (or, bearing) the same thematic role, although passivization clearly does not introduce an alteration in the semantics of the predicate which could be understood as an addition of an argument or role.

Alternatively, as suggested by Alsina (1996:54-56) for English, the linking between the original agent argument and the new argument adjunct is established through the lexical entry of the preposition *by* which "specifies some information about the c[onstituent]-structure realization of an argument, and, so it can access a suppressed argument". As a result, two types of obliques need to be identified, depending on the semantic participant they express: those that map onto arguments (such as the locative) and those that map onto adjuncts (such as the passive agent). In contrast with arguments, adjuncts are not represented at the a[rgument]-structure level – and this seems to be a solution to the problem posed by the standard LFG analysis of the passive agent outlined in the previous paragraph. This alternative analysis is based on the assumption – which I will discuss, and argue for, in more detail in section 7 of this paper – that argument structure is a level of representation that is derived from the lexical semantic representation, but it is in fact distinct, and separate, from the semantic level. However, it requires the presence of an argument-taking preposition which would license the syntactic realization of an argument. In view of the fact that in some languages (such as Russian, or some other Slavonic), the oblique agent is expressed simply in a nominal marked for an oblique case (which, in Russian, is the instrumental), I suggest that the analysis which should be favoured remains the one outlined earlier in (21) and (22).

4.2. Locative inversion

The revisions in my analysis of the passive preserve the standard LFG analysis of locative inversion. Locative inversion, discussed extensively in Bresnan & Kanerva (1989) and Bresnan (1994) and referred to in Ackerman & Moore (2001), can be exemplified by the following pairs of sentences in English (25) and Chicheŵa (26):

(25)
a. *Those visitors came to the village.*
b. *To the village came those visitors.*
(Ackerman & Moore 2001, 2a and b)

(26)
a. *Alendôwo* *anabwérá* *kumudzi.*
2-visitor-2 those 2 SB-REC PST-come-IND 17-3-village
'Those visitors came to the village.'
b. *Kumudzi* *kunabwérá* *alendôwo.*
17-3-village 17 SB-REC PST-come-IND 2-visitor-2 those
'To the village came those visitors.'
(Bresnan & Kanerva 1989, 2b and 1b)

According to the principles of syntactic classification of arguments, the argument structure of the predicate 'come' in both languages can be represented as:

(27)

$\langle$ *x*	*y* $\rangle$
[– r]	[– o]

where the [– r] classification of the first argument indicates that its semantic role is more patientlike than agentlike (and that the verb patterns with other unaccusative predicates – a point which will be taken up in some more detail in section 7 of this paper). The default, i.e. 'uninverted', assignment of grammatical functions in (a) sentences can be derived as follows:

(28)

$\langle$ *x*	*y* $\rangle$
[– r]	[– o]
|	|
SUBJ	OBL

On account of their morphological, syntactic and phonological properties, the nominals denoting the 'visitors' in both of the (b) sentences are identified as objects, while the nominals denoting the 'village' are subjects. This alternative, i.e. 'inverted', assignment of grammatical functions in the (b) sentences, which has been argued to occur in the context of presentational focus, can be modelled by LMT in the following way:

(29)

$\langle$ *x*	*y* $\rangle$	**locative inversion**
[– r]	[– o]	
|	|	
OBJ	SUBJ	

As in the case of passivization, the diagram in (29) represents only the final result of the application of locative inversion to a predicate, and this part of the analysis of locative inversion does no longer seem controversial – at least from a lexicalist perspective. However, similarly to the debate which was carried out a decade or two ago about the passive – namely, whether the primary operation of the passive is the ‘advancement’ of the logical object to subject, or the ‘demotion’ of the subject – it is possible to put forward two, empirically verifiable, hypotheses regarding the primary operation which is at work in locative inversion. A detailed discussion and the assessment of these hypotheses are beyond the scope of this paper. However, I would like to outline them very briefly and highlight those aspects of the preferred solution which will, unsurprisingly, show parallels with passivization and help bring out distinctive characteristics of these types of operation in contrast with the operation of impersonalization.

It was suggested by Bresnan & Kanerva (ibid.:26-28) that the mechanism which is involved in the process of assigning alternative grammatical functions in locative inversion is a special case of ‘subject default’: the first argument of a verb will be subject, unless special conditions (such as e.g. the first argument being [– r]) enable a lower [– o] argument to become subject by optionally characterizing it as thematically unrestricted ([– r]). When the location argument is a subject, the first (theme) argument must, then, be mapped onto object:

(30)

$$\begin{array}{ll} \langle\; x & y\; \rangle \\ [-\mathrm{r}] & [-\mathrm{o}] \\ | & \mathbf{[-r]} \\ | & | \\ \mathrm{OBJ} & \mathrm{SUBJ} \end{array}$$

locative inversion

Although technically possible, the solution seems counterintuitive in that it imposes a ‘thematically unrestricted’ classification on an argument which had received an intrinsic syntactic classification of [– o] precisely because of its specific, non-core, thematic status in the argument structure. Although it is the second argument in the argument structure (which is often the position of the logical object), it did not qualify for the [– r] classification because it was associated with a specific, locative, semantic role. The suggested solution assumes that the primary operation in the locative inversion is the ‘promotion’, or ‘advancement’, of a locative argument to subject. It seems to be motivated by the ‘Subject Condition’ which requires every predicate to have a subject and organizes the mapping from arguments to grammatical functions SUBJ-centrically (cf. (16) versus the alternative proposal in (17)).

The alternative hypothesis views locative inversion not as triggered by ‘promotion’ of a lower argument to subject, but – similar to the passive – as ‘demotion’ of the highest argument to a lower grammatical function. In the old ‘re-mapping’ terminology this could be expressed as ‘demotion of subject to object’. In current LMT terms, this can be achieved by specifying that the highest (underspecified) argument must be ‘objective’ ([+ o]). According to the Mapping Principle, the remaining (underspecified) argument will be mapped onto subject:

(31)

$$\begin{array}{ll} \langle\; x & y\; \rangle \\ [-\mathrm{r}] & [-\mathrm{o}] \\ \mathbf{[+o]} & | \\ | & | \\ \mathrm{OBJ} & \mathrm{SUBJ} \end{array}$$

locative inversion

Apart from the fact that restricting, rather than un-restricting, of the mapping seems to be a theoretically more plausible analysis of a process resulting in a marked, non-default, construction, the solution in (31) seems to be corroborated by some empirical evidence. Namely, the hypothesis in (31) predicts that locative inversion may be found with predicates which subcategorize for only one argument, since – like in the passive – ‘demotion’ of an argument involves a concomitant ‘promotion’ of another (lower) argument only if there is something to be ‘promoted’. This prediction seems to be confirmed by possible pairs of sentences such as:

(32)
a. *And then, those visitors came.*
b. *And then - came those visitors.*

where the locative inversion in (32b) can be accounted for as follows:

(33)

$\langle\ x\ \rangle$ **locative inversion**
[– r]
[+ o]
|
OBJ

4.3. Complementarity of passivization and locative inversion

The representations of both locative inversion and passivization which were suggested in the previous two parts of this section follow straightforwardly from the assumptions and principles of LFG's Lexical Mapping Theory and are in accordance with monotonicity. The proposed analyses seem to account for the distinctive properties of the two resulting constructions and, when considered together, they emerge as complementary processes which are part of a larger system of operations occurring in the argument structure of predicates.

Specifically, it has been observed that there are cross-linguistic restrictions on the applicability of both passivization and locative inversion which are based on the distinction between unergative and unaccusative predicates. The operation of passivization applies only to unergative predicates – that is, those predicates whose most prominent argument is their underlying, or 'initial', subject (Perlmutter 1978). In the terminology of LMT, passivization is restricted to predicates whose first argument is classified as non-objective ([– o]), as in (15). The analysis suggested in this paper correctly predicts that the passive rule cannot be applied to unaccusative predicates – that is, those whose initial argument is classified as [– r], as in (27) – because imposing a restricted marking ([+ r]) on an unrestricted argument would violate the principle of monotonicity.

On the other hand, locative inversion has been demonstrated to apply only to unaccusative predicates (Bresnan & Kanerva 1989). Since, according to the analysis offered in this paper, the locative inversion rule specifies that the highest argument of the predicate must be 'objective' ([+ o]), the rule cannot be applied to unergative predicates, as this would also violate the principle of monotonicity. The rule can, therefore, apply only to predicates in which the highest argument is classified as unrestricted ([– r]).

Essentially, both operations seem to serve the same purpose: they both target the highest argument of the predicate in order to degrade it to a lower grammatical function. The resulting alternative mapping of grammatical functions provides a means to take a different perspective on truth-functionally equivalent situations (Ackerman & Moore 2001:3) which is motivated by discourse considerations such as the choice of syntactic pivot and presentational focus. The difference between passivization and locative inversion lies in the scope of their operation, since they apply, respectively, to two complementary classes of predicates: unergatives and unaccusatives.

Analysing both passivization and locative inversion as 'demoting' rather than 'promoting' operations allows us, therefore, to see them as part of a system – which seems to be confirmed by the restrictions of their applicability. Since, on this account, the 'promotion' of an argument is only opportunistic in both of these operations, another advantage of this analysis is that it uniformly accounts for their applicability to both two-place and one-place predicates. The fact that no language seems to contain a passive strategy that solely defines impersonal passives (Blevins 2001:3), as well as – possibly – there is no language in which the only variant of 'locative inversion' is the one which does not contain a locative nominal, seems to indicate that the second, 'promoted' argument is somehow essential to these operations. This, however, can probably be explained by resorting to the discourse function of these operations: they may have arisen out of the need to shift the presentational focus or syntactic pivot onto an argument other than the highest one, in a situation where there is more than one argument present in the unaltered argument structure of the predicate.

5. Suppression of arguments

Unlike operations such as passivization and locative inversion which are function-changing operations, the Polish *-no/-to* impersonal does not involve an alteration in syntactic function assignment. In a transitive predicate, the underlying object is a surface object in a *-no/-to* impersonal, and – as was demonstrated in section 2 – the construction appears to have a syntactically active 'covert' subject which corresponds to its implied logical subject. It seems, therefore, that the *-no/-to* construction results from 'blocking' of the subject position – that is, 'holding up' the subject and preventing it from being mapped onto a categorial argument.

In LMT, a role, or argument, which cannot be mapped onto a categorial argument, is said to be 'suppressed'. The most straightforward formulation of a suppressing rule is: "Do not map an argument to the syntax" (e.g. Falk 2001: 111), and it is notated:

(34)

$$\begin{array}{c} \theta \\ | \\ \varnothing \end{array} \qquad\qquad \text{suppression}$$

If it is understood as a rule suppressing not the argument itself – in which case it would affect the lexical semantics of the predicate – but the argument's syntactic function, then suppression seems to be just the way to account for what occurs in impersonalization. The *-no/-to* impersonal has a covert subject, but it is a subject which has been blocked by suppression. Suppression prevents the subject from being mapped onto a categorial argument and in this construction it is associated with dedicated impersonal morphology.

Since impersonalization operates outside the syntactic classification of the arguments, it should not be sensitive to the distinction between unaccusative and unergative predicates. This is indeed the case. In contrast with passivization which is restricted to unergatives, *-no/-to* impersonalization can be applied to both unergative and unaccusative predicates (including the verb 'to be' used with passive participles, as illustrated by example (6)), and with very few exceptions it is fully productive.

I suggest, therefore, that *-no/-to* impersonalization is an instance of the rule of suppression which blocks the final subject and prevents it from taking a syntactic argument:

(35)

$$\begin{array}{cccc} \langle & x & y & \rangle \\ & [-\mathrm{o}]/[-\mathrm{r}] & [-\mathrm{r}] & \\ & | & | & \\ & \mathrm{SUBJ} & \mathrm{OBJ} & \\ & | & & \\ & \varnothing & & \end{array} \qquad\qquad \textbf{-\textit{no/-to} impersonal}$$

Similarly to the morphosyntactic operations discussed in the previous sections, *-no/-to* impersonalization can be applied to a predicate regardless of whether there is another argument present in the argument structure or not, as long as the argument structure contains a final subject which the rule targets (cf. examples (3), (5) and (6)).

6. The passive and the impersonal: preliminary conclusions

The aim of this paper has been to discuss two distinct constructions in Polish which result from alterations in the argument structure of the predicate: the passive construction, and the impersonal construction ending in *-no/-to*. The latter one, often misclassified as an impersonal passive, raises important descriptive and theoretical issues.

The fist issue concerns the existence of impersonal constructions and the way of accounting for them. I have demonstrated that, apart from overlapping with the meaning of the passive construction in some contexts, the *-no/-to* impersonal is indeed not passive, and – unlike in the true impersonal passive – its subjectlessness is very superficial. In this respect, the *-no/-to* impersonal patterns with analogous constructions in Finno-Ugric languages in which the primary voice opposition can be demonstrated to be personal/impersonal rather than active/passive. The existence of this type of construction in Balto-Slavic languages means that the approach to subjectlessness in theoretical frameworks needs to be reconsidered even for Indo-European languages.

The other issue, which is a consequence of this finding, is the need to distinguish between two types of operation in the argument structure of the predicate: demotion (as in passivization) and suppression (as in impersonalization). I have demonstrated that, in spite of the initially pessimistic view treating all subject-affecting operations as suppression, Lexical Mapping Theory reserves the resources with which it is possible to capture this distinction. Using the descriptive conventions of LTM, I have offered analyses of passivization and impersonalization which can be extended to other constructions believed to be part of the system of operations on the argument structure, and which preserve all the previous sub-claims about these constructions.

In the remaining part of the paper I will suggest further improvements to these analyses by refining the notion of the 'argument' and by arguing for the restoration of the early LFG distinction within the argument structure between semantic roles and argument positions. However, for reasons of space, the argumentation presented in the remaining sections will be considerably compressed.

7. The independence of argument positions and argument roles

Most of the current LFG analyses of passivization, as well as the outline of the revised analysis of passivization which I presented above, assume a two-tiered representation of argument structure. One of the tiers – the tier of the final classification of the arguments into syntactic functions (SUBJ, OBJ, ...) – is clearly a syntactic one and feeds directly into final syntactic structure. The other tier, however – the tier at which the semantic arguments of the predicate are identified (as *x*, *y*, etc.) and syntactically pre-specified ([– o], [– r], ...) – compresses two different levels of information: the semantic level of thematic roles and the syntactic level of argument positions subcategorized for by the predicate.

7.1. Unaccusativity and passivization

The need to separate semantic information from the syntactic representation of predicate's valency presumably manifests itself most strongly in attempts to account for unaccusativity. In syntactic accounts of unaccusativity, the arguments of unergative and unaccusative predicates can be distinguished by resorting to the notion of their 'underlying' grammatical functions (e.g., respectively, the 'initial' subject and object of Relational Grammar; cf. Perlmutter 1978), or – as in LMT – by recognizing that the highest argument of an unergative predicate is non-objective ([– o]), while the highest argument of an unaccusative predicate is unrestricted ([– r]). LMT's intrinsic syntactic classification makes it possible to refer to arguments independently of their thematic roles and grammatical functions, and captures the generalization that an unaccusative argument is not an object, but at the same time it is, underlyingly, not a subject.

The difficulty in applying this idea in analyses of constructions lies in the fact that, in most current LFG accounts, the arguments of the predicate are, in fact, identified with their thematic roles, even though unaccusativity and operations which are sensitive to it – such as passivization – have been recognized as essentially syntactic phenomena, and the notion of an 'underlying slot which comes first, but which is not a subject' is not easily expressible in thematic terms. In fact, it has been demonstrated that is it impossible to find a common semantic denominator for either the class of syntactically unaccusative, or unergative verbs (e.g. Rosen 1984; Wechsler 1995). Since it is uncontroversial that the two types of predicates display surface syntactic contrasts in a great number of languages, and since passivization and locative inversion occur – as has been argued here – completely at the syntactic level, it might be beneficial to restore the earlier LFG distinction between semantic roles and argument positions. I suggest, therefore, representing the two levels of argument structure as independent (though formally related through mapping):

(36)

$$\begin{array}{ccccl} & x & y & & \text{- thematic roles} \\ & | & | & & \\ \langle & \theta_1 & \theta_2 & \ldots \rangle & \text{- valency/argument positions} \end{array}$$

– with argument positions further mapped onto final grammatical functions. Such a model is in agreement with the scheme which underlies the design of LFG as well as other lexicalist syntactic frameworks (Bresnan 2000:306):

(37)

lexical semantics	
↓	Lexico-semantic projection
a[rgument]-structure	
↓	Lexico-syntactic projection
final syntactic structure	

Once the semantic component is extracted from the syntactic argument structure and posited as a separate level of representation, I would also want to argue, following Alsina (1996:37), that "although arguments are ordered in the a[rgument]-structure according to their thematic role, thematic role information is not represented at a[rgument]-structure" since "if thematic information is represented in the lexical semantic representation of predicates, it would be redundant to replicate this information elsewhere, as in the a[rgument]-structure".

Such a model allows us to posit that the passive rule needs to refer only to the syntactic information about the arguments and that, in fact, thematic information is inaccessible to it. This is confirmed by the fact that passivization is restricted to a syntactically, not semantically, distinct class of predicates. In order to reformulate LFG's passive rule in purely syntactic terms, we only need to state that instead of applying to the semantically most prominent role on the thematic hierarchy, it applies to the 'underlying' subject – that is, only to the unergative argument pre-specified syntactically as [– o][2]. By analogy, the locative inversion rule applies only to the unaccusative argument pre-specified syntactically as [– r].

7.2. Raising verbs

The distinction within the argument structure between semantic roles and argument positions is implicit in current LFG work concerning 'empty' (athematic) argument roles of raising verbs (Zaenen & Engdahl 1994; Bresnan 2000). In the a-structures of the subject-raising verb *seem* and the object-raising verb *believe*, given in (38a) and (38b), respectively:

(38)

a.
$$\begin{array}{cccc} \_\_ & \langle\; x & y \;\rangle \\ [-r] & [-o] & [-o] \end{array}$$

b.
$$\begin{array}{ccc} \langle\; x & y \;\rangle & \_\_ \\ [-o] & [-o] & [-r] \end{array}$$

the athematic arguments are represented outside of the angled brackets, which indicates that they do not belong to the set of semantic participants of the action denoted by the predicate. They nevertheless have a specific position in the argument structure relative to the other hierarchically ordered roles, which gives them greater or lesser priority in the mapping to grammatical functions. Having no semantic content, they receive the inherent syntactic classification of [– r].

It could be argued that, due to the nature of the athematic argument, both of these representations imply the existence of a distinct level of argument positions separate from the semantic level, and that the representations in (38a) and (38b) can be straightforwardly translated to the notation in (39a) and (39b):

(39)

a.
$$\begin{array}{ccc} & x & y \\ & | & | \\ \langle\; \theta & \theta & \theta \;\rangle \\ [-r] & [-o] & [-o] \end{array}$$

b.
$$\begin{array}{ccc} x & y & \\ | & | & \\ \langle\; \theta & \theta & \theta \;\rangle \\ [-o] & [-o] & [-r] \end{array}$$

These representations preserve very clearly the insight that the raising verbs subcategorize for three syntactic argument positions whilst they involve only two semantic participants.

---

[2] As argued by Blevins (2001), this restriction on passivization has not been invalidated by alleged passives of unaccusatives in languages such as Lithuanian or Turkish: the forms, related diachronically to the passive, which occur in 'unaccusative' and 'double' passives in Lithuanian (Timberlake 1982) have an evidential meaning that identifies them as part of the mood, rather than the voice, system of the language.

7.3. The anticausative

It is standardly assumed that in inchoative-causative pairs of verbs in English, the inchoative – that is the intransitive – variant is basic (*The jar broke*), while the causative – transitive – variant is derived by adding an agent to the lexical semantic structure of the intransitive predicate (*I broke the jar*) (though cf. Levin & Rappaport Hovav 1995 for detailed discussion). In the roughly corresponding class of verbs in Polish which participate in the causative alternation, the transitive causative form of the verb is morphologically unmarked, while the intransitive form is accompanied by the reflexive marker *się*, as in:

(40)
Słoik *się* *zbił.*
jar.MASC.NOM REFL broke.3SG.MASC
'The jar broke.'

It is, therefore, plausible to argue that the intransitive verb form is derived from the inherently causative transitive verb, and that the reflexive marker is the morphological exponent indicating the nonexpression of the cause.

Arguing for a causative analysis of certain, externally caused, intransitive verbs in English (including *break*) which participate in the causative alternation, Levin & Rappaport Hovav suggest that the binding of the external cause takes place in the mapping from the lexical semantic representation to argument structure. They argue that while the semantic and syntactic structure of the basic, transitive, variant of the verb *break* can be represented as in (1995:108):

(41)

Lexical semantic representation	[[*x* DO-SOMETHING] CAUSE	[*y* BECOME *BROKEN*]]
Linking rules	↓	↓
Argument structure	*x*	⟨*y*⟩

the semantic and syntactic structure of the intransitive form of *break* can be represented as in (ibid.):

(42)

Lexical semantic representation	[[*x* DO-SOMETHING] CAUSE	[*y* BECOME *BROKEN*]]
	↓	
Lexical binding	∅	
Linking rules		↓
Argument structure		⟨*y*⟩

Binding of a position in the lexical semantic representation prevents the projection of that position into argument structure in a similar way as binding, or suppression, of a position in argument structure prevents that position from being projected onto the syntax.

Using the model of argument structure suggested in this paper, I would like to propose that the anticausative rule in Polish targets and deletes the argument position associated with the most prominent thematic role (the agent/cause). Analysed in this way, the anticausative is a morphosemantic operation which deletes a component of meaning from the argument structure of the predicate (cf. Sadler & Spencer 1998), and can be seen as the inverse of the lexical causative rule which has been suggested for many languages with morphological causatives. It can be represented as follows:

(43)

$$\begin{array}{ccc} x & y & \\ & | & \\ \langle & \theta & \rangle \end{array} \qquad \textbf{anticausative}$$

As argued by Levin & Rappaport Hovav, the fact that in English sentences with the intransitive variant of the verb *break* the external cause cannot license a *by*-phrase or control a purpose clause constitutes evidence supporting the hypothesis that the operation of binding the external cause must take place at a level before argument structure. This observation is true of Polish anticausatives too, although the issue requires a brief clarification which regards the different status of the *by*-phrase as opposed to the various volitional expressions.

The fact that a *by*-phrase is disallowed in both English and Polish sentences with the intransitive *break* or *sink* is unsurprising:

(44)
a. **The ship sank by Bill.* (Roeper 1987:268, 2a)
b. **The window broke by Pat.* (Levin & Rappaport Hovav 1995:109, 65a)
c. **Słoik się zbił przez Piotra.*
jar.MASC.NOM REFL broke.3SG.MASC by Peter
'The jar broke by Peter.' (meaning: 'Peter broke the jar.')[3]

Whether analysed as a basically intransitive, or a derived intransitive verb, the verb's argument structure does not contain an argument which could map onto a passive oblique – there is no argument position which could be re-classified as a passive oblique and realized as a prepositional phrase which normally expresses a passive agent. The licensing of a *by*-phrase is essentially a syntactic phenomenon, and the sentences in (44) are syntactically ill-formed, i.e. ungrammatical.

The control of purpose clauses and volitional adverbials seems to be a different phenomenon in that in both English and Polish, sentences with the intransitive *break* or *sink* which contain purpose clauses or adverbials are not ungrammatical, but implausible in most semantic contexts. They are not syntactically deviant, but are uninterpretable unless the referent of the unaccusative argument is personified:

(45)
a. *#The boat sank to collect the insurance.* (Roeper 1987:268, 3a)
b. *#The window broke to rescue the child.* (Levin & Rappaport Hovav 1995:109, 65b)
c. *#Słoik się zbił celowo.*
jar.MASC.NOM REFL broke.3SG.MASC on purpose
'The jar broke on purpose.'

The fact that there are semantic contexts in which purpose clauses and volitional adverbials in anticausatives are acceptable – that is, that they do not need to be licensed by the original sentient agent – seems to point to the conclusion that what licenses them is not just the presence of a specific semantic participant at the semantic level of representation of the predicate. They seem to require the presence of an argument position which is linked to a semantic participant whose referent is, or can be construed as, an agent. Control of purpose and other adverbial (e.g. temporal) clauses seems, therefore, to be a syntactic phenomenon, sensitive to syntactic differences at the level of argument structure, as much as a semantic phenomenon. Purpose and other adverbial clauses are sanctioned at the syntactic level of argument positions, and interpreted according to the semantics of the role which is linked to the controlling argument position. It can be argued that what sanctions them is the 'logical subject', understood as the first argument present in the argument structure which is accompanied by its original thematic role. In contrast with the logical subject, an 'agent' is identified solely at the semantic level and – as we have seen – its characteristics can be transferred onto a semantic participant which is not normally sentient or volitional.

Neither the unacceptability of the *by*-phrase, nor the control phenomena discussed above, seem, therefore, to exclude the possibility that the original agent role is still present in derived argument structures of intransitive (inchoative) verbs – since, as I have argued, both phenomena can be explained without referring to the agent of the basic transitive verb form. The presence of the original agent role is indeed confirmed by the possibility in Polish of expressing this role overtly in an (arguably) additional oblique argument similar to a secondary object or 'Dative':

[3] The Polish sentence in (44c) is grammatically correct and semantically plausible if understood as: 'The jar broke because of Peter' – that is, for example, because Peter had put it in an unsafe place where it could easily be broken by someone else or by some other external cause.

(46)
Słoik mi się zbił.
jar.MASC.NOM me.DAT REFL broke.3SG.MASC
'The jar broke to me/in my hands.'

(47)
Piotrowi wylała się zupa.
Peter.DAT spilt.3SG.FEM REFL soup.FEM.NOM
'The soup has spilt to Peter.' (meaning: 'Peter has spilt the soup.')

In most contexts there is no doubt about the fact that the dative nominal is to be interpreted as the real agent of the action, as opposed to the nominative argument which, though being a logical subject, is – at best – a 'pseudo-agent'.

I suggest that in order to account for both the presence of the original agent role (which, at least in Polish, is available to be picked up by the syntax) in anticausatives, and at the same time for the absence of the argument position which would normally be linked to the original agent, we need to distinguish in argument structure the level of semantic roles as separate from syntactic argument positions. The separation of the semantic level from the syntactic level of representation allows us to hypothesize in what way the delinked agent role may become re-associated with another argument. One possible solution is illustrated in (48) which represents the sentences in (46) and (47), where *y* is the patient/theme, and *x* is the re-aligned agent[4]:

(48)

$$\begin{array}{cccc} & y & x & \\ & | & | & \\ \langle & \theta & \theta & \rangle \\ & [-\,r] & [+o] & \\ & | & | & \\ & \mathrm{SUBJ} & \mathrm{OBJ}_{\theta} & \end{array}$$

anticausative

7.4. The impersonal reflexive

In a similar way to the anticausative, the analysis of another Polish impersonal construction, the impersonal reflexive, can be argued to require separating the semantic level from the syntactic level in the argument structure. The construction can be exemplified by:

(49)
Tutaj się tańczyło.
here REFL danced.3SG.NEUT
'There was dancing here./The dancing was done here.'

By analogy with the *-no/-to* impersonal, this construction can also be seen as resulting from the suppression of the final subject. It similarly possesses a covert syntactic subject and does not allow it to be expressed by a categorial argument in nominative case. With transitive predicates, it retains the logical object in the accusative case, and it is not restricted to either the unergative or the unaccusative class of predicates. The suppression of the final subject in this construction can be accounted for in the same way as in the *-no/-to* impersonal (cf. (35)):

[4] The scope of this paper does not permit a more detailed discussion of this hypothesis.

(50)

$$\begin{array}{cccc} & x & y & \\ & | & | & \\ \langle & \theta & \theta & \rangle \\ & [-\text{o}]/[-\text{r}] & [-\text{r}] & \\ & | & | & \\ & \text{SUBJ} & \text{OBJ} & \\ & | & & \\ & \varnothing & & \end{array}$$

impersonal

Where the impersonal reflexive differs from the *-no/-to* impersonal is not only in its morphological exponent, but also in that it can, arguably, involve a change in the lexical semantics of the predicate concomitant with the primary operation of suppression of the final subject. In brief, just as the anticausative, the impersonal reflexive may contain an overt syntactic expression of its agent role mapped onto a secondary object, or ‘Dative’:

(51)
Łatwo się kapralowi biło Piotra.
easy REFL corporal.DAT beat.3SG.NEUT Peter.MASC.ACC
‘The corporal found it easy to beat Peter.’

I suggest that the agent role of the predicate which has undergone this type of impersonalization may become available to be picked up by the syntax in a way analogous to that proposed for the anticausative (cf. (48)), where *y* is the patient, and *x* is the re-aligned agent:

(52)

$$\begin{array}{ccccc} & & y & x & \\ & & | & | & \\ \langle & \theta & \theta & \theta & \rangle \\ & [-\text{o}]/[-\text{r}] & [-\text{r}] & [+\text{o}] & \\ & | & | & | & \\ & \text{SUBJ} & \text{OBJ} & \text{OBJ}_{\theta} & \\ & | & & & \\ & \varnothing & & & \end{array}$$

impersonal reflexive

The agent role can be re-associated (with a different argument) because the operation allows it to be freed up from the original, blocked, argument position. Although the syntactic classification of argument positions in an impersonal reflexive remains unaltered, the fact that the agent role is re-aligned, and a new argument position is added to the argument structure to accommodate it, means that the operation has to be regarded as morphosemantic – that is, meaning-altering.

The interpretation and the overt expression of the agent in the impersonal reflexive is, therefore, similar to that of the anticausative discussed in the previous section. The subject argument of the impersonal reflexive without the dative nominal is covert but interpretable (as agent). In the variant of this construction with the dative nominal, the covert subject is retained but its agentive interpretation is transferred onto an additional argument. As with the anticausative, the explanation of this fact requires referring to the semantic level of representation in the argument structure as distinct from the syntactic level of argument positions.

8. A revised analysis of the passive and the impersonal

This concluding section will provide a brief summary of the analysis of the passive and the impersonal which was given in earlier sections, this time taking into account all the revisions which have been suggested so far.

I have demonstrated that the Polish *-no/-to* impersonal is a distinct, non-passive, construction whose subjectlessness is very superficial. Although it does not allow an overt expression of the subject or agent, it contains an interpretable, syntactically active covert subject. When applied to a transitive predicate, the -*no/-to* impersonal retains its logical object marked for accusative case. To account for this type of construction, I have suggested distinguishing between two types of operation in the argument structure of the predicate: suppression (as in impersonalization) versus demotion (as in passivization).

I have argued that passivization targets the 'underlying', or 'initial', subject of the predicate – it is, therefore sensitive to the inherent syntactic classification of the arguments of the predicate. It can be seen as 'chômeurizing' the most prominent, non-objective, argument of the predicate by imposing on it a [+ r] marking. I have also suggested that since the restriction on the application of the passive rule to unergative predicates follows from the principle of monotonicity, it does not, in fact, have to be posited as a separate syntactic constraint.

Impersonalization, on the other hand, targets the 'final' subject of the predicate, preventing it from being realized in the overt syntax. As it operates on arguments which have been specified for their final grammatical functions, it is not sensitive to the unaccusative/unergative distinction, and it unproblematically retains the accusative object. It can be viewed as 'suppressing', or 'blocking' the subject in the way which has so far been suggested in LFG for the passive.

While passivization occurs completely at the syntactic level, altering grammatical function assignment but not the lexical semantic structure of the predicate, impersonalization is a function-preserving operation, though it can, in some constructions, involve an additional, concomitant alteration in the lexical semantics of the predicate – it can, therefore, be meaning-altering.

9. References

Ackerman, F. 1990. "Locative alternation vs locative inversion", *Proceedings of the West Coast Conference on Formal Linguistics* 9: 1-13. CSLI Publications

Ackerman, F. & J. Moore 2001. *Proto-properties and Argument Encoding: A Correspondence Theory of Argument Selection.* CSLI Publications

Alsina, A. 1996. *The Role of Argument Structure in Grammar.* CSLI Publications

Blevins, J. 2001. "Passives and impersonals". Ms. RCEAL, University of Cambridge

Bondaruk, A. & M. Charzyńska-Wójcik 2000. "Is there expletive *pro* in impersonal passives: A reappraisal", in: Stanulewicz, D. (ed.) *PASE Papers in Language Studies.* Gdańsk: Wydawnictwo UG

Borsley, R. 1988. "A note on impersonal passives", *Journal of Linguistics* 24: 483-487

Bresnan, J. 1980. "Polyadicity", in: Hoekstra, T., H. van der Hulst & M. Moortgat (eds) *Lexical Grammar.* Foris. In revised form in: Bresnan, J. 1982. (ed.) *The Mental Representation of Grammatical Relations.* MIT Press

Bresnan, J. 1994. "Locative inversion and the architecture of universal grammar", *Language* 70: 72-131

Bresnan, J. 2000. *Lexical-Functional Syntax*. Blackwell

Bresnan, J. & J.M. Kanerva 1992. "Locative inversion in Chichewa: A case study in factorization in grammar", in: Stowell, T. & E. Wehrli (eds) *Syntax and Semantics 26: Syntax and the Lexicon*. Academic Press

Bresnan, J. & A. Zaenen 1990. "Deep unaccusativity in LFG", in: Dziwirek, K., P. Farrell & E. Mejias-Bikandi (eds) *Grammatical Relations. A Cross-Theoretical Perspective*. CSLI Publications

Comrie, B. 1977. "In defense of spontaneous demotion: The impersonal passive", in: Cole, P. & J.M. Sadock (eds) *Syntax & Semantics 8: Grammatical Relations*. Academic Press

Dyła, S. 1983. "Some further evidence against an impersonal passive analysis of Polish impersonal constructions", *Studia Anglica Posnaniensia* 15: 123-128

Dziwirek, K. 1994. *Polish Subjects*. Garland

Falk, Y. 2001. *Lexical-Functional Grammar.* CSLI

Franks, S. 1995. *Parameters of Slavic Morphosyntax*. OUP

Keenan, E.L. 1985. "Passive in the world's languages", in: Shopen, T. (ed.) *Language Typology and Language Description. Vol.1.* CUP

Kibort, A. 2000. "The Polish *-no/-to* impersonal as an active indefinite: a reappraisal", a paper given at the *Autumn Meeting 2000 of the LAGB*, Durham, UK

Kibort, A. 2001. "Morpholexical rules in Polish revisited", in: Przepiórkowski, A. & P. Bański (eds) *GLiP 2: Proceedings of Generative Linguistics in Poland*. Warsaw: Polish Academy of Sciences

Kibort, A. (in preparation) “The form and function of the Polish *-no/-to* impersonal”
Levin, B. & M. Rappaport Hovav 1995. *Unaccusativity. At the Syntax-Lexical Semantics Interface.* MIT Press
Neubauer, P. 1979. “The score on ‘impersonal passive’: Dutch – 1, Polish – 0”. An unpublished paper presented at the 17th International Conference on Polish-English Contrastive Linguistics in Boszkowo
Perlmutter, D.M. 1978. “Impersonal passives and the Unaccusative Hypothesis”, *BLS* 4: 157-189
Roeper, T. 1987. “Implicit arguments and the head-complement relation”, *Linguistic Inquiry* 18: 267-310
Rosen, C. 1984. “The interface between semantic roles and initial grammatical relations”, in: Perlmutter, D.M. & C. Rosen (eds) *Studies in Relational Grammar 2.* University of Chicago Press
Sadler, L. & A. Spencer 1998. “Morphology and argument structure”, in: Spencer, A. & A.M. Zwicky (eds) *The Handbook of Morphology*. Blackwell
Siewierska, A. 1988. “The passive in Slavic”, in: Shibatani, M. (ed.) *Passive and Voice*. John Benjamins
Timberlake, A. 1982. “The impersonal passive in Lithuanian”, *BLS* 8: 508-524
Tommola, H. 1998. “Zametki k tipologii ambipersonala”, in: Kozinceva, N.A. & A.K. Ogloblin (eds) *Tipologija. Grammatika. Semantika. K 65-letiju Viktora Samuilovicha Xrakovskogo.* Sankt-Petersburg: Nauka
Wechsler, S. 1995. *The Semantic Basis of Argument Structure*. CSLI
Wierzbicka, A. 1966. “Czy istnieja zdania bezpodmiotowe”, *Jezyk Polski* 46(3): 177-196
Zaenen, A. & E. Engdahl 1994. “Descriptive and theoretical syntax in the lexicon”, in: Atkins, B.T.S. & A. Zampolli (eds) *Computational Approaches to the Lexicon*. OUP

“Optimal” Linking for Modern Greek Psych Verb Constructions

Valia Kordoni

University of Saarland

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

1 Introduction

Modern Greek (henceforward MG) distinguishes three classes of Psych Verb Constructions (henceforward PVCs):

1. The Experiencer-Subject Psych Verb Constructions (henceforward ESPVCs). This class of PVCs includes verbs like *miso* (hate), *agapo* (love), or *latrevo* (adore), which feature a nominative *experiencer* in agreement with the verb, and an accusative *theme*:

   (1) O Gianis misi to sholio.
the Gianis.N hate.3S the school.A
"John hates school."

   (2) O Gianis agapa tin Maria.
the Gianis.N love.3S the Maria.A
"John loves Mary."

   (3) O Gianis latrevi tin musiki.
the Gianis.N adore.3S the music.A
"John adores music."

2. The Experiencer-Object Psych Verb Constructions (henceforward EOPVCs), which feature a nominative *theme* in agreement with the verb, and an accusative *experiencer*:

   (4) I Maria eksorgizi ton Giani.
the Maria.N enrage.3S the Giani.A
"Mary enrages John."

   (5) I kategides to fovisan to pedi.
the thunderstorms.N,PL cl.A frighten.PAST.3PL the child.A
"The thunderstorm frightened the child."

3. The last class of PVCs in MG includes the verbs *aresi* (likes) and *ftei* (bothers/matters), which feature a nominative *theme* in agreement with the verb, and an *experiencer*, either in morphological genitive or as the complement of a prepositional phrase:[1]

   (6) To sholio aresi ston Giani.
the school.N like.3S to-the Giani.A
"John likes school."

[1] This class of PVCs in MG is parallel to the so-called *piacere* class of Italian:

(1) A Gianni piace questo.
to Gianni pleases this
"This pleases John."

(2) Questo piace a Gianni.
this pleases to Gianni
"This pleases John."

for which Belletti and Rizzi (1988) have argued that the *experiencer* argument bears a *lexical* dative case marker.

(7) To sholio tu aresi tu Giani.
the school.N cl.G like.3S the Giani.G
"John likes school."

This paper focuses on the semantic properties and the syntactic behaviour of constructions like the following:

(8) I Maria fovate tis kategides.
the Maria.N fear.3S the storms.A
"Mary is afraid of the storms."

(9) I Maria fovate me tis kategides.
the Maria.N fear.3S with the storms
"Mary is afraid of the storms."

(10) I Maria.N fovate ton Giani.
the Maria.N fear.3S the John.A
"Mary is afraid of John."

(11) I Maria.N fovate me ton Giani.
the Maria.N fear.3S with the John.A
"Mary is afraid of John."

These constructions belong to MG ESPVCs (cf., also examples (1)-(3)). Interesting about them is the fact that the "experienced" (henceforward EXPD)[2] semantic role[3] is syntactically realized as the object of the sentence in constructions like (8) and (10), while in constructions like (9) and (11) it is syntactically realized as the object of a prepositional phrase.

We should underline here that examples (8) and (9) convey the same meaning. The same is true for examples (10) and (11). That is, the two variants in each pair do *not* differ semantically, although the first variant of each pair (examples (8) and (10)) is used more often in the language. In other words, in order to express the meaning *Mary is afraid of the storms* native speakers do prefer the construction in (8), rather than the one in (9).[4]

Our aim, thus, here is twofold:

1. First we will account for the semantic and syntactic properties of the predicates' arguments in (8)-(11).

2. Then we will show that the linking of the semantic roles in (8) and (10) to the respective syntactic arguments might be related to the constructions in (8) and (10) ranking higher than the constructions in (9) and (11) in the native speakers' preference.

To do this, we rely on Wechsler's (1995) argument structure theory, as well as the linking architecture proposed by Butt, Dalrymple, and Frank (1997).

[2]We adopt here Markantonatou's (1995) terminology.

[3]I.e., the theme.

[4]Our source is the Greek ECI corpus, which contains a broad mixture of different texts, such as a number of novels (translated mostly from English originals) including 5 cowboy stories by Louis L'amour and Sherlock Holme's Last Bow, a group of 8 files to do with astrology, a group of legal case reports from the Supreme Court of Cyprus, and a number of technical and/or philosophical books or extracts from books.

2 ESPVCs in Modern Greek

2.1 Overview

As shown in Section (1), MG ESPVCs include verbal predicates whose common characteristic is that they feature a nominative *experiencer* in agreement with the verb.

Both the literature on MG PVCs and the literature on PVCs in other languages have paid more attention to Experiencer-Object rather than to Experiencer-Subject predicates, which we are interested in here. This as such would have been unproblematic, if it had not had the consequence that Experiencer-Subject PVCs have either been left unaccounted for, or the accounts provided for them by the different Lexical Semantics and Linking theories are to a great extent stipulative.

In the generative tradition, for instance, Grimshaw's (1990) linking theory, which on the one hand relies heavily on thematic roles, but on the other hand suggests that argument selection is determined by a causal aspectual structure on a separate "tier" from thematic structure, fails to provide a consistent and parsimonious account of Experiencer-Subject PVCs, since it stipulates counter-intuitively that Experiencer-Subject predicates are no different than normal causative verbs of any natural language:

> ...The case of psychological state verbs like *fear* is considerably more delicate. The desired result will follow if their Experiencer qualifies as the aspectually most prominent argument...However, it must be admitted that in this case there is no independent evidence that the aspectual analysis will give this result, so for the present purposes we must simply stipulate it (Grimshaw (1990, pp. 17-18)).

With this Grimshaw acknowledges that the interaction between her thematic and aspectual hierarchies proves to be problematic in the case of ESPVCs. That is, prominence must be stipulated in the case of ESPVCs.

Grimshaw's (1990) account of ESPVCs also leads her to the conclusion that the *experiencer* argument in such constructions, supposedly being the most prominent one on both the thematic and the aspectual hierarchies, qualifies as an EXTERNAL ARGUMENT.[5] According to her, the fact that the Experiencer-Subject predicates have an EXTERNAL ARGUMENT in their a-structure has the consequence that they can be related to passive sentences, since EXTERNAL ARGUMENT status does predict the availability of PASSIVIZATION in her theory. And although this might be true for this kind of construction in English, which is Grimshaw's case study:

(12) John admired the car parked next to his.

(13) The car parked next to his was admired by John.

this claim does not hold for MG ESPVCs (witness the ungrammaticality of (15) and (17)):

(14) I gonis tu agapun ton Giani.
the parents.N,PL his love.3PL the Giani.A
"His parents love John."

(15) *O Gianis agapiete apo tus gonis tu.
the Gianis.N love.PASS.3S by the parents his
"John is loved by his parents."

[5] Grimshaw (1990) defines the notion of EXTERNAL ARGUMENT as the argument that is most prominent on both hierarchies, i.e., the thematic and the aspectual. If the two dimensions do not pick out the same argument as the most prominent, then, in Grimshaw's account, the predicate lacks an external argument. Thus, according to this, ESPVCs do have an external argument in their a-structure, since the *experiencer* is the most prominent argument on the thematic hierarchy, and it is stipulated to be the most prominent argument on the aspectual hierarchy, as well.

(16) I Maria zilepse to spiti ton gitonon.
the Maria.N envy.PAST.3S the house.A the neighbours.G,PL
"Mary envied the neighbours' house."

(17) *To spiti ton gitonon zileftike apo tin Maria.
the house.N the neighbours.G,PL envy.PASS.PAST.3S by the Maria
"The neighbours' house was envied by Mary."

As far as the literature on MG PVCs is concerned, Tsimpli (1989, p. 246) has argued that some of the Experiencer-Subject predicates can be considered to be the passive forms of the Experiencer-Object predicates that we have seen in (4) and (5) in Section (1):

> ...As to experiencer verbs[6] I argue that they contain a single theta-role in their argument structure, which can be assumed to be either external or internal. *Passives of experiencer verbs are passives of causativised forms of the experiencer verbs*. The forms that enter passivization in the syntax are forms that have already undergone a process of causativization in the lexicon which has introduced an additional external argument to the original argument structure of the experiencer verbs which consists of only the experiencer argument[7] (Tsimpli (1989, p. 289)).

Tsimpli's (1989) analysis of MG ESPVCs is based on two assumptions:

1. that the forms ending in -*ome*[8] are passives, and

2. that most of these forms admit an *apo*-PP as an optional dependent.

These assumptions, though, are not unproblematic:

1. Markantonatou (1995) has shown that

> "...although it is true that for most of the MG ESPVCs ending in -*ome* one could find an active EOPVC counterpart, there are at least three ESPVCs – i.e., *vari-eme* (be bored), *onirev-ome* (dream of), and *her-ome* (enjoy/be happy) – which do not have any active EOPVC counterpart. One possible explanation for this that Tsimpli (1989) does not seem to have taken into consideration is that in MG verbs ending in -*ome* are not necessarily PASSIVE; they can be deponent

[6] That is, both ESPVCs, and EOPVCs (Experiencer-Object Psych Verb Constructions).

[7] To draw a connection to Grimshaw's (1990) account of ESPVCs that we have just outlined above, under Tsimpli's (1989) analysis those ESPVCs in MG which she takes to be passives of EOPVCs do not have an EXTERNAL ARGUMENT in their a-structure; thus, they cannot be predicted (on some accounts) to undergo passivization. This conclusion is compatible with what we have shown in examples (14)-(17).

The only case which might constitute a counter-example is the case of the "passive" form of the verb *agapo* (love) (*agapieme* (be loved)) (the examples are from Markantonatou (1995, p. 290)):

(1) To tragudi afto agapithike apo tus anthropus tis epohis tu.
the song.N this love.PASS.PAST.3S by the people the time its
"This song was popular among the people of its time."

(2) O Gianis agapithike *(apo tin adelfi tu).
the Gianis.N love.PASS.PAST.3S by the sister his
"John was loved by his sister."

But for this case we will agree with Markantonatou (1995, p. 290) that the meaning of the verb *agapieme* (be loved) in (1) is more something like "to be popular".

[8] Which is the typical ending of the passive verb constructions in MG.

verbs (e.g., *erhome* (to come)), or middle verbs (e.g., *diavazete (efkola)* (reads easily), cf., Condoravdi (1989)), or have a reflexive (e.g., *htenizome* (to comb myself)) or reciprocal meaning (e.g., *voithiomaste* (we help each other), cf., Theophanopoulou-Kontou (1985))" (Markantonatou (1995, p. 289)).

2. It is unclear to us what Tsimpli's (1989) analysis would be in the case of ESPVCs like *agapo* (love), *epithimo* (desire/want), *zilevo* (envy), *thavmazo* (admire), *thelo* (want), *latrevo* (adore), *miso* (hate), *nostalgo* (long for), *simpono* (sympathise with) which are neither passive-in-form, nor do they have an EOPVC counterpart.

3. Finally, according again to Markantonatou (1995),

   > "Tsimpli's assumption that most of the MG ESPVCs ending in -*ome* admit an *apo*-PP as an optional dependent is false, since the "suppressed" argument (such an argument should exist, if MG ESPVCs were indeed passive forms) can be expressed with a variety of PPs: *gia*-PP ("for"-PP), *me*-PP ("with"-PP), and *apo*-PP ("from"-PP).[9] In addition, different prepositions are associated with different interpretations. Furthermore, some of the MG ESPVCs ending in -*ome* do not accept at all an *apo*-PP dependent. Such predicates are: *endiaferome* (be interested), and *stenahorieme* (be upset)" (Markantonatou (1995, p. 289)).

The conclusion falling out from the discussion above is that pure a-structure accounts like Grimshaw's (1990), as well as analyses like the one proposed by Tsimpli (1989) which specifically predict that MG ESPVCs ending in -*ome* are the passive forms of the corresponding EOPVCs cannot account for the syntactic behaviour of MG ESPVCs.

2.2 Attempts at unified linking accounts of MG ESPVCs

A unified linking account of constructions like the ones in (1)-(3) and (8)-(11) has been the aim of Markantonatou's (1995) approach to MG ESPVCs.

Markantonatou (1995) focused for this purpose on the EXPD semantic argument of MG ESPVCs and proposed that this argument can be either semantically underspecified, or syntactically restricted. In the former case it bears the intrinsic classification (IC) feature [-o], while in the latter it bears the intrinsic classification (IC) feature [+r]:[10]

Intransitive ESPVCS			
predicate	<EXPR	EXPD>	
	-r	-o	Intrinsic Classification (IC)
	SUBJ	OBL	Mapping Principles

(18) O Gianis endiaferete gia sena.
the Gianis.N be-interested.3S for you
"John is interested in you."

[9]Passive forms in MG take only an *apo*-PP dependent; no other PP is licensed to encode the suppressed argument. Clearly, the situation is different with MG ESPVCs ending in -*ome*. This is one more argument against Tsimpli's (1989) claim that MG ESPVCs ending in -*ome* are the passive forms of the corresponding EOPVCs.

[10]The examples are from Markantonatou (1995, p. 296).

Transitive ESPVCs.

predicate	<EXPR	EXPD>	
	-r	+r	Intrinsic Classification (IC)
	SUBJ	OBJ_θ	Mapping Principles

(19) O Gianis agapa tin Maria.
the Gianis.N love.3S the Maria.A
"John loves Mary."

This specific Intrinsic Classification (IC) of the EXPD semantic argument of MG ESPVCs is stipulative and relies on the assumption that semantic arguments which are not related to any Dowty-like Proto-Agent or Proto-Patient entailments at all are intrinsically classified [-o], as well as on the assumption that the EXPD semantic argument of the "transitive"[11] MG ESPVCs is syntactically restricted. This latter assumption is based on the typological principle that

> "in languages in which SUBJ and (OBJ?) is encoded through case-marking and agreement (and not via word order) lexically case marked participants are always syntactically restricted" (i.e., intrinsically classified as [+r]) (Zaenen (1993, p. 152)).

To show that stipulative as it might be such an assumption holds for the EXPD semantic argument of MG ESPVCs, Markantonatou (1995) claims that

1. the surfacing accusative NP of the "transitive" MG ESPVCs is not related to passive adjectives, and
2. MG ESPVCs
   (a) do not passivize; and
   (b) lack an EXTERNAL = $\hat{\theta}$ [-o] a-structure argument.

As far as her first claim is concerned, Markantonatou (1995) relies on Ackerman (1992), who claims that passive adjectives are related only to predicates which have a [-r] argument and bear the feature [+telic]. However, Bresnan (1996) has shown that the ability of nominals to be related to passive adjectives has nothing to do with their intrinsic classification (IC) features. Rather, it has to do with the semantics of the base verb the surfacing accusative NP combines with, which has to denote a *result state* (cf., Bresnan (1996)).

As far as Markantonatou's second claim is concerned, it is true that MG ESPVCs do not passivize (see also examples (14)-(17)). But this does not fall from her account, at least as far as the variant of MG ESPVCS in example (18) is concerned. That is, according to her analysis, this variant of MG ESPVCS has a [-o] argument in its a-structure (see example (18)). Thus, it should have been able to passivize, which is obviously not the case.

Closing we want to underline that Markantonatou's (1995) analysis is the first attempt at a unified account of MG ESPVCs (cf., (8)-(11) in Section (1)), based on the semantic and syntactic properties of the EXPD argument of these constructions. The unified linking account we propose in the following for MG ESPVCs tries to overcome the problematic aspects of this analysis.

3 Towards a robust linking account of MG ESPVCs

To recapitulate, as far as the syntax of MG ESPVCs is concerned, we saw in Section (2) that these constructions:

1. do not passivize,

[11] To follow Markantonatou's (1995) terminology.

2. are not the passive forms of the corresponding EOPVCs, and
3. realize syntactically the EXPD semantic role either as the object of the sentence, or as the complement of a prepositional phrase.

Now, in order to account for the semantic and linking properties of the arguments of the constructions in (1)-(3) and (8)-(11), we focus on both the EXPR ("experiencer") and the EXPD semantic arguments of MG ESPVCs, and rely on Butt, Dalrymple, and Frank (1997), and Wechsler (1995).

3.1 The EXPR ("experiencer") semantic argument

Zaenen (1993) has shown that semantic arguments with equal number of Dowty-like Proto-Agent and Proto-Patient properties are assigned the Intrinsic Classification (IC) feature [-r]. The EXPR semantic role of MG Psych Verb Constructions is such an argument (cf., examples (1)-(5) in Section (1)), since the Dowty-like entailments related to it – *sentience and/or perception* and *undergoing a change of (mental) state* – classify it neither as a pure Proto-Agent semantic role, nor as a pure Proto-Patient one.

That the EXPR ("experiencer") semantic argument of MG ESPVCs must link to the IC feature [-r] is also supported by the fact that MG ESPVCs do not passivize, as has been shown in (14)-(17) in Section (2.1) above, repeated here for convenience:

(20)

I	gonis	tu	agapun	ton	Giani.
the	parents.N,PL	his	love.3PL	the	Giani.A

"His parents love John."

(21)

*O	Gianis	agapiete	apo	tus	gonis	tu.
the	Gianis.N	love.PASS.3S	by	the	parents	his

"John is loved by his parents."

(22)

I	Maria	zilepse	to	spiti	ton	gitonon.
the	Maria.N	envy.PAST.3S	the	house.A	the	neighbours.G,PL

"Mary envied the neighbours' house."

(23)

*To	spiti	ton	gitonon	zileftike	apo	tin	Maria.
the	house.N	the	neighbours.G,PL	envy.PASS.PAST.3S	by	the	Maria

"The neighbours' house was envied by Mary."

What (20)-(23) show is that there is no $\hat{\theta}$ [-o] argument in the a-structure of MG ESPVCs available to be suppressed by the morphosyntactic (*Suppression*) operation of passivization:

(24) *The Suppression Morphosyntactic Operation*

$$\begin{array}{c} \hat{\theta} \\ | \\ \emptyset \end{array}$$

In this case, the argument which is realized as the subject of constructions like MG ESPVCs must bear the IC feature [-r], according to the following default mapping principles proposed by Bresnan and Zaenen (1990):

(25)
- $\hat{\theta}$ [-o] is mapped onto SUBJ; otherwise,
- θ [-r] is mapped onto SUBJ
- Other roles are mapped onto the lowest compatible function on the markedness hierarchy.

and interpreted by Butt, Dalrymple, and Frank (1997) as follows:

> "...the intrinsic classifications [of Lexical Mapping Theory (LMT)] are augmented by a set of *default mapping principles* which induce full specification of the grammatical function of the thematic role. That is, the default principles are taken to resolve the disjunctive possibilities specified by the intrinsic features. The default principles...can be read: if available, the external argument...has to be mapped onto the subject; if there is no external argument, an internal argument is mapped onto the subject. All other roles are mapped onto the lowest compatible function on the markedness hierarchy..." (Butt, Dalrymple, and Frank (1997, p. 4)).

From the above we conclude that the EXPR ("experiencer") semantic argument of MG ESPVCs is a [-r] argument.

3.2 The EXPD ("experienced") semantic argument

In the account of MG ESPVCs that we propose here, the EXPD semantic argument of these constructions is classified either as [+o] or as [+r].

[+o] is the intrinsic linking possibility we propose for the EXPD ("experienced") semantic argument in the following MG ESPVCs:

(26) O Gianis agapa tin Maria.
the Gianis.N love.3S the Maria.A
"John loves Mary."

(27) I Maria.N fovate ton Giani.
the Maria.N fear.3S the John.A
"Mary is afraid of John."

(28) I Maria fovate tis kategides.
the Maria.N fear.3S the storms.A
"Mary is afraid of the storms."

This proposal is based on the fact that themes[12] link traditionally to the IC feature [+o] (see a.o. Bresnan and Zaenen (1990), Butt, Dalrymple, and Frank (1997)).[13]

[+r] is the intrinsic linking possibility we propose for the EXPD ("experienced") semantic argument in the following MG ESPVCs:

(29) I Maria fovate me tis kategides.
the Maria.N fear.3S with the storms
"Mary is afraid of the storms."

(30) I Maria.N fovate me ton Giani.
the Maria.N fear.3S with the John.A
"Mary is afraid of John."

[12] Which is what the EXPD ("experienced") semantic argument in (26)-(28) is. In classical LMT, *theme* is the role that is assumed to be semantically most neutral. Thus, our choice to adopt *theme* for the second argument of MG ESPVCs is in agreement with Markantonatou's (1995) assumption that the EXPD semantic argument of MG ESPVCs may be semantically underspecified (see Section (2.2)).

[13] To objections that the surfacing accusative NP of MG ESPVCs like the ones in (26)-(28) does not behave as a NORMAL OBJ because MG ESPVCs do not passivize (cf., Markantonatou (1995)), we repeat what we have shown in Section (3.1) above: MG ESPVCs do not passivize because they lack a $\hat{\theta}$ [-o] in their a-structure.

This proposal[14] is based on Wechsler's (1995) suggestion that "a feature [+r] [is]...for "restricted" or oblique roles, i.e., roles which may be linked to a semantically restricted complement such as a PP..." (cf., Wechsler (1995, p. 63)). That is the case in examples (29) and (30) above.

3.3 Linking MG ESPVCs

Thus, taking into consideration the proposals we have made in Sections (3.1) and (3.2) above for the intrinsic linking possibilities of the EXPR ("experiencer") and the EXPD ("experienced") semantic arguments, the linking of MG ESPVCs like the ones in (26)-(30) is as follows:

	agapo	<EXPR	EXPD>
(31)	Intrinsic Classification (IC)	-r	+o
	Mapping Principles (MP)	SUBJ	OBJ

O Gianis agapa tin Maria.
the Gianis.N love.3S the Maria.A

"John loves Mary."

	fovame	<EXPR	EXPD>
(32)	IC	-r	+o
	MP	SUBJ	OBJ

I Maria.N fovate ton Giani.
the Maria.N fear.3S the John.A

"Mary is afraid of John."

	fovame	<EXPR	EXPD>
(33)	IC	-r	+o
	MP	SUBJ	OBJ

I Maria fovate tis kategides.
the Maria.N fear.3S the storms.A

"Mary is afraid of the storms."

	fovame	<EXPR	EXPD>
(34)	IC	-r	+r
	MP	SUBJ	OBL_θ

I Maria fovate me tis kategides.
the Maria.N fear.3S with the storms

"Mary is afraid of the storms."

[14]Our proposal deviates from Butt, Dalrymple, and Frank's (1997) intrinsic classification rules for *theme*. They assume that *theme* links alternatively either to [-r] or [+o]. This, though, would have falsely predicted SUBJ-OBJ assignment for both (26)-(28) and (29)-(30). The mismatch between the two systems is due to the fact that Modern Greek needs to account for themes which may link to a restricted complement such as a PP.

	fovame	<EXPR	EXPD>
(35)	IC	-r	+r
	MP	SUBJ	OBL_θ

I Maria.N fovate me ton Giani.
the Maria.N fear.3S with the John.A

"Mary is afraid of John."

3.4 Summary

Up to here we have investigated the semantic properties of the participants of the states denoted by Experiencer-Subject Psych Verb Constructions (ESPVCs) in Modern Greek, and the effects that these properties have on the syntactic realization of the semantic roles involved in these constructions. The linking from the semantics to the syntax of MG ESPVCs is achieved by the intrinsic linking feature [-r] for the EXPR ("experiencer") semantic role and the intrinsic linking features [+o] or [+r] for the EXPD ("experienced") semantic role of these constructions (see examples (31)-(35) in Section (3.3) above).

In the following we will show that the specific linking of the semantic roles in (32) and (33) to their respective syntactic arguments is related to the constructions in (32) and (33) ranking higher than the constructions in (35) and (34) in the native speakers' preference.

To do this, we rely on the linking theory proposed by Butt, Dalrymple, and Frank (1997). The reason we adopt this approach is that the linking architecture proposed by Butt, Dalrymple, and Frank (1997) incorporates naturally a set of preference constraints which impose an ordering on the available linking possibilities of given predicators. This is actually what helps us support the observation that the constructions in (32) and (33) rank higher than the constructions in (35) and (34) in the native speakers' preference.

4 "Optimal Linking"

Following the principles of Lexical Mapping Theory (LMT) in their overall spirit, Butt, Dalrymple, and Frank (1997) reformulate the precise statement of linking in Lexical Mapping Theory in terms of a theory of under-specification, and preference rankings.

They propose that

> "argument structure is an additional level of representation, related directly to the c-structure, and the f-structure of Lexical Functional Grammar (LFG), and indirectly to other levels".

The grammatical architecture they propose differs from usual assumptions in that

> "argument structure is projected directly from c-structure. That is, the α projection function maps nodes of the c-structure tree to pieces of the argument structure. Argument structures are mapped to f-structures by the linking function λ, which in a sense represents the integration of linking theory into the projection architecture. Thus, the familiar ϕ projection relating the c-structure to the f-structure can be seen as a composition of the α and λ functions" (Butt, Dalrymple, and Frank (1997, pp. 1-2)).

In their grammatical architecture, the relationship between the a-structure and the c-structure is a many-to-one relationship, as has been the case in traditional LFG for the relationship between the c-structure and the f-structure.

Their approach, as they declare,

"departs most radically from the Lexical Mapping Theory (LMT) literature in that they do not assume that a-structure roles are deterministically and uniquely linked to grammatical functions via a set of default principles. Instead, they propose a set of preference constraints which impose an ordering on the available linking possibilities; the most preferred possibility or possibilities are chosen" (Butt, Dalrymple, and Frank (1997, p. 5)).

In the theory they propose, the core of preference constraints is formed by two basic considerations: one is the preference ranking of the individual grammatical functions:

(36) SUBJ$>$OBJ$>$OBJ$_\theta$ / OBL$_\theta$

According to (36) above, subjects are most highly ranked in that they represent the function that is universally required in every clause, as the Subject Condition dictates. Next come objects, and then obliques, and semantically r(estricted) objects, on a par. They restate this preference ranking as in the following:

(37) [-r]$>$[+r]

(38) [-o]$>$[+o]

"in order to capture the intuition that unrestricted functions, like, for instance, the function [-r], are more preferred than restricted functions, and that within those unrestricted functions, it is the non-objective functions — that is, the subjects — which are more highly preferred" (Butt, Dalrymple, and Frank (1997, p. 6)).

The second consideration forming the core of preference constraints has to do with the relation between the a-structure and the f-structure. That is, Butt, Dalrymple, and Frank propose

"another kind of preference constraint, which is sensitive to this very relation: the SUBJ is preferentially linked to the highest non-suppressed argument. This constraint crucially relies on the notion of the thematic hierarchy, and makes use of the relation of *outranking* presented in the following hierarchy:

AGENT$>$BENEFICIARY$>$EXPERIENCER / GOAL$>$INSTRUMENT$>$PATIENT / THEME $>$ LOCATIVE" (Butt, Dalrymple, and Frank (1997, p. 6))

Thus, what they propose is that

"the linking that will be chosen is the one that best satisfies the preference constraints (36), (37), and (38). In addition, language particular preference constraints can be added to these basic constraints in order to reflect the case marking, or syntactic properties peculiar to each individual language" (Butt, Dalrymple, and Frank (1997, p. 6)).

Their view of linking differs from the fully deterministic linking principles of the traditional Lexical Mapping Theory (LMT) in that the preference constraints for linking that they propose presuppose an integration of argument structure into LFG's projection architecture. In their view, this provides for more flexibility in the treatment of argument alternations, as well as for a more natural representation of the influence of extra-thematic information such as discource structure (topic, focus) on the realization of grammatical functions.

4.1 "Optimal Linking" and MG ESPVCs

In this section we apply the linking theory of Butt, Dalrymple, and Frank (1997), which was briefly outlined in Section (4) above, on MG ESPVCs like the ones in (27)-(30) that we have shown in Section (3.2).

We need to add here that in Butt, Dalrymple, and Frank (1997) the preference constraints mentioned in Section (4) above are by way of illustration expressed in terms of numeric weights like in (39). This means of encoding of preference ranking is maintained in the analysis of MG ESPVCs below.

	SUBJ:	+3
	OBJ:	+2
(39)	OBJ_θ:	+1
	OBL_θ:	+1
	SUBJ linked to thematically highest argument (Subject Preference (SP)):	+1

Applying, thus, Butt, Dalrymple, and Frank's (1997) linking theory on MG ESPVCs like the ones in (27)-(30) we get the following results:

	fovame	<EXPR	EXPD>			
	Intrinsic Classification (IC)	-r	+o			
(40)	Mapping Principles (MP)	SUBJ	OBJ			
				SP	Total	Optimal
	Numeric Weights (NW)	+3	+2	+1	6	√

I Maria.N fovate ton Giani.
the Maria.N fear.3S the John.A

"Mary is afraid of John."

	fovame	<EXPR	EXPD>			
	IC	-r	+r			
(41)	MP	SUBJ	OBL_θ			
				SP	Total	Optimal
	NW	+3	+1	+1	5	

I Maria.N fovate me ton Giani.
the Maria.N fear.3S with the John.A

"Mary is afraid of John."

	fovame	<EXPR	EXPD>			
	IC	-r	+o			
(42)	MP	SUBJ	OBJ			
				SP	Total	Optimal
	NW	+3	+2	+1	6	√

I Maria fovate tis kategides.
the Maria.N fear.3S the storms.A

"Mary is afraid of the storms."

	fovame	<EXPR	EXPD>				
	IC	-r	+r				
(43)	MP	SUBJ	OBL_θ				
				SP	Total	Optimal	
	NW	+3	+1	+1	5		

I Maria fovate me tis kategides.
the Maria.N fear.3S with the storms

"Mary is afraid of the storms."

The analysis of the constructions in (40)-(43) builds on the investigation of the semantic properties of the participants of the states denoted by Experiencer-Subject Psych Verb Constructions (ESPVCs) in Modern Greek, which has been presented in Sections (3.1) and (3.2). The results of that investigation have led to the assignment of the intrinsic classification feature [-r] to the EXPR ("experiencer") argument of MG ESPVCs, and to the assignment of the intrinsic classification features [+o] or [+r] to the EXPD ("experienced") argument of the same constructions.

In (40)-(43) the specific intrinsic linking features of the EXPR ("experiencer") and the EXPD ("experienced") arguments of MG ESPVCs do not stand alone responsible for the mapping from the semantics to the syntax of these constructions, but are part of the bigger linking architecture proposed by Butt, Dalrymple, and Frank (1997). This linking architecture incorporates in a natural way a set of preference constraints (see (36), (37), and (38) in Section (4)), which

1. constrain the mapping from the semantics to the syntax of given predicators directly, since they are integrated in the linking algorithm,[15] and

2. impose a natural ordering on the available linking possibilities of given predicators.

Applied to MG ESPVCs as shown in (40)-(43), the "optimal linking" architecture results in:

1. making the correct predictions as far as the linking of the EXPR ("experiencer") and the EXPD ("experienced") semantic arguments of MG ESPVCs to the syntax is concerned, and

2. predicting that the variant of MG ESPVCs which realizes syntactically the EXPD ("experienced") semantic argument as an accusative NP (examples (40) and (42)) ranks higher (is "more optimal" ($\surd$)) than the variant which realizes syntactically the EXPD ("experienced") semantic argument as the complement of a PP phrase (examples (41) and (43)).

The variant of MG ESPVCs which ranks lower does not correspond to an ungrammatical construction (examples (41) and (43)). In fact, within the "optimal linking" system this variant is treated as the more *marked* option of MG ESPVCs which is only chosen in case it is triggered by the presence of the preposition (*me* (with), for instance) in the "environment" of the verb *fovame* (fear).[16,17] Thus, the complete range of the linking possibilities for the EXPD ("experienced") argument of the verb *fovame* (fear) can be constrained through other requirements of the specific ESPVCs. In general, according to Butt, Dalrymple, and Frank (1997),

> "[this allows] for a more flexible correlation between argument arrays and possible linkings to grammatical functions" (Butt, Dalrymple, and Frank (1997, p. 7)).

[15] This is not the case in LMT, where the preference constraints are activated after the linking from the semantics to the syntax of given predicators has been achieved by the intrinsic classification features [+/-o] and [+/-r] solely.

[16] In the Greek ECI corpus there are 24 instances of the *marked* variant of the kind we have seen in the sentences (41) and (43) in a total of 431 MG ESPVCs headed by the verb *fovame* (fear). That is, the rate of the *marked* variant is 5.568%.

[17] For more on the linking of indirect arguments in English see Markantonatou and Sadler (1996). For more on the way that the prepositions *me* (with) and *gia* (for) are treated when appearing in the "environment" of the verb *fovame* (fear) see Kordoni (1999), Kordoni (2001).

5 Conclusion

This paper focused on the semantic properties and the syntactic behaviour of MG ESPVCs like the following:

(44) I Maria fovate tis kategides.
the Maria.N fear.3S the storms.A
"Mary is afraid of the storms."

(45) I Maria fovate me tis kategides.
the Maria.N fear.3S with the storms
"Mary is afraid of the storms."

(46) I Maria.N fovate ton Giani.
the Maria.N fear.3S the John.A
"Mary is afraid of John."

(47) I Maria.N fovate me ton Giani.
the Maria.N fear.3S with the John.A
"Mary is afraid of John."

We accounted for the semantic and syntactic properties of the constructions in (44)-(47) by relying on the linking architecture that Butt, Dalrymple, and Frank (1997) have proposed as an alternative to the fully deterministic principles of standard LMT (see Sections (4)). In Section (4.1) we showed that the "optimal linking" theory makes the correct predictions for the linking of the EXPR ("experiencer") and the EXPD ("experienced") semantic arguments of MG ESPVCs and predicts that the variant of MG ESPVCs which realizes syntactically the EXPD ("experienced") semantic argument as an accusative NP (examples (44) and (46)) ranks higher (is "more optimal") than the variant which realizes syntactically the EXPD ("experienced") semantic argument as the complement of a PP phrase (examples (45) and (47)). This ranking reflects in a way the fact that in order to express the meaning *Mary is afraid of the storms* native speakers prefer the construction in (44), rather than the one in (45), which is also mirrored in the rate of the *marked* variant of MG ESPVCs ((45) and (47)) in the Greek ECI corpus (5,568%).

Future work includes the formalization of the proposal we have presented in Section (4.1) in a OT setting with stochastic evaluation (Boersma (1998), Boersma and Hayes (2001), Bresnan, Dingare, and Manning (2001)), which brings along the useful for our purposes advantage that it can generate both categorical (cf., examples (44) and (46)) and variable outputs (cf., examples (45) and (47); Bresnan, Dingare, and Manning (2001)).

References

Ackerman, F. (1992). Complex Predicates and Morphological Relatedness: Locative Alternation in Hungarian. In I. A. Sag and A. Szabolcsi (Eds.), *Lexical Matters. CSLI Lecture Notes no. 24*, pp. 55–84. Stanford, Calif.: CSLI Publications.

Belletti, A. and L. Rizzi (1988). Psych Verbs and Theta-Theory. *NLLT 6*, 297–352.

Boersma, P. (1998). *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives*. Ph. D. thesis, University of Amsterdam, Amsterdam.

Boersma, P. and B. Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry 32.1*, 45–86.

Bresnan, J. (1996). Lexicality and Argument Structure. Invited paper given at the Paris Syntax and Semantics Conference, October 12-14, 1995. Corrected version: April 15, 1996. 27 pages. Available at:http://www-lfg.stanford.edu/lfg/bresnan/download.html.

Bresnan, J., S. Dingare, and C. Manning (2001). Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi. In M. Butt and T. H. King (Eds.), *Proceedings of the LFG01 Conference, University of Hong Kong, Hong Kong*, Stanford, Calif. CSLI Publications. Available at: http://csli-publications.stanford.edu.

Bresnan, J. and A. Zaenen (1990). Deep Unaccusativity in LFG. In K. Dziwirek, P. Farrell, and E. Mejias-Bikandi (Eds.), *Grammatical Relations: A Cross-Theoretical Perspective*, pp. 45–57. Stanford, Calif.: Stanford Linguistics Association and CSLI Publications.

Butt, M., M. Dalrymple, and A. Frank (1997). An architecture for linking theory in LFG. In M. Butt and T. H. King (Eds.), *Proceedings of the LFG97 Conference*, University of California, San Diego. CSLI Publications.

Condoravdi, C. (1989). The Middle: where semantics and morphology meet. *MIT Working Papers in Linguistics* (II), 16–30.

Dowty, D. (1991). Thematic Proto-Roles and Argument Selection. *Language 67*, 547–619.

Grimshaw, J. (1990). *Argument Structure*. Cambridge, Massachusetts: MIT Press.

Kordoni, V. (1999). Lexical Semantics and Linking in HPSG: the case of Psych Verb Constructions. In V. Kordoni (Ed.), *Tübingen Studies in Head-Driven Phrase Structure Grammar*, SFB 340, Bericht Nr. 132, pp. 494–527, (Vol. II). Universität Tübingen: Seminar für Sprachwissenschaft.

Kordoni, V. (2001). Linking Experiencer-Subject Psych Verb Constructions in Modern Greek. In D. Flickinger and A. Kathol (Eds.), *Proceedings of the 7th International HPSG Conference, UC Berkeley (22-23 July, 2000)*, Stanford, Calif., pp. 198–213. CSLI Publications. Available at: http://csli-publications.stanford.edu.

Markantonatou, S. (1995). Modern Greek deverbal nominals: an LMT approach. *Journal of Linguistics 31*, 267–299.

Markantonatou, S. and L. Sadler (1996). Linking Indirect Arguments. *Essex Research Reports in Linguistics 9*, 24–63.

Theophanopoulou-Kontou, D. (1985). Patient vs. non-patient orientation of the action and the voice distinction in MG. *Glossologia 3*, 75–90.

Tsimpli, I.-M. (1989). On the properties of the passive affix in Modern Greek. *Working Papers in Linguistics, University College London 1*, 235–261.

Wechsler, S. (1995). *The Semantic Basis of Argument Structure*. Stanford: CSLI Publications. Series: *Dissertations in Linguistics*, Joan Bresnan, Sharon Inkelas, William J. Poser, and Peter Sells (eds.).

Zaenen, A. (1993). Unaccusativity in Dutch: Integrating Syntax and Lexical Semantics. In J. Pustejovsky (Ed.), *Semantics and the Lexicon*, pp. 129–162. Dordrecht: Kluwer Academic Publishers.

Valia Kordoni <kordoni@coli.uni-sb.de>
University of Saarland
Computerlinguistik
Postfach 15 11 50
D-66041 Saarbrücken, Germany

ANOTHER LOOK AT PARTICIPLES AND ADJECTIVES IN THE ENGLISH DP

Tibor Laczkó
Department of English Linguistics
University of Debrecen
laczko@tigris.klte.hu

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu

1. Introduction

In this paper[1] I will address the following main question. Are all the formally participial (verbal) elements modifying NP heads in English adjectives or do they (or any of them) retain their original participial (verbal) category? Interestingly, in recent approaches (cf. Bresnan (1982, to appear), Levin and Rappaport (1986) and Ackerman and Goldberg (1996)), it seems to be a strong tendency to consider these prenominal (and, often, even postnominal) participial modifiers to be adjectives uniformly in English.[2] In these works, the explicitly stated or tacitly assumed generalization is that participles obligatorily undergo participle → adjective conversion prior to being used as modifiers in NPs.

On the basis of a systematic comparison of the relevant English and Hungarian data, I will challenge this view and I will propose a more articulated analysis of the relevant English phenomena. I will argue that under clearly definable circumstances even prenominal modifiers must be taken to be participles and not adjectives and, consequently, two separate rules are needed: one for capturing the use of participial modifiers in NPs and another for accounting for participle → adjective conversion. Thus, it will be obvious that this whole issue is not merely terminological, because it has consequences for our grammar, as regards the number and actual nature of the relevant rules.

The structure of the paper is as follows. In section 2, I will offer a critical evaluation of Bresnan's (to appear) view on the nature and behaviour of participle-looking words in the English NP. In section 3, I will propose an alternative approach to the English phenomena, based on Laczkó's (2000) analysis of the corresponding Hungarian data. This will be followed by some concluding remarks in section 4.

2. On participle → adjective conversion

Bresnan (to appear) offers a battery of tests for distinguishing between participles and adjectives, and, on these grounds, she argues for obligatory participle → adjective conversion in the NP domain.[3] She regards the participle-looking words in the following examples, among others, as adjectives.

(1) a. a smiling child
    b. a fallen leaf
    c. an opened can

---

[1] This paper is a revised and extended version of Laczkó (to appear). The extension will deal with the theoretical consequences of the empirical generalizations I will make on the basis of comparing the relevant English and Hungarian phenomena as well as the theoretical implications of the history of -*en* participles in English. I am grateful to Bozena Cetnarowska for very helpful comments on a previous version of this paper, and to Tracy King, David Kovacs, George Seel and Ted Riordan for their grammaticality judgements.

[2] For a brief terminological overview of the literature, see Cetnarowska (to appear).

[3] In actual fact, at least those *-ing* participles derived from transitive verbs that retain the input predicates' object argument must be considered participles (verbs) in her system as well, cf. the discussion below. However, she does not elaborate on this issue.

In this section I will take a closer look at Bresnan's tests and I will comment on them by also comparing the relevant English data with their Hungarian counterparts. She discusses the following five diagnostics.

A) Adjectives but not verbs (including participles) can be negated by *un-* prefixation,[4] cf.:

(2) a. happy ~ unhappy; clear ~ unclear
b. *untouch; *ungo

For instance, given that there is no verb like **unqualify*, we have to assume that in the expression *an unqualified denial* the word *unqualified* is an adjective: the participle *qualified* was first converted into an adjective and then (adjectival) *un-* prefixation took place.[5] Note, however, that this test can only safely identify certain *un-* prefixed words as adjectives and it does not tell us anything decisive about the categorial status of their "*un*-less" counterparts. It can be argued, for example, that *qualified* is an adjective in (3a) and it is a participle (verb) in (3b).[6]

(3) a. a qualified tourist guide
b. a superbly qualified tourist guide

It is also noteworthy that although Bresnan assumes that words like *smiling* in (1a) are adjectives as well, this test, contrary to one's expectations, in a great number of cases cannot be applied to support such an assumption, despite the fact that there seem to be no semantic factors that could be invoked to explain this.

(4) Yesterday I saw four people in the park. Three men were jogging and a woman was sitting on a bench. I watched one of the jogging men for a while and then I went up to the *unjogging woman.

B) Adjectives but not verbs (including participles) can be prenominal modifiers: *A N* vs. **V N*. The essence of this test is that any participle-looking word must be taken to be an adjective if it premodifies the noun head (cf. all the three examples in (1), for instance). It does not say anything about postmodifying elements. This is all the more surprising, because ordinary (that is, underived) adjectives must also, as a rule, follow the noun head when they take complements. Compare:

(5) a. the proud mother ~ the mother proud of her daughter
b. the smiling boy ~ the boy smiling at the girl

Thus, premodification is not a sufficient test even in the case of adjectives. Naturally, one could try to save the test by assuming that if a participle-looking element, without a PP

---

[4] In English there are two *un-* prefixes. One of them can only attach to adjectives (as is shown in (2)) and it expresses negation and the other can only combine with verbs and the prefixed verb refers to an action that is the reverse of the action denoted by the input verb, cf. *lock ~ unlock*, *button ~ unbutton*.

[5] For a detailed discussion (and the relevance of this to the description of passivization), see Bresnan (1982).

[6] For details, see section 3.

complement or adjunct, is capable of premodifying a noun head then it has to be taken to be an adjective in this position and also in a postnominal position when it has a PP complement or adjunct. Then we could conclude that *smiling* is an adjective in both examples in (5b). There is, however, a problem with this strategy. Certain participle-looking words have an obligatory PP complement. Therefore, the prenominal test simply cannot be applied to them, cf. (6).

(6) a. *the belonging boy (to the group)
b. the boy belonging to the group

C) Adjectives can be modified by *too* without *much*:

(7) *too (*much)* AP vs. VP *too *(much)*: *it is too flat* vs. **it is too much flat*

Given that this test, as formulated by Bresnan (cf. (7)), is fundamentally concerned with the predicative use of words, it is not directly relevant to our problem, which is the category of modifying elements within the NP domain. In addition, the test is limited to words (whether adjectives or participles) that are gradable. If, despite these difficulties, one wanted to use this diagnostic then they could, perhaps, try and check the category of *smiling* in (1a) indirectly in the following way.

(8) a. *The child was too crying last night.
b. The child was crying too much last night.

It should be obvious that the test in this case does not yield conclusive results. On the one hand, (8a) is not felicitous, thus we have no strong evidence for the adjectival status of *crying*. On the other hand, the pattern in (7) and the example in (8b) might formally suggest that *smiling* is in actual fact a participle. However, the most natural analysis of (8b) is not that it contains the combination of the copula and a participle-looking element whose categorial status we are investigating but that it contains a past progressive tense form (and, thus, the participial nature of the *-ing* word is trivial, but this is not an issue we are primarily interested in).

On the basis of these considerations, I think that the following alternative form of the test would be more appropriate in the case of gradable words within the NP domain. If the participle-looking word can be used without any complements then we could check whether it can be intensified by *very* and precede the noun head or it rather combines with *very much* and they postmodify the head. Compare:

(9) a. *I was woken up by the very crying child.
b. I was woken up by the child crying very much.

The contrast in (9) suggests then that *crying* should rather be taken to be a participle.

D) Adjectives cannot have direct NP complements (as opposed to transitive verbs, including participles), cf.:

(10) *[A NP]$_{AP}$ vs. [V NP]$_{VP}$: **supportive my daughter* vs. *support my daughter*

From the perspective of the present paper, the two most important aspects of this diagnostic are as follows. (i) Even in Bresnan's system, postmodifying transitive *-ing* forms must be regarded as participles. (ii) The test offers us a tool for a proper treatment of participle-looking elements derived from ditransitive verbs. These are two issues I will return to in section 3.2.

E) Adjectives (as opposed to verbs, including participles) can head concessional relative phrases beginning with *however*:

(11) *however* AP vs. **however* VP: *however supportive of her daughter she may have been* vs. **however supporting her daughter she may have been*

It appears to me that the nature and the applicability of this test are very similar to those of Test C) (the *too* vs. *too much* diagnostic). Here, too, the predicative function of the participle-looking elements is involved, which is not directly related to their pre- or postmodifying use in the NP domain. In several cases this diagnostic does not seem to provide even indirect evidence for the alleged adjectival status of a participle-looking element, either. Consider:

(12) *However crying the baby was, I refused to give her another bottle of tea.

Partially on the basis of the tests above and partially on the basis of some further assumptions and considerations, the most important aspects of Bresnan's (to appear) view on participle-looking words in the NP domain are as follows.[7] Below, I summarize these points and I also comment on them.

1) *All premodifying participle-looking elements are adjectives in English.* This assumption is so pervasive that it has been made a diagnostic (cf. Test B)). I am not convinced that this is a well-founded assumption for the reasons I mentioned in the discussion of Test B) above. My fundamental problem is that non-derived adjectives are not confined to a prenominal position in English. For the details see the next point.

2) *Some postmodifying participle-looking elements are also adjectives and some others are participles.* Consider the following examples.

(13) a. feathers still unstuffed into their pillows
b. the boy eating popcorn noisily
c. the boy shouting (angrily) at the girl

(13a) is Bresnan's example. On the basis of Test A) *unstuffed* can only be taken to be an adjective. In (13b), however, *eating* has a direct NP complement and, thus, it must be regarded as a participle in accordance with Test D). We cannot immediately apply either of these tests to *shouting* in (13c), because it does not contain the *un-* prefix and, being used intransitively, it has no direct NP complement. So as a first approximation we could conclude that the categorial status of this participle-looking word is neutral between participle and adjective. Then we could attempt to use some of the diagnostics discussed above. First of all, we could try *un-* prefixation (Test A). This might be considered the most direct and most

---

[7] It is important to note that this view is quite dominant in the literature, whether generative or non-generative. For instance, Levin and Rappaport (1986) appear to share Bresnan's empirical generalizations, but in their GB framework they offer an analysis different from Bresnan's LFG approach. (Here I cannot discuss and compare these two accounts.) In addition, Ackerman and Goldberg (1996) also consistently talk about English "deverbal adjectives based on past participles and used attributively".

reliable test, because it leaves the entire constituent headed by the participle-looking word in its original postmodifying position. Consider:

(14) *the boy unshouting (angrily) at the girl

Thus, on the basis of this test, our conlusion could be that the participle-looking word in (13c) is not an adjective because it does not admit *un-* prefixation.[8] For the reason mentioned above, Test D) cannot be applied at all. Test B) is not really applicable because a participle-looking word with a complement (just like any non-derived adjective with its complement) cannot be used prenominally. Consider:

(15) a. *the (angrily) shouting at the girl boy
b. *the shouting (angrily) at the girl boy

As regards Test C) and Test E), it appears that they lend some support to the participial interpretation of *shouting* in (13c), cf.:

(16) a. *Yesterday the boy was too shouting (angrily) at the girl.
b. *However shouting at the girl the boy was, I decided not to intervene.

On the basis of all these considerations we may conclude that *shouting* is a participle in (13c) just like *eating* in (13b).

3. *All premodifying participle-looking words in English* (whether derived by *-en* or *-ing*) and *those postmodifying ones that are taken to be adjectives* (cf. Test B and (13a)) *are stative*. Bresnan (to appear) puts this in the following way. "It is clear that adjective conversion in general denotes a state derived from the semantics of the base verb. This seems to be true for all types of conversion, including the present participles (*a smiling woman*)" (p. 12). This assumption, again, is wide-spread in the literature (whether generative or non-generative). For instance, in his theory-neutral, typological survey, Haspelmath (1994) expresses a similar idea: "Both past passive participles (as in *the abused child*) and past unaccusative participles (as in *the wilted dandelion*) characterize their head by expressing a state that results from a previous event. The fact that they express a state has to do with the fact that they are adjectives" (p. 159)... "Habitual situtations are realis, non-past, non-future and dynamic (non-stative) but nevertheless so time-stable that it is worthwhile for languages to have verb-derived adjectives that characterize things in terms of them: **active participles**, e. g. English *traveling salesman, running water,* etc... It is probably only secondarily that present participles are used for current events, like *flying planes*, *a crying baby* ... because current events are not very time-stable" (p. 164). As regards their views on English *-ed* participle-looking words in the NP domain, Bresnan (to appear) and Haspelmath (1994) are fully identical. As for *-ing* participle-looking words, Bresnan (to appear) simply makes the claim cited above, but she does not elaborate at all on the sense in which these forms are to be considered stative. Haspelmath (1994) assumes that the fundamental (primary) function of *-ing* "adjectives" in the NP domain is the expression of habituality. Habituality can be regarded as a general property attributed to the entitity expressed by the NP (head) and, thus, this

[8] Note that there is no general ban on participle → adjective conversion and then *un-* prefixation in the case of *-ing* participles: *uninteresting, untiring, unrelenting,* etc.

characterization is stative. According to him the "current" use of *-ing* "adjectives" is only a secondary development.[9]

Now I would like to claim (contra Bresnan (to appear) and Haspelmath (1994)) that true participles are not stative. Consider (1) again, repeated below for convenience.

(1) a. a smiling child
b. a fallen leaf
c. an opened can

I think the participle-looking words in (1), which I consider to be true participles (as opposed to Bresnan (to appear) and Haspelmath (1994)), do not, by themselves, express states. In my view, these participles are as dynamic as their input verbs.

Naturally, (1b) and (1c) denote entities that are in a state preceded by a certain change of state. However, I believe that this semantics is composed of the following two major elements: *a)* the meaning of the participle, which only expresses a change of state and *b)* the meaning of the entity. From these two components the "present" state of the entity denoted by the NP (head) directly follows, but, strictly speaking, this is not the semantic function of the participle itself. The following simple argument can be invoked to support my claim. If we took seriously the idea that the participles in (1b) and (1c) are stative, then, by the very same token, we could be forced to assume that the predicates in (17a) and (17b) are also stative, because they also imply that the leaf and the can are now in a particular state.

(17) a. Another leaf has fallen.
d. This can was opened an hour ago.

I find the stative interpretation of the participle in (1a) even more counter-intuitive. Consider the following example.

(18) When I entered, one of the children started smiling. I went up to the smiling child.

The only feasible way of defending the stative *-ing* participle assumption, I think, would be to claim that the *-ing* word expresses a state in an extended sense: 'the child in a smiling state'. However, this would directly lead us to the possibility of assuming that the past progressive tense form in (19) also expresses a state in this sense.

(19) When I entered, the child was smiling.

In addition, in the case of *-ing* participle-looking words, I find Haspelmath's (1994) "habitual" → "current" directionality both counter-intuitive and, at least as regards the Hungarian counterparts, historically incorrect. On the basis of Károly (1956), among other sources, it appears that "present" participles were originally and primarily used in the "current" sense and then this use was extended to a habitual function.

In my opinion, Bresnan's (to appear) generalized adjectival approach to participle-looking words in the NP domain would receive some general theoretical motivation and

---

[9] I think it is appropriate to compare Haspelmath's (1994) notions of "habitual" and "current" to the following two uses of the non-derived adjective *polite*:
(i) You are (always) polite.
(ii) (Now) You are only being polite.

justification if it could be applied to cover all these words. However, this is not the case, because, as has been pointed out above, participle-looking words in *-ing* taking a direct NP complement are regarded as participles in her system, too.

3. An alternative proposal

In this section, first I will present a brief analysis of the relevant Hungarian facts on the basis of Komlósy (1994), as far as the participle vs. adjective distinction is concerned, and (fundamentally) on the basis of Laczkó (2000): 3.1. Then, motivated by this account, I will offer a novel approach to the English phenomena,: 3.2. This will be followed by the discussion of some major theoretical aspects and consequences of this approach supplemented by my speculation related to the historical development of *-en* participles in English: section 3.3.

3.1. A brief look at Hungarian

In Hungarian NPs, both adjectival and participial constituents, whether with or without complements, precede the noun head. Compare (5a) with (20) and (5b) with (21).

(20) a. a büszke anya
the proud mother
'the proud mother'

b. a lány-á-ra büszke anya
the daughter-her-onto proud mother
'the mother proud of her daughter'

(21) a. a mosolyg-ó fiú
the smile-ing boy
'the smiling boy'

b. a lány-on mosolyg-ó fiú
the girl-on smile-ing boy
'the boy smiling at the girl'

Obviously, this means that the prenominal position in Hungarian NPs is not available as a diagnostic for the adjectival status of a constituent (cf. Bresnan's Test B)). Komlósy (1994) proposes a battery of tests for distinguishing adjectives from participles in this language.[10] The most important ones, for our present purposes, are as follows.

A) Only adjectives can be used predicatively.
B) Only adjectives can serve as input to adverb-formation by *-An* '-ly'.
C) Only adjectives can have comparative and superlative forms.
D) Only participles can retain the input verb's arguments (except for the unexpressed subject argument of the participle, which is understood as being controlled by the NP head).

[10] For examples and discussion, see Komlósy (1992: 386–8).

Diagnostic A) means that when a participle-looking word in Hungarian corresponding to either *-ing* or *-en* words in English is combined with the copula and is predicated about a subject, it is an adjective, because participles of such forms cannot be used in this predicative way.

(22) a. A kudarc kiábrándít-otta János-t.
the failure disappoint-Past.3sg.def John-acc
'The failure dissappointed John.'

b. A kudarc kiábrándít-ó volt.
the failure disappoint-ing was
'The failure was disappointing.'

c. *A kudarc János-t kiábrándít-ó volt.
the failure John-acc disappoint-ing was
'The failure was disappointing John.'

Naturally, this diagnostic cannot have a real English counterpart, because *-ing* participles and passive *-en* participles do combine with the copula and they together make up progressive tense forms and passive predicates, respectively. In a sense, Bresnan's Test B) would be a mirror image of this diagnostic; however, I have already pointed out some of its problematic aspects.

Diagnostic B) is applicable to the English data as well, cf.:

(23) *A bírálat János-t kiábrándít-ó-an reális volt.
the criticism John-acc disappoint-ing-ly realistic was
'*The criticism was realistic disappointingly John.'

Diagnostic C) can also be applied to the corresponding English data, cf.:

(24) *The boy more liking music than Mary has bought a guitar.[11]

The examples in (25) illustrate Diagnostic C) and Diagnostic D) for Hungarian.

(25) a. a tavaly-i-nál is kiábrándít-ó-bb kudarc
the last.year-AFF-at even disappoint-ing-COMP failure
'the failure even more disappointing than last year's'

b. *a János-t a tavaly-i-nál is kiábrándít-ó-bb kudarc
the John-acc the last.year-AFF-at even disappoint-ing-COMP failure
'*the failure even more disappointing John than last year's'[12]

[11] In this example, *more* is used as the comparative marker for adjectives. It should not be confused with adverbial *more*, meaning to a greater extent, cf.:
(i) The boy liking music more than Mary has bought a guitar.
[12] Cf. Footnote 11.

(25a) is grammatical, because it contains a departicipial adjective in a comparative form without any argument other than its unexpressed subject. By contrast, (25b) is unacceptable, because the very same adjective has another argument.[13] It seems that Diagnostic D) does not have a full English counterpart, because, at least in certain cases, the unquestionably adjectival *-ed* forms can preserve arguments other than the subject, cf. (13a), repeated here for convenience.

(13) a. feathers still unstuffed into their pillows

Komlósy (1994) proposes the following condition on participle → adjective conversion from *-Ó* '-ing' ("present") participles. The *-Ó* participle has to have a theme subject and an experiencer object, cf. *kiábrándító* 'disappointing' in (22b) and (25a), and the examples in (26).

(26)	fáraszt-ó,	bosszant-ó,	felháborít-ó,	szórakoztat-ó
	make.tired-ing	annoy-ing	appal-ing	amuse-ing
	'tiresome/tiring'	'annoying'	'appalling'	'amusing'

As far as the function of the *-(V)(t)t* '-en' participial suffix and the conditions on participle → adjective conversion are concerned, Laczkó (2000) makes the following generalizations.

A) The *-(V)(t)t* affix cannot attach to unergative verbs (see (27)), hence no participle → adjective conversion can take place either.

(27) a. *a keményen dolgoz-ott fiú
the hard work-(V)(t)t boy
'*the hard worked boy'

b. cf. the boy who has worked hard

B) In the case of unaccusative verbs, there are three main possibilities.

Ba) If the verb is atelic without a telic counterpart, then no participle formation (see (28)) and, hence, no participle → adjective conversion can take place.

(28) *az áraml-ott víz
the flow-(V)(t)t water
'*the flowed water'

Bb) If the verb is telic and it has no atelic counterpart, both *-(V)(t)t* participle formation and participle → adjective conversion are possible.

(29) a. a kudarc-on felháborod-ott játékos (participle)
the failure-on become.indignant-(V)(t)t player
cca. 'the player (who has become) indignant at the failure'

b. az edzõ-nél is felháborod-ott-abb játékos (adjective)

---

[13] In this case the argument is a direct NP complement, but oblique complements are also disallowed in combination with Hungarian departicipial adjectives.

the coach-at even become.indignant-(V)(t)t-COMP player
'the player even more indignant than the coach'

Clearly, (29a) contains a *-(V)(t)t* participle, because it has retained a non-subject argument, cf. Diagnostic D), and (29b) contains an adjective converted from this participle, because it is in its comparative form.[14]

Bc) If there is an unaccusative verb pair, one of them with a perfectivizing preverb and the other without it, then the possibilities are as follows.

Bci) The version with the preverb always has a telic interpretation. It can combine with *-(V)(t)t*; however, the participle derived in this way cannot serve as input to participle → adjective conversion.

Bcii) The version without the preverb has both telic and atelic uses. It can combine with *-(V)(t)t* only in its telic use. Participle → adjective conversion is also possible.

(30) a. A meg-sül-t hús-t be-te-ttem a hûtõ-be. (participle)
the PERF-get.roasted-(V)(t)t meat-acc in-put-Past.1sg.def the fridge-into
cca. 'I put the meat that had got roasted into the fridge.'[15]

b. *A meg-sül-t-ebb hús-t kér-em.[16] (participle)
the PERF-get.roasted-(V)(t)t-COMP meat-acc want-Pres.1sg.def
cca. 'I want the meat that has got more roasted.'

c. A sül-t-ebb hús-t kér-em. (adjective)
the get.roasted-(V)(t)t-COMP meat-acc want-Pres.1sg.def
cca. 'I want the meat that has got more roasted.'

d. A hosszú nyárs-on sül-t hús-t be-te-ttem a hûtõ-be. (participle)
the long spit-on get.roasted-(V)(t)t meat-acc in-put-Past.1sg.def the fridge-into
cca. 'I put the meat that had got roasted on the long spit into the fridge.'

e. *Tegnap az étterem-ben meg-sül-t hús-t rendel-tem. (adjective)
yesterday the restaurant-in PERF-get.roasted-(V)(t)t meat-acc order-Past.1sg.def
'Yesterday I ordered roast meat in the restaurant.'

f. Tegnap az étterem-ben sül-t hús-t rendel-tem. (adjective)
yesterday the restaurant-in get.roasted-(V)(t)t meat-acc order-Past.1sg.def

---

[14] The constituent *az edzõnél* 'than the coach' is an argument, but such an element is the typical argument of any adjective in a comparative form, and it has nothing to do with the input verb.

[15] *(Meg-)sül* 'PERF-get.roasted' is a middle verb in Hungarian and it has no simplex verbal counterpart in English, that is why in (30a) I can only give a periphrasic English translation. Of course, *the roasted meat* would be a possibility to consider, but it would be interpreted as containing a passive participle derived from a transitive verb, as opposed to the Hungarian unaccusative participle.

[16] It is possible to express the same meaning with a participle in the following way:

(i) A jobb-an meg-sül-t hús-t kér-em. (participle)
the better-ly PERF-get.roasted-(V)(t)t meat-acc want-Pres.1sg
cca. 'I want the meat that has got roasted to a greater extent.'

In this example the participle is modified by a constituent (*jobban* 'to a greater extent') which can be used with verbal predicates to express a comparative meaning.

'Yesterday I ordered roast meat in the restaurant.'

(30a) exemplifies the participial use of the preverbed version. In such cases the participle expresses anteriority. (30b) and (30e) show that this form cannot be converted into an adjective. In (30b) the comparative form clearly suggests that we are dealing with an adjective, and in (30e) the intended adjectival meaning is to denote a kind of meat.[17] In both cases, it is only the participial form without the preverb that can be input to participle → adjective conversion, see (30c) and (30f). Finally, (30d) illustrates the fact that even the preverbless form can be used as a participle with a telic reading if it has a complement or adjunct.[18]

C) In the case of transitive verb stems there are also three major possibilities, which correspond to but are not fully identical with the unaccusative possibilities.

Ca) If an atelic verb with no preverb has no preverbed (telic) counterpart, it can combine with *-(V)(t)t*, and participle → adjective conversion is also possible. Consider:

(31) a. a mindenki által kedvel-t sorozat (participle)
the everybody by like-(V)(t)t series
'the series liked by everybody'

b. a Dallas-nál is kedvel-t-ebb sorozat (adjective)
the Dallas-at even like-(V)(t)t-COMP series
'the series even more liked than Dallas'

(31a) contains a participle, because it is modified by a *by*-phrase, while (31b) contains an adjective, because it is in its comparative form.

Cb) If the verb is telic (often containing a preverb) and it has no atelic counterpart, then *-(V)(t)t* can attach to it and participle → adjective conversion is also possible. Consider:

(32) a. az áruház által becsap-ott vásárló (participle)
the store by deceive-(V)(t)t customer
'the customer deceived by the store'

b. a János-nál is becsap-ott-abb vásárló (adjective)
the John-at even deceive-(V)(t)t-COMP customer
'the customer even more deceived than John'

---

[17] In Hungarian certain (obligatorily unmodified) adjectives make up compound-like constituents with their noun heads, and they denote the kind of the entity expressed by the head. One of the indications of the compound-like nature of this combination is that it follows a stress pattern typical of compounds (the noun head loses its word-initial stress). When a participle-looking form is used in such a construction, then this provides evidence that it has undergone participle → adjective conversion.

[18] It is not entirely clear to me why this modification is necessary when we want to use the preverbless form as a participle with a telic interpretation, but I think part of the explanation can be sought along the following lines. It seems that the preverbed and preverbless forms tend to be in complementary distribution when there are no additional (modifying) elements involved: a preverbed form can only be a participle and a preverbless form can only be an adjective, cf. (30a) and (30f).

(32a) contains a participle, because it is modified by a *by*-phrase, while (32b) contains an adjective, because it is in its comparative form.

Cc) If there is a transitive verb pair, one of them with a perfectivizing preverb and the other without it, then the possibilities are as follows.

Cci) The version with the preverb always has a telic interpretation. It can combine with *-(V)(t)t*; however, the participle derived in this way cannot serve as input to participle → adjective conversion.

Ccii) The version without the preverb has both telic and atelic uses. It can combine with *-(V)(t)t* in both its uses (as opposed to the corresponding accusative verbs). Participle → adjective conversion based on the telic use is also possible.

(33) a. A meg-darál-t hús-t be-te-ttem a hûtõ-be. (participle)
the PERF-mince-(V)(t)t meat-acc in-put-Past.1sg.def the fridge-into
cca. 'I put the minced meat (=the meat that had been minced) into the fridge.'

b. Tegnap a bolt-ban ve-ttem egy kis (*meg-)darál-t hús-t. (adjective)
yesterday the store-in buy-Past.1sg.def a little PERF-mince-(V)(t)t meat-acc
'Yesterday I bought a little minced meat in the store.'

c. Jó illat-a van a János által darál-t hús-nak. (participle)
good smell-its is the John by mince-(V)(t)t meat-dat
'The meat minced/being minced by John has a good smell.'

d. Jó illat-a van a darál-t hús-nak. (adjective, ?participle)
good smell-its is the John by mince-(V)(t)t meat-dat
'The minced meat/The meet being minced has a good smell.'

(33a) exemplifies the participial use of the preverbed version. In such cases the participle expresses anteriority. (33b) shows that this form cannot be converted into an adjective, and the preverbless version must be used for conversion purposes. The intended adjectival meaning is to denote a kind of meat.[19] (33c) illustrates the fact that a preverbless participle, when it is modified, can be compatible with either anteriority or simultaneity (as opposed to the obligatorily anterior reading of the corresponding unaccusative participle).[20] In (33d) the much more natural interpretation of the *-(V)(t)t* form is adjectival. On the far more marked participial reading it can only express simultaneity.[21]

As should be clear from the discussion above, conversion from *-(V)(t)t* participles derives two types of adjectives. One of them belongs to ordinary (non-derived) adjectives: this type can be used in the comparative and superlative forms, it can be modified by adverbs

---

[19] Cf. Footnote 17. It is also noteworthy that *darált* 'minced' as a kind-denoting adjective does not have comparative and superlative forms, so this test cannot be applied here as opposed to (30c).

[20] The semantics of the modifier may restrict the relative temporal interpretation of these constructions. Consider:

(i) Jó illat-a van a tegnap darál-t hús-nak. (participle, anteriority)
good smell-its is the yesterday grind-(V)(t)t meat-dat
'The meat ground yesterday has a good smell.'

[21] For some speculation on why the preferred interpretation of unmodified preverbless forms is adjectival, see Footnote 15.

capable of modifying adjectives in general and usually it can be used predicatively. This type is exemplified in (29b), (31b) and (32b). The other type normally lacks these adjectival features[22] and it makes up a compound with the noun head and it specifies the kind of the entity expressed by the noun. Such adjectives can be found in (30f) and (33b).

3.2. Participle-looking words in English NPs revisited

On the basis of the discussion in sections 2 and 3.1, I would like to make the following assumptions and generalizations.

A) Prenominal modifiers are not necessarily adjectives: they can also be participles; and certain postmodifying participle-looking words must be taken to be participles in any analysis.

B) Thus, it is not possible to capture the use of all prenominal participle-looking words in NPs by the help of a single general rule (contra Bresnan's (to appear) adjectival and stative generalization). We need two (sets of) rules: one for the use of *-ing* and *-en* participles and another for participle → adjective conversion.

C) The relevant generalizations for *-ing* forms are as follows.

Ca) Participles derived from any one of the three major verb types (unaccusative, unergative and transitive) can be used in NPs, cf.:

(34) a. I couldn't catch the tile falling from the roof.
b. I couldn't catch the falling tile.
c. The boy shouting at that girl used to be my friend.
d. The shouting boy used to be my friend.
e. The boy reading a newspaper used to be my friend.

This participial usage cannot be appropriately characterized as 'stative' (whether a result or non-result state is posited).

Cb) Participles ending in *-ing* can serve as input to participle → adjective conversion:

(35) a. a surprising fact, an amusing story, etc.
b. a travelling salesman, a wandering minstrel, etc.

It seems that in English, just like in Hungarian, two kinds of adjectives can be derived: ordinary, see (35a), and kind-denoting, see (35b).

D) The relevant generalizations for *-en* forms are as follows.

Da) When passivizing *-en* attaches to transitive verbs, it passivizes them and, when they are used in NPs, both atelic and simultaneous interpretations are available, depending on the (a)telicity of the input verb.

(36) a. The photograph taken by my friend is a real masterpiece.
b. The teacher respected by all his students is going to resign.
c. The highly respected teacher is going to resign.

[22] Although some departicipial adjectives of this sort can occasionally take a comparative form, see (30c).

Db) Perfectivizing *-en* can attach to all the three major verb types (unaccusative, unergative and transitive) and in clauses it combines with the perfect auxiliary *have*. However, it is only perfect *-en* participles derived from unaccusative verbs that can be used in NPs.

(37) a. The river frozen last night is still dangerous.
b. The (recently) frozen river is still dangerous.
c. *The man worked a lot at the office is tired.
d. *The man discovered the murder notified the police.

Dc) It is only in the case of this participle type that the "result state" condition (generalized to all kinds of participle-looking words in numerous works) really holds. The reason for this is simple. It is only in this instance that both perfect *-en* and present *-ing* can attach to the same stem and the derived form can be used in fundamentally the same ("active") function in NPs.[23] As their functions are the same (they both have a theme subject), their relative time features come into play, and the *-ing* forms typically express simultaneity and the perfect *-en* forms express anteriority, cf.:

(38) a. We couldn't catch the falling tiles.
b. We picked up the fallen tiles.

Dd) Although unaccusative perfect *-en* participles can occasionally serve as input to participle → adjective conversion (*fallen angels*), this is not as frequent as in the case of the corresponding Hungarian participles derived by *-(V)(t)t*. Passive *-en* participles, by contrast, are much more often a source for such a conversion, and both "ordinary" and "kind-denoting" adjectives can be derived, see (39a) and (39b), respectively.

(39) a. satisfied, exhausted, disappointed
b. salted peanuts, minced meat, granulated sugar

3.3. Theoretical considerations

From the foregoing discussion it should be obvious that the status of participle-looking *-(V)(t)t* and *-Ó* words in Hungarian and their possible LFG-style treatment, once the necessary empirical generalizations have been correctly made, is straightforward and unproblematic. There are two main reasons for this. On the one hand, these words can only be used within NPs as participles (they cannot be combined with various auxiliaries, unlike their English counterparts), thus, the question of whether participle-looking words ending in *-(V)(t)t* and *-Ó* can occur in NPs as participles at all simply does not arise, and, on the other hand, there are very reliable tests even within the NP domain to tell the participial and the adjectival uses apart. In English, by contrast, *-en* and *-ing* participles can be combined with auxiliaries, and the categorial status of a number of participle-looking words within NPs is far from being easy to decide. I hope to have been able to demonstrate in this paper that a relatively clearly definable set of English participle-looking words in the NP domain are true participles and

[23] Although both can attach to unergatives and transitives as well in the same "active" function, the *-en* version, as has been pointed out above, cannot occur in NPs; and although both passive *-en* and (active) *-ing* can attach to transitive verbs, the meanings of the derived forms are radically different.

the complementary set is made up by adjectives converted from *-en* and *-ing* adjectives. I have suggested that the rules of this conversion are fundamentally similar to the Hungarian pattern described in section 3.1. Below, I would like to discuss some theoretical issues related to the use of the participles in question in the two languages.

In Laczkó (2000) I propose that *-Ó* participle formation follows the ordinary active pattern: the suffix can attach to all the three major types of verbal predicates (transitives, unaccusatives and unergatives) and it follows the same mapping procedure as the input verbs. By contrast, *-(V)(t)t* is "semi-ergative": as a rule, it maps the [–r] argument of the predicate onto SUBJ: it attaches to unaccusatives, transitives (obligatorily passivizing them) and it cannot combine with unergatives (that is, there is no elsewhere case, that is why the pattern is semi-ergative).

If we take the view I have argued for in this paper for granted as regards the use of the English counterparts of these Hungarian participles, then the following generalizations can be made about the present state of affairs. Active *-ing* participles can be used either predicatively, in combination with *be* in progressive tense forms, or as modifiers within NPs without any restriction. "Passive" *-en* can only attach to transitive predicates and it passivizes them and it can combine with *be* to form passive constructions and perfect *-en*, just like *-ing*, can attach to the three major types of verbs and the output combines with perfective *have*. The crucial point for our present purposes (in the context of the foregoing discussion in general and the Hungarian–English comparison in particular) is that, as a first approximation, we can claim that when *-en* participles are used as modifiers in NPs, they "jointly" follow the semi-ergative pattern manifested by Hungarian *-(V)(t)t*. A very obvious way of capturing this would be the adoption of Bresnan's (1982) rule for participle → adjective conversion, except that here her general conditions on conversion would be applied to describe the rules regulating the use of *-en* participles in the NP domain.[24] Naturally, then we would need another rule for participle → adjective conversion in the same spirit. This is one possible way of accounting for the relevant facts. One of its most salient aspects is that it has to refer to two different *-en* participles: passive *-en* attaching to transitive verbs and perfective *-en* attaching to unaccusatives. Below, I will sketch an alternative by the help of which a single generalization (reference to only one *-en* participle type) can be applied.

On the basis of the historical development of *-en* participles, I would like to speculate that in the Old English and in the Middle English periods even their predicative use in combination with the auxiliaries *have* and *be* showed striking similarity to the present day pattern followed by Hungarian *-(V)(t)t* in NPs described in section 3.1. In OE and ME, after the perfect auxiliary use of *have* and *be* had developed, there was a well-attested, very strong tendency in the division of labour between them.[25] Its essence was as follows: *have* was used in combination with participles derived from transitive and unergative verbs and *be* was combined with unaccusatives. Later on *have* ousted *be* even from the unaccusative domain. In addition, *be* as a passive auxiliary was compatible with participles derived from intransitive verbs. This state of affairs is typically characterized along the following lines in historical and traditional descriptive work.

(40) *have* + V_{tr} *-en*
V_{unerg} *-en*
be (perf) + V_{unacc} *-en*

[24] Of course, Bresnan's (1982) *theme* condition has to be translated into my [–r] condition in the context of LMT.
[25] For discussion and references, see Traugott (1992), Fischer (1992), and Rissanen (1999).

	be (pass) +	V_{tr}	*-en*

That is, one (past) participle is postulated and its combination with the auxiliaries in various functions yields the required perfect or passive meanings. By contrast, Bresnan (1982), for instance, assumes the following system of *-en* participles.

(41)		V_{tr}	$-en_1$
	perf	V_{unerg}	$-en_1$
		V_{unacc}	$-en_1$
	pass	V_{tr}	$-en_2$

If we adopt this fundamental perfect vs. passive distinction and at the same time wish to account for the distributional facts in (40), characterizing the OE and ME periods, then I think the following solution naturally suggests itself.

(42)	*have* +	V_{tr}	$-en_1$
		V_{unerg}	$-en_1$
	be (perf) +	V_{unacc}	$-en_2$
	be (pass) +	V_{tr}	$-en_2$

The basic generalization here is that $-en_2$, just like Hungarian *-(V)(t)t* even today, is passive in an extended sense: it attaches to transitive verbs verbs, obligatorily passivizing them, and it can also combine with unaccusatives. As regards the compatibility of the two auxiliaries with these two participles, the natural conclusion is that *have* is active (V_{tr} and V_{unerg}), while *be*, in its dual function, is passive (V_{tr} and V_{unacc}).

As is well known, after the ME period *have* ousted perfect *be* even from the domain unaccusatives. However, I envisage this as a process resulting in an instance of functional overlap between $-en_1$ and $-en_2$ in the case of unaccusatives.[26] My motivation for this approach is that in this way I can capture the use of *-en* participles in the two distinct domains in present-day English in a simpler and, perhaps, more principled way. Consider:

(43)		V_{tr}	$-en_1$	
	have +	V_{unerg}	$-en_1$	
		V_{unacc}	$-en_1$	
			$-en_2$	modifier of
	be (pass) +	V_{tr}	$-en_2$	NP head

The picture that (43) depicts is as follows. In present-day English, perfect *have* combines with $-en_1$ and it is $-en_2$ participles that can modify noun heads.

4. Concluding remarks

In this paper I have challenged the quite general view that the overwhelming majority of participle-looking words in English NPs (including all the prenominal ones) are adjectives and claimed that they are participles (non-finite verb forms). I have also compared Bresnan's

[26] In Laczkó (2000), I assume exactly the same kind of overlap between active *-Ó* and (extended) passive *-(V)(t)t* in precisely the same unaccusative domain in Hungarian.

(to appear) participle vs. adjective tests for English and Komlósy's (1992) participle vs. adjective diagnostics for Hungarian. In addition, I have argued against the generalized (result) state approach to all prenominal participle-looking words and proposed instead that two separate (sets of) rules are needed: one for the description of the use of *-ing* and *-en* participles in the NP domain and another for participle → adjective conversion.

It is to be noted that in addition to the differences discussed in section 3.1, there are further rather systematic contrasts between English and Hungarian. Compare:

(44) a. *the arrived guests
b. the newly arrived guests

(45) a. *the built house
b. the recently built house

(46) a. *az érkez-ett vendég-ek
the arrive-(V)(t)t guest-pl
'*the arrived guests'

b. a meg-érkez-ett vendég-ek
the PERF-arrive-(V)(t)t guest-pl
'*the arrived guests' cf.
cca. 'the guests who have arrived'

c. az újonnan (meg-)érkez-ett vendég-ek
the newly PERF-arrive-(V)(t)t guest-pl
'the newly arrived guests'

(47) a. *az épít-ett ház
the build-(V)(t)t house
'*the built house' or:
'the house being built'

b. a fel-épít-ett ház
the PERF-build-(V)(t)t house
'*the built house' cf.
cca. 'the house that has been built'

c. a nemrég (fel-)épít-ett ház
the recently PERF-build-(V)(t)t house
'the recently built house'

My explanation for these facts is as follows. As I pointed out in section 3.1, in Hungarian there are a great number of *-(V)(t)t* participles[27] derived from morphologically distinct verb pairs whose members only differ in one respect: whether they contain a perfectivizing preverb or not. The preverbed version is always telic and it can only be used as a participle, see

[27] Recall that *-(V)(t)t* can only attach to unaccusative and transitive verbs, typically passivizing the latter (there is a systematic exception to this passivizing effect, see Laczkó (2000), but this does not concern us here).

(46b,c) and (47b,c). The preverbless version of unaccusatives can only be used in a telic sense and the (passive) preverbless version derived from transitive verbs can be used in either a telic or an atelic sense, and (in its telic sense) it can be input to participle → adjective conversion.[28] It is interesting to note that although the preverbless participle can also have a telic interpretation, it cannot be used on its own (without any modification) in this sense. Compare (46a) and (47a) with (46c) and (47c). My observation is that this is because there is a very strong tendency to interpret bare participle-looking words as adjectives (the output of participle → adjective conversion). The problem with (46a) and (47a) is that in the case of these participles no conversion can take place. Participle → adjective conversion is governed, among other things, by certain pragmatic conditions which have been argued for in English in a detailed fashion by Ackerman – Goldberg (1996). Thus, in Hungarian there are no adjectives like **érkezett* '*arrived' and **épített* '*built'. In addition, I claim that participle → adjective conversion has fundamentally similar aspects in English as well including the tendency to interpret single (unmodified) participle-looking words as adjectives. That is why (44a) and (45a) are ungrammatical. The participle-looking words in them are not interpreted as participles because they stand there by themselves, although their inherently telic interpretation would suffice for a participial use, if that use was available. The difference between English and Hungarian is that in the former there are no pairs like the preverbed and preverbless pairs in the latter.[29] As (46b) and (47b) show, in Hungarian the preverbed version (which is obligatorily taken to be a participle) can be used by itself. In the case of English participle-looking words and in the case of Hungarian preverbless participle-looking words some modification (characteristic of verbal predicates) is required for the participial interpretation to be available. If this condition is satisfied then the relevant forms in both languages can be used as participles, see (44b), (45b), (46c) and (47c).

References

Ackerman, Farrell – Goldberg, Adele E. (1996) Constraints on adjectival past participles, in: Goldberg, Adele E. (ed.) *Conceptual Structure, Discourse and Language*. Stanford Calif.: CSLI Publications, 17–30.

Bresnan, Joan (1982) The passive in lexical theory, in: Bresnan, Joan (ed.) *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: The MIT Press, 3–86.

Bresnan, Joan (to appear) Lexicality and argument structure, in: *Proceedings of the Paris Syntax and Semantics Conference, 1995*.

Cetnarowska, Bozena (to appear) Adjectival past participle formation as an unaccusativity diagnostic in English and Polish, in: *Proceedings of "Vienna 2000", the 9th International Morphology Meeting.*

Fischer, Olga (1992) Syntax, in: Blake, Norman (ed.) *The Cambridge History of the English Language. Volume One. 1066-1476.* Cambridge: Cambridge University Press, 207–408.

Haspelmath, Martin (1994) Passive participles across languages, in: Fox, Barbara – Paul J. Hopper (eds.) *Voice: Form and Function*. Amsterdam: Benjamins, 151–177.

---

[28] See the relevant examples in: (31c,d,f) and (34b,c,d).

[29] Cetnarowska (to appear) reports similar contrasts between Polish adjectival past participles derived from preverbed and preverbless verbs. It is interesting, however, that according to her all these forms in Polish formally behave like adjectives, contra my claim about Hungarian and English.

Károly, Sándor (1956) *Igenévrendszerünk a kódexirodalom elsõ szakaszában [The System of Participles in the First Period of Hungarian Codices]. Nyelvtudományi Értekezések* 10 [*Dissertations in Linguistics* 10]. Budapest: Akadémiai Kiadó.
Kiefer, Ferenc (1992) (ed.) *Strukturális magyar nyelvtan 1. Mondattan [Structural Grammar of Hungarian 1. Syntax]*. Budapest: Akadémiai Kiadó.
Kiefer, Ferenc (2000) *Strukturális magyar nyelvtan 3. Morfológia [Structural Grammar of Hungarian 3. Morphology]*. Budapest: Akadémiai Kiadó.
Kiefer, Ferenc–É. Kiss, Katalin (1994) (eds.) *The Syntactic Structure of Hungarian. Syntax and Semantics* 27. New York: Academic Press.
Komlósy, András (1992) Régensek és vonzatok [Predicates and complements], in: Kiefer (1992:299–527).
Komlósy, András (1994) Complements and adjuncts, in: Kiefer–É. Kiss (1994:91–178).
Laczkó, Tibor (2000) A melléknévi és határozói igenévképzõk [Adjectival and adverbial participial suffixes], in: Kiefer (2000:409–452).
Levin, Beth – Malka Rappaport (1986) The formation of adjectival passives, *Linguistic Inquiry* 17, 623–661.
Rissanen, Matti (1999) Syntax, in: Lass, Roger (ed.) *The Cambridge History of the English Language. Volume Three. 1476-1776.* Cambridge: Cambridge University Press, 187–331.
Traugott, Elizabeth Closs (1992) Syntax, in: Hogg, Richard M. (ed.) *The Cambridge History of the English Language. Volume One. The Beginnings to 1066.* Cambridge: Cambridge University Press, 168–289.

Effect of Head-Marking on Constituent Order in Chichewa

Sam Mchombo
University of California, Berkeley

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)\

2001

CSLI Publications
http://csli-publications.stanford.edu

EFFECTS OF HEAD-MARKING ON CONSTITUENT ORDER IN CHICHEWA[*]

INTRODUCTION

Standard analyses of the morphological structure of the verb in Chichewa, and other Bantu languages have the verb comprising a verb root (VR) to which verbal extensions such as causative, applicative, reciprocal, passive, stative, etc. are suffixed, and to which prefixes are added. The prefixes, analyzed as clitics in some studies, cf. Givón 1971, Mchombo (in press), include elements that encode information pertaining to agreement with the subject and the object, tense/aspect, negation, modality etc. The morphology of suffixation and prefixation (or cliticization) is sufficiently well established as to require no further comment. The following provides a typical example of such morphological organization:

1.	Mkángo	u-da-ómb-án-íts-á	alenje	ndí	asodzi
	3-lion	3SM-pst-hit-recip-caus-fv	2-hunter	and	2-fishermen
	‘The lion made the hunters and the fishermen hit each other.’				

In this example the VR **–omb-** ‘hit’ supports the reciprocal extension **–an-** and the causative extension **–its-,** the final vowel (fv) **–a,** as well as the subject marker **u** and the past tense marker **da**. This paper will focus on the subject and object markers, which appear on the verb head, and the consequence of their appearance on constituent order of the nominal phrases in the sentence.

[*] I am grateful to Andreas Kathol for suggesting the topic and to him, Adams Bodomo, Eyamba Bokamba, Joan Bresnan, Ron Kaplan, Al Mtenje, Steven Schaufele, and participants at the 6th International Lexical Functional Grammar Association conference held at the University of Hong Kong, June 2001, for comments on the material discussed in this paper. Preliminary versions of this material were delivered to audiences at the University of Sonora in Hermosillo, Mexico during a course on Bantu morphosyntax (May 2001) and at a colloquium held at El Colegio de Mexico in Mexico City, D.F. (May 2001). I am grateful to the participants, most especially to Isabel Barreras, Zarina Estrada Fernández, Maria Eugenia Vazquez Laslop, and Paulette Levy, and many others, for comments. I remain responsible for all errors.

HEAD-MARKING IN CHICHEWA

Chichewa shows both subject and object agreement in its verbal morphology. In finite verb forms the S[ubject] M[arker] is obligatory, while the single O[bject] M[arker] is optional. This is exemplified by sentence 2 below:

11 a. Njûchi zi-na-lúm-á alenje
10-bees 10SM-pst-bit-fv 2-hunters
'The bees hit the hunters"

b. Zi-na-lúmá alenje njûchi.
10SM-pst-bit-fv 2-hunters 10-bees
'The bees hit the hunters'

In this sentence the SM **zi** in the verbal morphology contains the ϕ-features of the nominal **njûchi** 'bees.' Note that nominal **njûchi** whose ϕ-features are duplicated by the SM need not appear initially. It can appear in post-verbal position, for as long as it does not disrupt the string adjacency between the verb and its object. The object NP **alenje** 'hunters' must appear in the immediate post-verbal position, adjacent to the verb. The adjacency requirement is demonstrated by the ungrammaticality of sentence 3 below, in which the expression **dzulo** 'yesterday' is inserted between the verb and the object:

3 *Njûchi zi-na-lúm-á dzulo alenje
10-bees 10SM-pst-bite-fv yesterday 2-hunters

This sentence would be well-formed if the expression **dzulo** ‘yesterday’ appeared after the object NP **alenje** ‘hunters.’

The verb can also be marked with an object marker (OM), an element that duplicates the ϕ-feature of the object NP. In the case of the nominal **alenje**, the relevant OM is **wa**. This is shown in sentence 4:

4	a. Njûchi	zi-ná-wá-lum-a	alenje.
	10-bees	10SM-pst-2OM-bite-fv	2-hunters
	‘The bees bit the hunters.’		

The presence of the SM and the OM has consequences on the word order requirements of Chichewa. For a start, the adjacency requirements between the verb and its object NP is relaxed and the ordering of the three major constituents of the sentences becomes free. Thus, the sentences in 4b-e below, showing the different ordering possibilities of the major constituents of the sentence 4a, are grammatical and have the meaning that the bees bit the hunters.

b. SVO Njûchi zi-ná-wá-lum-a alenje

c. OVS Alenje zi-ná-wá-luma njuchi

d. VOS Zi-ná-wá-lum-a alenje njuchi

e. SOV Njuchi alenje zi-ná-wá-lum-a

Further, the presence of the SM and OM license the omission of the actual noun phrases, as shown in (f) below:

f. Zi-ná-wá-lum-a

10SM-pst-2OM-bite-fv

‘They bit them’

In a detailed study of these morphological elements in Chichewa, Bresnan (1985), Bresnan & Mchombo (1986, 1987), and Mchombo (1984) analyzed these markers as incorporated pronominal arguments. The SM is ambiguously used for grammatical and anaphoric agreement whereas the OM is an incorporated pronominal argument. The analysis of the OM as an incorporated pronominal derives further support from the distribution in other Bantu languages such as Kikuyu (Bergvall 1986; Mugane 1997) or Lunda (Kawasha 2001). In those languages the OM is in complementary distribution with an overt nominal functioning as the object of the verb. Based on studies of Setawana, a dialect of Setswana, Demuth & Johnson (1989) analyzed the SM in that language as an incorporated pronominal which functions as the subject of the sentence. This is applicable to Chichewa on the pronominal argument analysis of the SM.

The seminal work of Bresnan & Mchombo focused on the status of the SM and OM, both of which are marked on the verbal head. Their comments on the effect of head-marking on constituent order was confined to the relative freedom of word order that accompanied the presence of the OM. This paper will take those observations a little further. Consider the following:

5	a. Njuchí		izi	zi-ná-lúm-á	álenje
	10-bees		10prox.dem	10SM-pst-bite-fv	2-hunter
	awa	ópúsa			
	2prox.dem	2SM-foolish			

‘These bees bit these foolish hunters.’

b. Mikángó		i-tâtu	i-ná-gúmúl-á
4-lions		4SM-three	4SM-pst-pull down-fv
makólá	ónse	a-nâyi	
6-corrals	all	6SM-four	

‘Three lions puled down all the four corrals.’

In these the subject and object NPs have internal constituents. Ordinarily, those constituents cannot be separated from the head noun. The sentences below are ungrammatical:

6

a. ***Awa**	njuchí	izi	zi-ná-lúmá	**alenje**
2proxdem	10-bees	10proxdem	10SM-pst-bite	2-hunters
ópúsa				
foolish				

b. ***Awa**	ópúsa	njuchí	izi	zi-ná-lúmá
2proxdem	foolish	10-bees	10proxdem	10SM-pst-bite
alenje				
2-hunters				

Without the object marker the object NP cannot be discontinuous. On the other hand, in the presence of the OM not only does the order of the constituent words of the sentences become free but, it seems that the internal components of the NPs equally tolerate discontinuity.

7

a. **Awa**	njuchí	izi	zi-ná-**wá**-lum-a
2-prox-dem	10-bees	10prox.dem	10SM-pst-2OM-bite-fv
alenje	**ópúsa**		
2-hunters	2SM-foolish		

b. **Alenje**	zi-ná-**wá-**lum-a	njuchí	izi
2-hunters	10SM-pst-2OM-bite-fv	10-bees	10-prox.dem
awa	**ópúsa**		
2-prox.dem	2SM-foolish		

c. Izi **awa** **ópúsa** zi-ná-**wá**-lum-a
10-prox.dem 2prox.dem 2sM-foolish 10SM-pst-2OM-bite-fv
alenje njûchi
2-hunters 10-bees
'These bees bit these foolish hunters.'

The data above show the possibility of discontinuity among the internal constituents of the NPs, induced by the presence of both SM and OM being marked on the verbal head. This kind of free word order as well as the possibility of syntactically discontinuous expressions has the hallmarks of non-configurationality. The question thus arises, is Chichewa a non-configurational language?

THE CONFIGURATIONALITY PARAMETER

Studies of Australian languages such as Warlpiri, initiated by Ken Hale, revealed properties about natural language that posed a problem for the conception of grammatical theory proposed within generative grammar. These languages display properties that were dubbed 'non-configurational.' The salient properties of non-configurational languages are:

i. free word order
ii. syntactically discontinuous expressions, and,
iii. null anaphora.

By syntactically discontinuous expressions is meant that non-adjacent nominals may correspond to a single verbal argument; and, null anaphora, according to Hale (1983) is the situation in which an argument (subject, object) is not expressed by an overt nominal in phrase structure.

Hale proposed to deal with the facts about Australian languages by separating lexical structure from phrase structure and positing a configurationality parameter for the application of the projection principle. In configurational languages, such as English, the projection principle was assumed to hold of the pair LS (lexical structure), PS (phrase structure). In non-configurational languages, the projection principle holds of the LS alone.

In further analyses of configurationality, Jelinek posited the pronominal argument hypothesis (PAH) to account for the observed facts. The analysis reduces to the claim that the projection principle uniformly applies to lexical structure cross-linguistically. In configurational languages grammatical relations are marked by the order of the syntactic constituents, which satisfy the argument structure of the predicate. In non-configurational languages, null anaphors satisfy argument structure, hence, grammatical relations may be marked in the morphology as well as in the syntax. The nominal expressions that appear in the sentence are not verbal arguments but freely added expressions that may be referentially linked to pronominal arguments inside the verbal complex.

The pronominal argument hypothesis was reviewed in studies conducted within the framework of LFG (Austin & Bresnan 1996) and rejected as the optimal account for non-configurationality. Austin & Bresnan propose a dual structure hypothesis, where natural language is decomposable into linked parallel informational structures, such as constituent structure (c-s), functional structure (f-s), argument structure (a-s), etc, each of a different formal character. The grammar, within LFG "consists of a set of local, co-descriptive constraints on partial structures. There are no derivational or transformational operations involved: grammatical structures are defined by constraint satisfaction." (Bresnan 1982).

Further, the parallel structures of LFG model different facets of the structure of language. The constituent structure, also called categorial structure, models the overt structure of forms of expression, encoding such surface structure relations as precedence and dominance, while the functional structure models the grammatical relations among syntactic functions. Argument structure deals with the grammatically expressible participants of eventualities (cf. Austin & Bresnan 1996). The various structures are associated by principles of functional correspondence, sometimes called linking or mapping principles.

The idea of a configurationality parameter is also rejected by Nordlinger (1998). Focusing on constructive case within Australian languages Nordlinger claims that "…there is no parametric distinction between configurationality and nonconfigurationality. Rather, these two language types merely represent the extremes of a continuum: languages may identify grammatical relations in the syntax (fully configurational), in themorphology (fully nonconfigurational), or (more usually) by some mixture of the two." (Nordlinger 1998:26). It seems that Chichewa lends some credence to this claim by displaying characteristics that place it somewhere between the extreme positions.

CHICHEWA VERBAL MORPHOLOGY AND THE PRONOMINAL ARGUMENT HYPOTHESIS.

The configurationality parameter of Hale rested on the separation of lexical structure from phrase structure and determination of whether the projection principle holds of both structures. By lexical structure Hale refers to predicates and their argument arrays which correspond to variables specified in the dictionary definition of a verb. Jelinek's pronominal argument hypothesis derives from the claim that the argument variables are satisfied by pronominal arguments, which in the Australian languages may lack phonological realization. This account appears to be readily applicable to Bantu languages. Thus, consider sentence 4a again, repeated below as 8:

8 Njûchi zi-ná-wá-luma alenje
'The bees bit the hunters'

It could be argued that the SM and OM are the pronominal arguments and the NPs **njûchi** 'bees' and **alenje** 'hunters' are not arguments of the verb. This is technically true. The fact that they can be omitted without inducing ungrammaticality testifies to that. Further, the complementarity between the OM and an overt nominal argument in some languages such as Gikuyu demonstrated the argument status of the OM. How then do the word order facts get and the discontinuous constituency of the nominal expressions get derived? Jelinek's suggestion was to appeal to referential linking, treating the constituents of the NPs as nominals that get linked to the incorporated arguments. This approach is rejected by Reinholtz, discussing comparable facts in Swampy Cree, on the grounds that "…the formation of discontinuous constituents does not depend on 'referential linking' and that it has the all the hallmarks of wh-movement in so called configurational languages" (Reinholtz 1999).

The determination of grammatical relations in Chichewa definitely appears to involve a mixture of syntactic order and morphological marking. The SM and OM as pronominal arguments marked on the verbal head are strictly ordered. The determination of grammatical relations based on their distribution involves ordering which may be linked to dominance relations. The OM is always sister to the verb stem, indicative of its status as an internal argument. It is in complementary distribution with an overt nominal, even in Chichewa where *inter alia* phonological cues such as tonal patterning, mark the overt NP as outside the VP configuration when the OM is present (see Bresnan & Mchombo 1987, Bresnan & Kanerva 1989). In brief, Chichewa has a VP configuration. It is the requirements of the VP that mandate the ordering relation between the verb and its internal argument. What is clearly the case is that the when the verb has a nominal phrase as its argument, the constituent integrity of the NP must be respected and retained, relaxed only when the OM is the argument. Why?

DISCOURSE FUNCTIONS, GRAMMATICAL FUNCTIONS, AND CONSTITUENT ORDER

Head-marking in Chichewa, just like pronominal argument marking, in Australian languages, has the result that the pronominal arguments bear the syntactic functions of subject and object. The nominal expressions that may appear are indeed extra-sentential and have discourse functions such as Topic. This is certainly the case with the NP that is linked to the OM. Such discourse elements have to be linked to elements bearing grammatical functions for the expression to be well-formed. This is where the theory of LFG provides the architecture for an elegant account of the linguistic facts. The theory posits constraints such as the completeness condition, coherence condition, extended coherence condition, which need to be satisfied. In the f-s grammaticized discourse functions such as Topic (TOP) and Focus (FOC) will be included, with their linkage to the syntactic functions either through functional or anaphoric binding resolved. For instance the f-s of sentence 8 above may be represented as follows:

S -> NP, NP, VP
VP -> SM T/AOM-VS

Besides the anaphoric agreement that holds between the pronominal arguments and the NPs in the discourse structure, Bantu languages also display agreement between a nominal head and its modifiers. Consider the NPs **alenje awa ópúsa** 'these foolish hunters' and **njuchí izi zópúsa** 'these foolish bees.' In both, the form of the demonstrative 'these' is different, reflecting agreement with the head noun in number and gender (nominal class features). The same applies to the form of word 'foolish', which is derived from the verb **–pusa** 'be foolish' through the prefixation of the associative marker **–a** to which the class concord of the noun is applied. This yields the different forms of the word. Such agreement will be reflected in the f-s representation of the NP, where the constituents of the NP will have the same values for the specification of the gender and number attributes. The disparate bits of information associated with the constituents of the NPs are unified within the f-s which is where the anaphoric relation between the TOP nominal and the pronominal argument bearing the syntactic function is specified, satisfying the extended coherence condition. This is the condition that requires that all syntactic functions (including adjuncts and grammaticalized discourse functions) must be appropriately integrated into the f-structure (cf. Farsi-Fehri 1984; Nordlinger 1998). As such, the constituents of the NP do not have to be 'referentially linked' to the pronominal argument separately. The discontinuity of the constituents is possible in the c-s "[b]ecause of the many-one correspondence between the parallel c- and f-structures in violation of the projection principle, true discontinuous constituents are allowed. The principle of functional uniqueness, together with the free association of functions with constituents of the non-projective S node, can create a single functional constituent in f-structure corresponding to a 'scattered' set of c-structure nodes. Such a constituent will have the 'merged' interpretation. (Austin & Bresnan 1996:237).

THE LIMITS OF DISCONTINUITY

Although there is a degree of discontinuity among the constituents of the nominal expressions when the verbal head is marked with the subject and object pronominal arguments, there are limits on the possibilities of such discontinuity. Consider the following sentence:

9	Mkángo	u-méné	ú-ma-saká	mbûzi
	3-lion	3SM-relpro	3SM-hab-hunt	10SM-goats

ú-ma-wa-sautsa	alenje	a-méné	á-ma-gwetsá
3SM-hab-2OM-bother	2-hunters	2SM-relpro	2SM-hab-fell

mitêngo
4-trees
'The lion which hunts goats bothers the hunters who fell trees'

The formation of relative clauses in Chichewa is, largely, comparable to that of English. Chichewa is a head-initial language and, within relative clauses, the realtive clause follows the head noun. There is a relative marker **–mene** which is marked for agreement with the head noun and introduces the relative clause. If the relativized Nominal is the object of the verb, the OM is optionally absent, but normally present, functioning as a resumptive pronoun. The presence of the relative marker **–mene** also has the phonological effect of marking the verb within the relative construction with a high tone. Consider the following:

10	a. mkángo	u-ku-sáká	mbûzi
	3-lion	3SM-pres-hunt	10-goats

'The lion is hunting goats'

b. mkángó	u-méné	ú-kú-sáká	mbûzi
3-lion	3SM-relpro	3SM-pres-hunt	10-goats

‘The lion which is hunting goats.’

In these examples the tone patterns on the verb **ukusaka** ‘it is hunting’ are different, in part because of the presence of the relative marker in (b). The fact that the tone marking functions as a phonological cue of the different constructions makes it possible for the relative marker **u-mene** to be deleted or dropped. Thus construction 10c below, which only differs tonally from sentence 10a above, is still construed as a relativized NP configuration:

c. mkángó	ú-kú-sáká	mbûzi
3-lion	3SM-pres-hunt	10-goats

‘The lion which is hunting goats.’

Returning to sentence 9 above, the verb **sautsa** ‘bother, trouble’ is marked with the SM **u** agreeing with **mkángo** ‘lion’ and the OM **wa,** agreeing with **alenje** ‘hunters.’. While the order of the nominal expressions **mkango u-mene u-ma-saka mbuzi** ‘the lion which hunts goats’ and **alenje a-mene a-ma-gwetsa mitengo** ‘the hunters who fell the trees’ is free, the constituents of those nominal expressions cannot be discontinuous. The following sentence is, at best, questionable. :

11	?**alenje**	mkángo	ú-ma-**wa-**sautsa		**a-méné**
	2-hunters	3-lion	3SM-hab-2OM bother		2SM-relpro
	á-ma-gwetsá	**miténgo**	u-méné	ú-ma-saká	mbûzi
	2SM-hab-fell	4-trees	3SM-rekpro	3SM-hab-hunt	10-goats

‘The hunters the lion bother them, who fell trees, that hunts goats.’

The nominal expressions are both relativized NPs. The possibility of extraposing the relative clause does not completely rescue the sentence from ungrammaticality. The object NPs within those relative clauses cannot be moved out largely because the verbal head is itself not marked with the OM. Now consider the following where the verbal heads within the relative clauses are marked with the OMs:

12 Mkángo uméné úma**zi**saka **mbûzi** úmawasautsa alenje améné amaigwetsa miténgo

‘The lion which hunts the goats bothers the hunters that fell the trees’

This sentence does not allow for the range of possible word orders that are normally associated with head-marking in Chichewa. Thus,

13 a*Mkángo mbûzi alenje miténgo úmawasautsa umêné umazisaka améné ámaigwetsa

The lion, the goats, the hunters, the trees, it bothers them which hunts them (goats) who fell them (trees)

b. *Mbûzi mkángo uméné úmazisaka miténgo alenje améné amaigwetsa úmawasautsa

The goats the lion which hunts them the trees the hunters who fell them(the trees) it (the lion) bothers them (hunters)

Part of the problem associated with 11b is that it seems to have mbuzi ‘goats’ as its topic yet the sentence has to do with the fact that the lion which hunts the goats bothers the hunters who fell trees. In brief, when the nominal expressions have relative clauses, discontinuity among the constituents is more difficult, certainly reducible to island effects.

In her study of discontinuous constituents in Swampy Cree Reinholtz (1999) proposes to treat discontinuous constituents as the output of a Focus mechanism, which “…picks out a nominal modifier belonging to a larger NP, and places this in a preverbal Focus position where it is separated from the noun it qualifies by other material” (Reinholtz 1999: 28). She proceeds to adopt the analysis of Swampy Cree proposed in Russell & Reinholtz (1996) that discontinuous NPs involve movement, which moves the movement of a nominal modifier out of a containing NP to a pre-verbal focus position. This movement-based approach leads Reinholtz to assimilate such discontinuous constituents to being an instance of Wh-movement. She notes that movement to Focus position is commonly grouped with wh-movement and the discontinuous constituents in Swampy Cree share common characteristics with wh-movement. Thus, they show the ability to span several clauses, cannot move material out of adverbial constituents, and have limited application in relative clauses or embedded questions.

It is an open question as to whether discontinuous constituents can be assimilated to wh-movement. Certainly failure to affect relative clauses can be handled in terms of island effects. Even if an analysis along the lines of wh-movement were to be adopted, within LFG there are no movement rules. The resolution of wh-effects can be handled in terms of functional uncertainty. The island effects associated with relative clauses are equally amenable to analysis in terms of functional uncertainty. Significantly, it is not the case that movement needs to be invoked since the domain where functional uncertainty applies is the f-s. The c-s to f-s mapping algorithm allows for the possible scattering in c-s of information which, in the f-s, is unified. Limits on discontinuity are the result of satisfying constraints on partial structures of parallel informational structures.

CONCLUSION

Head-marking in Chichewa results in discontinuous constituency of the nominal expressions construed as dependent on the incorporated pronominal arguments. While this may appear to accord languages like Chichewa the appearance of being non-configurational, the language seems to have a VP. The incorporated pronominal arguments also have a fairly rigid order, characteristic of clitics in Bantu languages in general. In this regard, the clitics differ from the verbal extensions or suffixes, which affect argument structure and are subject to an array of linguistic processes excluded from the pre-VS clitic domain (cf. Mchombo, in press). The verbal suffixes are subject to variable order, within limits. The rigidity of the clitics captures c-command relation (borrowing the terminology of the principles and parameters theory) between the incorporated pronominal arguments. Thus Chichewa offers an example of a language that is configurational yet allows for discontinuous constituents under the specific conditions of head-marking, comparable to non-configurational languages. This suggests the independence of discontinuity of constituents from non-configurationality, as noted by

Austin & Bresnan, as well as Nordlinger. It still remains to provide an exact specification of this and of the limits on discontinuity in Chichewa within the theory of LFG.

REFERENCES.

Austin, Peter & Joan Bresnan, 1996. Non-Configurationality in Australian Aboriginal Languages. *Natural Language & Linguistic Theory 14*.. 215-268.

Baker, Mark C. 1990. Pronominal Inflection and the Morphology-Syntax Interface. CLS-26. Chicago Linguistic Society, Chiacgo, Illinois. 25-48.

Bergvall, Victoria L. 1986. A Tyopology of Empty Categories for Kikuyu and Swahili. In Dimmendaal, Gerrit J. (ed.) 1986. *Current Approaches to African Linguistics Vol. 3 No. 6*. Dordrecht, Holland: Foris publications

Bresnan, Joan 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: The MIT Press

Bresnan, Joan 1994. Locative Inversion and the Architecture of Universal Grammar. *Language 70*, 72-131.

Bresnan, Joan & Jonni Kanerva 1989: Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry 20*:1-50

Bresnan, Joan & Sam Mchombo 1986. Grammatical and anaphoric agreement. Papers from the Parasession on Pragmatics and Grammatical Theory. Chicago Linguistic Society 22:2, 278-97. Chicago.

Bresnan, Joan & Sam Mchombo 1987. Topic, pronoun and agreement in Chichewa. *Language 63.4*. 741-82

Bresnan, Joan & Sam Mchombo 1995. The Lexical Integrity Principle: Evidene from Bantu. *Natural Language & Linguistic Theory 13*: 181-254

Demuth, Katherine & Mark Johnson 1989. Interaction between Discourse Functions and Agreement in Setawana. *Journal of African Languages and Linguistics.11*: 21-35.

Farsi-Fehri, Abdelkader, 1984. Agreement in Arabic, Binding and Coherence. In Barlow. Michael. & Charles A. Ferguson (eds.) 1984. *Agreement in Natural Language*. Stanford: CA: CSLI Publications. 107-158.

Givón, Talmy. 1971. On the verbal origins of the Bantu verbal suffixes. *Studies in African Linguistics 2.2* 145-63. .

Givón, Talmy 1976. Topic, pronoun, and grammatical agreement. In Li, Charles (ed) 1976. *Subject and topic*. New York: Academic Press 149-88

Hale, Ken 1983. Warlpiri and the Grammar of Non-configurational Languages. *Natural Language & Lnguistic Theory 1*. 5-47.

Jelinek, Eloise 1984. Empty Categories, Case, and Configurationality. *Natural Language & Linguistic Theory 2*. 39-76.

Kathol, Andreas & Richard Rhodes.1999. Constituency and Linearization of Ojibwe Nominals. In *Proceedings of WSCLA 4*. UBC Working Papers in Linguistics II. Vancouver: Dept. of Linguistics, UBC. 75–91.

Kawasha, Boniface 2001. Relativization and Grammatical Relations in Lunda. Unpublished paper, University of Oregon.

Matthewson, Lisa & Charlotte Reinholtz 1996. The Syntax and Semantics of Determiners: A comparison of Salish and Cree. In Davis, Henry (ed.) 1996. *Papers from the 31st International Conference on Salish and Neighbouring Languages*. Vancouver: University of British Columbia 211-237

Mchombo, Sam. 1984. The non-existence of verb-object agreement in Bantu. Unpublished paper. MIT.

Mchombo, Sam in press. Affixes, clitics, and Bantu morphosyntax. *Perspectives on Cognitive Science 3: Language Universals and Variation.* Ablex Publishing.

Mugane, John 1997. *A Paradigmatic Grammar of Gikuyu*. Stanford, CA. CSLI Publications (Stanford Monographs in African Languages)

Nordlinger, Rachel1998. *Constructive Case. Evidence from Australian Languages*. Stanford, CA: CSLI Publications

Reinhart, Tanya 1981. Pragmatics and Linguistics: An analysis of sentence topics. *Philosophica 27*. 53-94

Reinholtz, Charlotte 1999. On the Characterization of Discontinuous Constituents: Evidence from Swampy Cree. *International Journal of American Linguistics.*

Russell, Kevin, and Charlotte Reinholtz. 1996. Hierarchical Structure in a Non-configurational Language: Asymmetries in Swampy Cree. *Proceedings of the Fourteenth West Coast Conference in Formal Linguistics (WCCFL 14)*, 431-445.

Deriving the Directionality Parameter in OT-LFG*

Yukiko Morimoto
University of Düsseldorf

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu

*I am grateful to Peter Sells for helpful discussion on this work, and to the audience of LFG-01 conference for useful comments and questions. I am solely responsible for any errors and misrepresentations.

1 Introduction

Recent work in Optimality Theoretic Syntax (Samek-Lodovici 1996, Grimshaw 1997, Costa 1998, Sells 2001) successfully models positioning of syntactic elements that target a privileged position in a clause, such as topics, sentential adverbs, operators, and core syntactic dependents like the subject, by extending the mechanism of Generalized Alignment developed in OT phonology (McCarthy and Prince 1993) to the domain of clausal syntax. By allowing only left-alignment, Sells' (1999a,b, 2001) work takes a pioneering step towards recasting Kayne's (1994) insightful observation that phrase structure is fundamentally antisymmetric: there is a universal preference for the left-edge of the clause, so that the (unmarked) structure is predominantly right-branching.

Restricting alignment to the left-edge of a clause as in earlier proposals imposes serious limitations on the kinds of syntactic positioning and constructions we can explain within the constraint-based, output-oriented model of grammar. For example, there is a major class of constituents that exhibit affinity with the head (which I call 'head-attracted' constituents as opposed to 'edge-attracted' constituents like topics, operators, etc. mentioned above), and structural position of such constituents is dependent on the directionality of heads: head-initial languages typically place the head of the relative clause before the modifying clauses, negation before the verb, and objects and focus after the verb; head-final languages, on the other hand, typically place the head of the relative clauses after the modifying clause, negation after the verb, and object and focus before the verb. In some of these cases, the opposite patterns are quite rare or non-existent. These predictable patterns of syntactic positioning therefore deserve a principled explanation.

Focusing on a typology of sentential negation, my goal in this paper is to show how the distribution of head-attracted constituents like negation can be explained by a set of well-grounded universal constraints independently proposed to derive the typology of clause structure and the head directionality parameter.

2 Expression of Sentential Negation

According to Dahl (1979), negation is most commonly expressed in one of three forms. First, negation can be expressed as an inflectional morpheme, as illustrated in (1)–(3). (1) and (2) are examples from SOV languages, and the negative morpheme is expressed as a suffix. (3) is an SVO language, and the negative morpheme is prefixal.

(1) Turkish (SOV): suffix

John	elmalar-i	ser-**me**-di-∅
John	apples-ACC	like-NEG-PAST3SG

'John didn't like apples.'

(2) Japanese (SOV): suffix

Taroo-wa	asagohan-o	tabe-**na**-katta.
Taroo-TOP	breakfast-ACC	eat-NEG-PAST

'Taro didn't eat breakfast.'

(3) Bantu—Kinyarwanda (SVO): prefix

Sií-n-zi igihe á-záa-garuk-(ir)-a.
NEG-I-know time he-FUT-REL-return-APPL-ASP
'I don't know when he will return.' (Kimenyi 1980:69)

Second, negation can also be expressed as an invariant (morphologically uninflected) negative particle, as illustrated in (4)–(8) from a variety of languages that represent different word order types.

(4) English (SVO): negative particle

John does **not** speak German.

(5) French (SVO): negative particle

Jean (n') est **pas** venu.
Jean *ne* is not come
'Jean did not come.' (Haegeman 1995:229)

(6) German (SOV): negative particle

Warum gefiel unsere Lösung dem Peter **nicht**?
why pleased our solution Peter not
'Why did our solution not please Peter?' (Haegeman 1995:168)

(7) Swedish (SVO): negative particle

Jan köpte **inte** boken.
Jan bought NEG books
'John didn't buy books.' (Holmberg and Platzack 1988)

(8) Malagasy (VOS): negative particle

Tsy manasa lamba mihitsy ve Rakoto?
NEG wash clothes at.all Q Rakoto
'Does Rakoto not wash clothes at all?' (Rackowski and Travis 2000)

Another way to express negation is to use a negative auxiliary as shown in (9)–(10). Korean shows a postverbal negative auxiliary, and Finnish, preverbal. In these languages, the negative auxiliary and the main verb are morphologically independent words, forming a verbal complex.

(9) Korean (SOV): negative auxiliary

Na-nun phyenci-lul ssu-ci **anh-ass-ta**.
I-TOP letter-ACC write-COMP NEG-PAST-DECL
'I didn't write a letter.' (Kim 2000:2)

(10) Finnish (SOV): negative auxiliary

Minä **e-n** puhu-isi
I-NOM NEG-1SG speak-COND
'I would not speak.' (Mitchell 1991)

Several observations regarding both the form and placement of negation are noted. First, Dahl (1979) observes that negation is often attracted to the verb. Moreover, as illustrated in the Malagasy example in (8), verb-initial languages has a strong tendency to place NEG preverbally in sentence-initial position (Dahl 1979, Payne 1985, Dryer 1988).

Second, Dahl also notes that negative morphemes appearing immediately after the main verb (either as a suffix to the verb stem or as an independent word) are typically inflected (e.g. for tense, mood). This is illustrated by the examples from Turkish and Japanese in (1)–(2). On the other hand, negative morphemes appearing in (immediately) preverbal position tend to be uninflected, as exemplified by the variety of languages in (4)–(8).

Thirdly, Dryer (1988) observes that SVO languages most commonly (47 out of 67 languages in Dryer's typological survey) place NEG preverbally. On the other hand, SOV languages commonly place NEG postverbally (64 out of 117 languages in Dryer (1998)). It is also noted, however, that there are also a fair number of SOV languages (39 out of 117) in which negation is expressed preverbally. Hindi and Korean in (11) and (12), whose basic word order is SOV, exemplify this last pattern. In short, the placement of negation tends to be quite uniform among verb-initial languages, but among SVO and SOV languages, it is much more variable.

(11) Hindi (SOV) preverbal negation

anjum haar **nahĩĩ** banaa rah-ii hai
Anjum.F.NOM necklace.M.NOM NEG make STAT-PERF.F.SG is
'Anjum is not making a necklace.' (Butt 1994:79)

(12) Korean (SOV) preverbal negation

John-un ppang-ul **an** mek-ess-ta.
John-TOP bread-ACC NEG eat-PAST-DECL
'John didn't eat the bread.' (Kim 2000:1)

Of course not only is the exact structural position of negation variable in these language types, but forms of negation also vary (elaborated below). An obvious question is whether the cross-linguistic variation in form and position of negation can be systematically derived by a set of universal constraints on clause structure, by relating the distribution of negation with other elements in the clause. In what follows, I argue that it can and should be, and show how the question can be answered straightforwardly from the OT-LFG perspective.

The above data and the typological generalizations noted above are summarized in the table in (13). Going from the left to right in the columns, we have the languages considered in the analysis to follow, which presumably represent predominant language types with respect to the syntax of negation, syntactic category of the negative element, its position, the position of the main verb, and presence/absence of a functional category.

(13)

	Languages	NEG category	NEG position	V pos.	F?
I.	Japanese (SOV)	suffix	V^0	V	no
II.	Bantu (SVO)	prefix	V^0	V	C
III.	Korean *an* (SOV)	NEG^0	left-adjoins to V^0	V	no
IV.	Korean *anh-ta* (SOV)	V^0 (postverbal)	form V^0 w/ LexV^0	V	no
V.	English (SVO)	particle	IP domain	V	I,C
VI.	German (SOV)	particle	VP domain	V2	C
VII.	Swedish (SVO)	particle	IP domain	V2	I,C

The table indicates that one language type represented by Japanese and Turkish, referred to as Type I here, has a negative suffix morphologically attached to the verb stem. Importantly, this language type lacks any evidence of a functional category. Cho and Sells (1995) and Sells (1995), for example, present morphophonological and morphosyntactic evidence for the lack of functional categories (I, C, or D) in Japanese and Korean. The form of negation and presence/absence of a functional category exhibit an interesting correlation, and that will become clearer through the analysis.

The second language type, referred to as Type II hereafter, is represented by Bantu languages like Kinyarwanda illustrated earlier in (3). This language type has a negative prefix bound to V^0. The verb is in V in most clause types.[1] As for the presence/absence of functional categories, Bresnan and Mchombo (1987), for example, argue convincingly that the Bantu language Chicheŵa lacks evidence for I^0. Other Bantu languages such as Kirundi, Kinyarwanda, and Swahili also show no evidence for I (cf. Morimoto 2000).

As illustrated earlier Korean displays two types of negation. One type (Type III in the table) is the preverbal negative particle *an* (cf. (12)), and the other, the postverbal negative auxiliary (cf. (9)). The preverbal particle *an* is analyzed as Neg^0 which is adjoined to V^0 to form a verbal complex V^0, as shown in (14a); the postverbal negative auxiliary (V_{neg}) forms a V^0 with the main verb (V_{lex}), as shown in (14b) (Sells 1999a). Like Japanese, Korean lacks any evidence of functional categories I, C, and D (cf. Cho and Sells 1995, Sells 1995, Choi 1999).

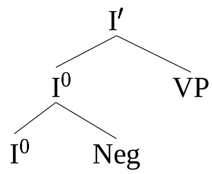
(14) a. V^0 [Neg^0 V^0]

b. V^0 [V_{lex} V_{neg}]

English, German, and Swedish, referred to in the table as Types V, VI, and VII respectively, have a morphologically invariant negative particle, but each exhibits a different pattern in terms of the positioning of the negative element. In English, sentential negation is in the domain of IP. Bresnan (2001b), for example, analyzes the English negative particle to be adjoined to I^0, as shown in (15).

[1] It is argued that in a subset of Bantu languages such as Dzamba, Swahili, and Kilega that display so-called "subject inversion" in relative clauses, the verb in the relative clause appears in C. See Morimoto (2001) and references on earlier work cited therein.

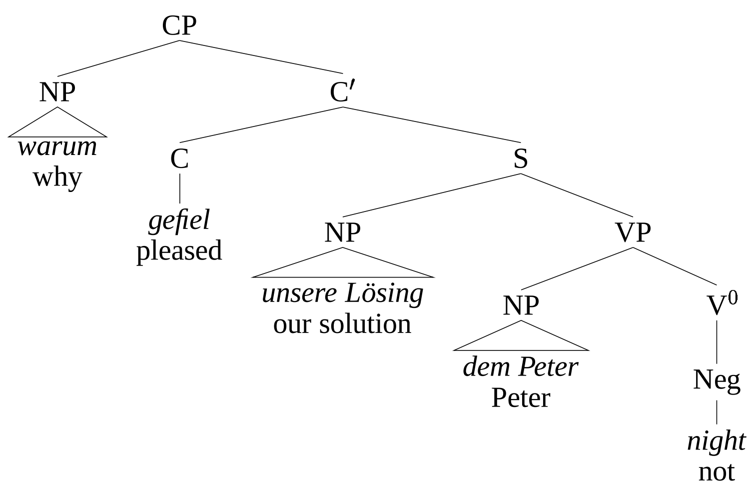
(15) a. Type V. English (Bresnan 2001b)



In German, sentential negation is apparently in the domain of VP. For example, the earlier example in (6), repeated here in (16a), can be assigned a structure like that in (16b). As argued most recently by Choi (1999) and Berman (2000), in German, phrasal elements in sentences like that in (6) are organized within the exocentric category S dominated by CP. The finite verb occupies C satisfying the verb-second requirement. Under CP we have a head-final structure: the subject is under S, and the object inside VP. Given the basic clause structure in German established in the LFG earlier work above, the negative particle naturally falls within VP, presumably adjoined to V^0 as shown.

(16) a. Warum gefiel unsere Lösung dem Peter **nicht**?
why pleased our solution Peter not
'Why did our solution not please Peter?' (Haegeman 1995:168)

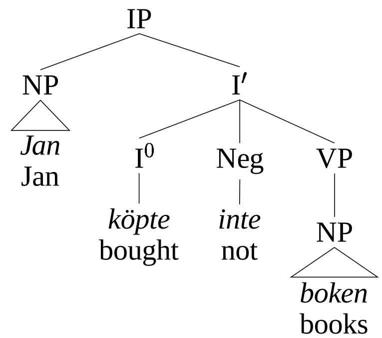
b.



In Swedish, negation is argued to be in the IP domain. Sells (2000) presents empirical evidence specifically showing that Swedish negation cannot be inside VP. The earlier example in (7), repeated here in (17a), thus can be assigned a structure like that in (17b). Swedish exhibits verb-second (V2) in main clauses, so the finite verb in (17a) occupies I, and the subject is in SpecIP.[2]

[2]English negation could perhaps be analyzed as having a structure like that in Swedish in (17b). For simplicity, in my analysis I will treat English and Swedish negation as having the structure in (17b).

(17) a. Jan köpte **inte** boken.
Jan bought NEG books
'John didn't buy books.'

b. 

An additional example from Swedish in (18) more clearly shows that negation can be placed anywhere within the IP domain but cannot be within VP, as pointed out by Sells (2000, ex.(6b)).

(18)	att	(inte)	Johan	(inte)	[$_{vp}$	gillar	(*inte)	prinsesstårta]	
	that	(not)	Johan	(not)	[$_{vp}$	likes	(*not)	princess cake]	(Holmberg 1993)

Based on the distribution of the negative particle in these languages, we might state more generally that languages choose between the domain of a lexical projection (VP) or that of a functional projection (IP) for the position of negation. In what follows, I provide an OT-LFG analysis to suggest how the cross-linguistic inventories in the form and position of negation can be viewed as a result of the systematic interaction of constraints on clausal skeleton, head-positioning, and those on structural economy.

3 Clausal Syntax in OT-LFG

Due to space limitations, I will omit much of the discussion on the nature of the INPUT and GEN by now standardly assumed in OT-LFG work. For detailed discussion, I refer the reader to Kuhn (2001) and Sells (2001a,b). In the present discussion, I concentrate on motivating the crucial constraints and deriving the typology of negation discussed in the previous section.

Particularly crucial to the analysis is the use of the constraint family of Generalized Alignment (McCarthy and Prince 1993). The core idea of Generalized Alignment has been extended to the domain of syntax to account for placement of clausal elements (e.g. heads and complements). Furthermore it has been proposed that there is preference for clausal elements to align to the left edge, thereby yielding clause structure that is predominantly right-branching rather than left-branching (cf. Sells 2001b, Grimshaw 2001).

This "leftness" tendency is essentially the observation that led Kayne (1994) to develop his theory of the antisymmetry of phrase structure. Kayne's key proposal that an asymmetric dominance relation (i.e. c-command) invariably maps onto linear precedence (referred to as the Linear Correspondence Axiom) makes a number of important typological predictions. In order for the theory to work, however, Kayne must treat a set of universal tendencies as absolute universals, or hard constraints, and derive more marked or non-canonical structures by movement of relevant syntactic elements. Motivating a series of movement operations required by the theory is often difficult, if not empirically unfounded.

Recent developments in OT syntax using the formal mechanism of Generalized Alignment, particularly the work of Sells (2001b) and Grimshaw (2001), are attempts to recast Kayne's insights from the non-derivational perspective. OT's capacity to explain "soft" generalizations like linguistic universals by exploiting violability of constraints is well-supported in other areas of grammar (e.g. Sells (2001a) and earlier work cited therein), and has also been shown to be a promising framework for explaining phrase structure universals and cross-linguistic variation. The analysis presented in this section builds on the earlier OT-LFG approach to clause structure and further refines the model by proposing a additional formal mechanism that effectively derives directionality parameters.

Clause Structure Constraints

Two alignment constraints have been proposed on clausal skeleton and head directionality (taken from Sells 2001b): left-alignment of heads ('Head-Left') is applied within an immediate projection of the head. Importantly, there is no Head-Right constraint. This essentially recaptures Kayne's proposal about the universal order of head-complement structure. Here, this is a soft constraint on representation, and obviously it is violated more severely by head-final structure. The only right-alignment constraint admitted in the present approach is Spine-R(ight): this requires that any head (e.g. V) and its extended heads (V′, VP, I, I′, IP), or co-heads (= spine), be final in each of their local subtree (hence 'Spine-Right'), preferring a fully right-headed structure.

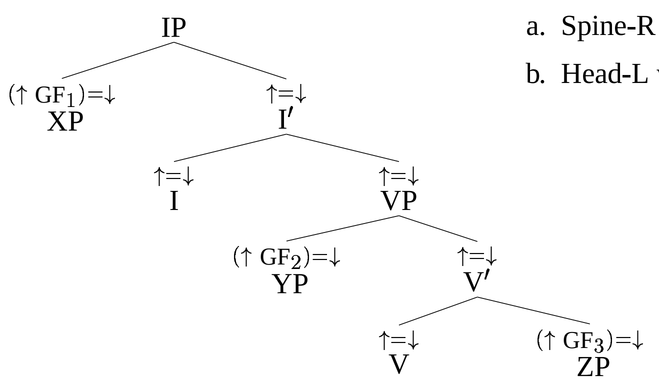
(19) Head positioning constraints

a. Head-L(eft): X^0 is left in its immediate constituent.

b. Spine-R(ight): A co-head aligns right in its immediate constituent.

To illustrate how these constraints figure in OT-LFG grammar, let us consider a schematic SIVO structure in (20).[3]

(20) SIVO structure



a. Spine-R violations: 2 (I and V).

b. Head-L violation: 0.

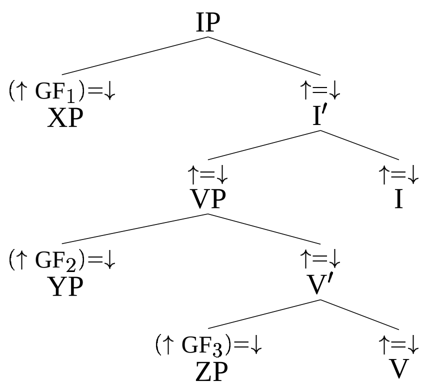
Starting from the bottom, within the immediate projection of V′, there is 1 Spine-R violation at V, as this node, annotated ↑ = ↓ is a co-head but does not align right within its immediate projection. V is also a c-structure X^0 head, and aligns left in its local subtree, so this respects the Head-L constraint. Second, going up one level in the structure within the subtree dominated by VP, there is no violation of Spine-R or Head-L: the co-head V′ aligns right (respecting Spine-R), and there is no X^0 at this level. Within the next higher subtree, I and VP are co-heads. The configuration given here thus violates Spine-R once for I but there is no Head-L violation. Note in passing that if the head is rightmost, and the VP co-head is leftmost in the structure, then this configuration would violate both Spine-R (degree of 1) and Head-L.

[3]The structure in (20) and the subsequent trees representing different word order types are originally discussed in Sells (2001).

Consequently, having functional projections is more preferable in head-initial languages than in head-final languages. In sum, in this structure, there are 2 Spine-R violations and no Head-L violation, as indicated in (20). Thus, in order to yield this structure as optimal, Spine-R must be ranked lower than Head-L.

Sells' proposal on the derivation of a clausal skeleton makes a number of important typological predictions, particularly when we contrast head-initial and head-final structures. For example, let us assume the basic clause structure for a head-final language to be the mirror-image of the SIVO structure in (20) with respect to head positioning:

(21) SOVI structure



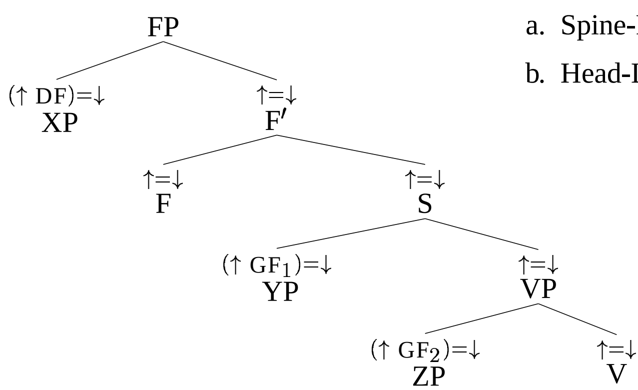
a. Spine-R violation: 1 (VP).

b. Head-L violations: 2 (I and V).

The structure in (21) violates Spine-R once at VP (a co-head) and Head-L twice, at I and V. More generally, whenever the structure contains a functional head (F^0), which is also a co-head (annotated $\uparrow = \downarrow$), the sister node will always be a co-head also; therefore whether F^0 is left or right, this local structure will violate Spine-R once. Having F^0 in a head-final structure also means Head-L violations will be more severe.

Is there a more optimal head-final structure than (21) that GEN can generate? Note that in addition to the endocentric XPs, OT-LFG also allows the exocentric category S (Bresnan 2001b). If we posit a structure like that given in (22) where there is a single functional head above S, violation of high-ranking Spine-R is fewer. This structure is observed in German, where SpecFP is reserved for a discourse function (e.g. TOPIC).

(22) SIOV structure (e.g. German nonsubject-initial clause)



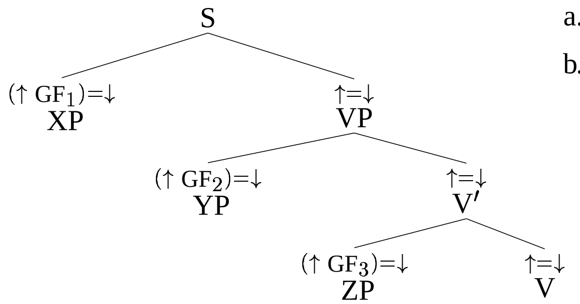
a. Spine-R violation: 1 (I).

b. Head-L violation: 1 (V).

Importantly, unless the presence of a functional projection is required (e.g. by V2, as observed in

German), the most optimal structure in head-final languages would be one in which the root node has the exocentric organization (S) and has no functional head above it, as shown in (23). In this structure, each co-head (ones annotated ↑ = ↓) aligns right in its local subtree. The degree of Spine-R violation is therefore 0. The structure has one lexical head (V) at the bottom of the structure, aligning right. This incurs 1 violation of Head-L.

(23) SOV structure



a. Spine-R violation: 0.

b. Head-L violation: 1 (V).

Indeed as briefly noted earlier, it is generally the case that some head-final languages lack lexically independent functional heads. That is, tense/aspect and complementizer are typically morphological affixes on the verb; see for example, Cho and Sells (1995), Sells (1995) for Korean and Japanese. Kayne's theory of the antisymmetry of phrase structure predicts that there is no V-I-C rightward movement. Instead, V universally moves to I and to C leftward; in head-final languages, IP complement of C moves leftward to yield the verb-final structure. The OT constraints on clausal skeleton derive exactly this asymmetry: Head-L ≫ Spine-R optimally yields a head-initial language; the presence of functional projections (at least one violation of Spine-R) would be tolerated. Spine-R ≫ Head-L optimally yields a head-final language; Spine-R being high-ranked, the presence of any functional projection will be dispreferred, as there is an alternative in head-final languages of having bound inflectional elements morphologically attached to a verb. This view of phrase structure is also suggested and formally implemented by Haider (1997a, 1997b) in a different framework.

As Sells (2001) points out, Spine-R also prefers specifiers and adjunctions to be leftward: specifier is sister to a bar-level category, necessarily a co-head, and the node sister to an adjoined element will always be a co-head also, whether adjunction takes place at XP, X', or X^0. No other constraints or movements need to be postulated to capture this 'leftness' tendency observed and modeled by Kayne.[4]

The alignment constraint in (24) prefers subject to be the leftmost in a clause. More precisely, the constraint requires alignment of the c-structure node that maps to the f-structure of SUBJ and the c-structure that maps to the immediately outer f-structure nucleus. The details of how this is achieved are omitted here due to space limitations.

(24) SUBJ-L(eft): Subject aligns left in the clause.

Constraints on Adjacency

While earlier proposals on syntactic positioning based on the idea of alignment (here strictly interpreted as aligning identical edges) have been successful in capturing some typological generalizations about word order, headedness of phrases, and realization of grammatical information that correlates

[4] For other word order types not illustrated here, see Sells (2001); also Morimoto (2000b, chapter 3).

with clausal organization, it is not clear if alignment alone can explain all types of syntactic positioning. There is a rather wide range of syntactic elements that prefer adjacency with their respective heads (referred to as 'head-attracted' constituents in section 1): negation, focus, and (XP/X^0) adverbs typically appear near the verbal head; a relative pronoun and possessor prefer adjacency with their respective nominal head. Moreover, positioning of these constituents correlates with head directionality: verb-initial and verb-final languages often (though not always) exhibit the mirror image with respect to positioning of these constituents. In this respect, these head-attracted constituents thus deserve a unified treatment.

The formal mechanism proposed in the present work, referred to as *abutment*, is modeled on Generalized Alignment, and is intended to derive directional parameters. Abutment is alignment of *opposite*, rather than identical, edges. According to (25), there exists some category C_1 (e.g. focus, adverb, relative pronoun) and C_2 (e.g. V-head, N-head); abutment is satisfied if, for example, the left-edge of C_1 is adjacent to the right-edge of C_2. C_1 can be a c-structure node, or a f-structure attribute which maps onto a particular c-structure node. On the other hand, C_2 is always a c-structure head. Crucially, the constraint does not specify particular edges of the elements; the only requirement is that two edges be opposite.

(25) **Abutment**

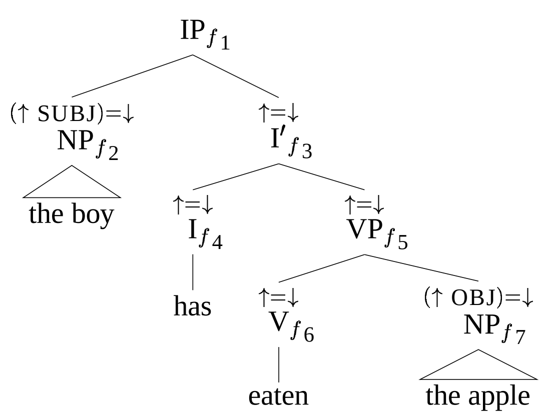
Abut(Cat_1,Edge_1, Cat_2,Edge_2) $=_{def}$
$\forall$ Cat_1 $\exists$ Cat_2 such that Edge_1 of Cat_1 shares Edge_2 of Cat_2.

Where $\text{Cat}_1 \in$ F-Cat $\cup$ C-Cat, $\text{Cat}_2 \in$ C-Cat
Edge_1, $\text{Edge}_2 \in \{$Right, Left$\}$ & $\text{Edge}_1 \neq \text{Edge}_2$.

Constraints instantiated by Abutment in (25) are abbreviated as "Abut-C_1(C_2)". To illustrate how Abutment works more concretely, let us consider OBJ abutment in (26) and the simple transitive sentence in (27); here F is instantiated by OBJ (f_7), and C by a verbal head (V-HD). C' will be the NP that corresponds to OBJ. According to abutment defined in (25), the verb-object configuration in (27) satisfies the constraint (26): there exists a node NP_{f_7} (= C' in (25)) which maps to the f-structure of of OBJ (= F in (25)), and the left edge of that node meets the right edge of V (= C in (25)).

(26) Abut-OBJ(V-HD): Abut E_1 of OBJ with E_2 of V-HD.

(27) [$_{IP}$ [$_{NP}$ The boy] [$_{I'}$ has [$_{VP}$ eaten an apple]]]



$$f_{1,3,5,6}: \begin{bmatrix} \text{PRED} & \text{`eat <}(f_1\ \text{SUBJ})\ (f_1\ \text{OBJ})\text{>'} \\ \text{SUBJ} & f_2: \begin{bmatrix} \text{PRED} & \text{`boy'} \end{bmatrix} \\ \text{OBJ} & f_7: \begin{bmatrix} \text{PRED} & \text{`apple'} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{ASP} & \text{PERF} \end{bmatrix}$$

Now consider the constraint on placement of negation in (28a). Here C_1 (Neg^0) and C_2 are both (part of) c-structure nodes. The constraint is satisfied as long as negation, however it is realized, is adjacent to the main verb as shown in (28b).

(28) a. Abut-NEG(V-HD): Abut E_1 of NEG with E_2 of V-HD.

b.

Neg Form	Before V	After V	Eval
Affixal	Neg-V	V-Neg	satisfied
	Neg-X-V	V-X-Neg	violated
Particle	V Neg	Neg V	satisfied
	V . . . Neg	Neg . . . V	violated
Auxiliary	V_{lex} V_{neg}	V_{neg} V_{lex}	satisfied
	$V_{lex} \ldots V_{neg}$	$V_{neg} \ldots V_{lex}$	violated

Constraints on X^0 Elements

Additional constraints needed to model the typology of negation discussed above include those on head positioning, given in (29). In English, *Lex-in-F is high-ranked (cf. Bresnan 2000b, Sells 2001), whereas V2 plays little role. In V2 languages, *Lex-in-F ranks below V2 so that a lexical head (inflected main verb) can occupy V2 position to respect the high-ranking V2 constraint.

(29) a. "V2": Satisfied by an inflected verbal element in the second position F^0. (cf. Sells 2001)

b. *Lex-in-F: Avoid lexical head in functional head position.

The constraint in (30) articulates the intuition that negation is associated with Infl (or F^0). At a superficial level for example, in derivational frameworks negation is often analyzed as the head of the function projection NegP. Although the data presented earlier revealed that languages apparently choose between the FP and LP domain for the placement of negation, the analysis presented below will show that negation in the LP domain is forced by other constraints, including a high-ranking constraint on structural economy (discussed below). In other words, if nothing else forces negation to be elsewhere, it will appear in the FP domain (under F' or adjoined to F^0).

(30) Neg-in-FP: Negation must in the FP domain.

Constraints on Economy and Markedness

Two constraints are introduced below on structural economy and morphological markedness. The economy constraint penalizes every X^0 and its projection. The underlying premise is that economical structure is preferred, and the most optimal output would be one with no structure at all. However, that would incur massive violation of faithfulness, as the input (semantic and syntactic) information would not be faithfully represented in the output. In other words, structure is driven by the need to satisfy faithfulness constraints. Economy constraints proposed in earlier work include $*X^0$ (Sells 1998) and Don'tProj (Bresnan 2000b) which penalizes empty nodes. ECONOMY in (31) is taken to be a more general constraint against any structure.

The constraint *Affix is a constraint against having a morphologically marked structure. The assumption is that bound forms are more marked than analytic forms. The relative markedness of morphologically bound forms is more clearly discernible in other areas of grammar such as pronominal systems. Bresnan's (2000a, 2001a) work on typological variation in pronominal systems reveals that there are languages with only free pronouns, as well as those with both free and reduced (bound) pronominal forms, but there are no languages with only reduced pronominals—displaying a classic markedness asymmetry. It is also well-known that newly created contact languages as a result of massive reduction and simplification of structures at all levels of grammar often exhibit recurring universal properties, such as a smaller inventory of consonant and vowel systems, preference for CV structure, simplified morphology, absence of complex sentences, and analytic syntax (Foley 1988, Thomason and Kaufmann 1988, Holm 1989, Bakker 1995). Preference of analytic forms over synthetic forms reported in the literature on pidgin languages (such as those cited above) is thus one such unmarked property (also see Bresnan 1998 for a discussion of these generalizations in the context of OT-LFG formalization).

(31) a. ECONOMY: Economical structure is preferred (violated by every X^0 and XP).

b. *AFFIX: Avoid affixes. (Analytic forms are preferred over synthetic forms.)

Below is a summary of the constraints discussed above. As indicated, some of the constraints that are not crucial in the analysis will be omitted from the tableaux for the sake of simplicity and compactness. They are nonetheless assumed to exist, and in the analysis presented below, those constraints will be satisfied by appropriate candidates.

(32) Summary of Constraints

a.

a.	Head-L	(19a)	f.	ECONOMY	(31a)	
b.	Spine-R	(19b)	g.	*AFFIX	(31b)	
c.	"V2"	(29a)	h.	Abut-OBJ(V-HD)	(26)	omitted from tableaux
d.	Neg-in-FP	(30)	i.	SUBJ-L	(24)	omitted from tableaux
e.	Abut-NEG(V-HD)	(28)	j.	*Lex-in-F	(29b)	omitted from tableaux

4 Deriving a Typology of Negation

Having established the constraints necessary to derive the basic clause structure and placement of negation, we now turn to the discussion of exactly how these constraints derive the cross-linguistic variation in the form and placement of negation. The languages considered here are again listed below.

(33) Candidates

I.	Japanese (SOV)	V.	English (SVO)
I I.	Bantu (SVO)	VI.	German (SOV, V2)
III.	Korean preverbal NEG (SOV)	VII.	Swedish (SVO, V2)
IV.	Korean postverbal NEG		

The rankings in (34) show abstractly which constraints play a crucial role in each language. I will discuss these rankings together with the tableaux presented below. The first to note in the overall analysis is that as mentioned earlier head directionality is determined by interaction of two constraints on clausal skeleton, Head-L and Spine-R: head-final structure in languages like Japanese is ensured by ranking Spine-R (b) above Head-L (a); head-initial structure in languages like Bantu and English is derive by the reverse ranking. In some languages (e.g. V2 languages such as German and Swedish) these constraints on clausal skeleton exhibit an interesting interaction with other constraints, particularly those on head positioning. In others (Types I–IV), such interaction is not obvious. The ranking of these constraints relative to others are therefore not crucial in the analysis for those languages.

(34) Constraint Rankings

a. Head-L	b. Spine-R	c. "V2"	d. Neg-in-FP
e. Abut-NEG(V-HD)	f. ECONOMY	g. *AFFIX	

I.	b ≫ a ≫ f ≫ e ≫ d ≫ g ≫ c
II.	a ≫ b ≫ f ≫ e ≫ d ≫ g ≫ c
III & IV.	b ≫ a ≫ h ≫ f ≫ g ≫ e ≫ c
V.	d ≫ a ≫ e ≫ b ≫ g ≫ f ≫ c
VI.	c ≫ g ≫ f ≫ b ≫ a ≫ e ≫ d
VII.	c ≫ d ≫ a ≫ b ≫ f ≫ e ≫ g

Focusing now on the rest of the constraints, in Types I and II, ECONOMY is a high-ranking constraint and ranks above *Affix: this ranking prefers a synthetic form of negation over an analytic form. Recall that clauses in these languages are taken to be organized around the exocentric category S, due to the lack of evidence for Infl (cf. Sells 1995 for Japanese; Bresnan and Mchombo 1987, Morimoto 2000 for Bantu). Given the preference for S over IP, Neg-in-FP and V2 are presumably low-ranked.[5] Reversing the ranking of Spine-R and Head-L (i.e. Head-L ≫ Spine-R), we get Type II languages.

(35) Type I: Japanese, Turkish

		Spine-R	Head-L	ECONOMY	ABUT-NEG	Neg-in-FP	*AFFIX	"V2"
I. ☞	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ V-**NEG**]]	0	1	7		∗	∗	
II.	$[_S$ NP$_{su}$ $[_{VP}$ **NEG**-V NP$_{oj}$]]	1!	0	7		∗	∗	
III.	$[_S$ NP$_{su}$ $[_{VP}$ $[_{V^0}$ **NEG**^{0} V^0] NP$_{oj}$]]	0	1	8!		∗		
IV.	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ $[_{V^0}$ V^0-COMP **NegV**^{0}]]]	0	1	8!		∗		
V.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]	2!	0	8				
VI.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 $[_{VP}$ NP$_{oj}$ $[_{V'}$ **NEG** V]]]]	1!	1	8		∗		
VII.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]	2!	0	8				

The ranking for Types III and IV yields both the structure with a preverbal negative particle and the one with a postverbal negative auxiliary in Korean. *Affix crucially ranks above ECONOMY, favoring

[5]Constraints that determine the choice between S and IP are omitted in the present discussion. A previously proposed constraint like *S-node (Morimoto 2000), for example, disfavors the exocentric node S, whereas OB-HD(FP) ("FP must have an overt head"), if ranked above *S-node, ensures the root node will be S in the absence of the overt head F^0.

an analytic form over a synthetic form. Like Japanese, Korean lacks evidence for Infl (cf. Cho and Sells 1995), and is thus taken to make use of the exocentric category S. The constraints Neg-in-FP and "V2" therefore are ranked low.

(36) Type III & IV: Korean

		Spine-R	Head-L	*AFFIX	ABUT-NEG	ECONOMY	Neg-in-FP	"V2"
I.	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ V-**NEG**]]	0	1	∗!		7	∗	
II.	$[_S$ NP$_{su}$ $[_{VP}$ **NEG**-V NP$_{oj}$]]	1!	0	∗		7	∗	
III. ☞	$[_S$ NP$_{su}$ $[_{VP}$ $[_{V^0}$ **NEG**^{0} V^0] NP$_{oj}$]]	0	1			8	∗	
IV. ☞	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ $[_{V^0}$ V^0-COMP **NegV**^{0}]]]	0	1			8	∗	
V.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]	2!	0			8		
VI.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 $[_{VP}$ NP$_{oj}$ $[_{V'}$ **NEG** V]]]]	1	1!			8	∗	
VII.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]	2!	0			8		

In English (Type V), Neg-in-FP is ranked above Abut-Neg(V-HD) even though here, neither is violated and the ranking is therefore not significant. The ranking of these constraints as proposed becomes crucial in utterances like those in (38)–(39), where negation must appear before the parenthetical phrase that is presumably adjoined to VP.

(37) Type V: English (present perfect)

		Neg-in-FP	Head-L	ABUT-NEG	Spine-R	ECONOMY	*AFFIX	"V2"
I.	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ V-**NEG**]]	∗!	1		0	7	∗	
II.	$[_S$ NP$_{su}$ $[_{VP}$ **NEG**-V NP$_{oj}$]]	∗!	0		1	7	∗	
III.	$[_S$ NP$_{su}$ $[_{VP}$ $[_{V^0}$ **NEG**^{0} V^0] NP$_{oj}$]]	∗!	1		0	8		
IV.	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ $[_{V^0}$ V^0-COMP **NegV**^{0}]]]	∗!	1		0	8		
V. ☞	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]		0		2	8		
VI.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 $[_{VP}$ NP$_{oj}$ $[_{V'}$ **NEG** V]]]]	∗!	1		1	8		
VII. (☞)	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]		0		2	8		

(38) a. John has **not**, $[_{VP}$ in fact, left the country].

b. ??John has, $[_{VP}$ in fact, **not** left the country].

(39) a. The meeting did **not**, $[_{VP}$ however, bring peace to the nation].

b. ??The meeting did, $[_{VP}$ however, **not** bring peace to the nation].

(40) Type V: German

		"V2"	*AFFIX	ECONOMY	Spine-R	Head-L	ABUT-NEG	Neg-in-FP
I.	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ V-**NEG**]]	$*!$	$*$	7	0	1		$*$
II.	$[_S$ NP$_{su}$ $[_{VP}$ **NEG**-V NP$_{oj}$]]	$*!$	$*$	7	1	0		$*$
III.	$[_S$ NP$_{su}$ $[_{VP}$ $[_{V^0}$ **NEG**^{0} V^0] NP$_{oj}$]]	$*!$		8	0	1		$*$
IV.	$[_S$ NP$_{su}$ $[_{VP}$ NP$_{oj}$ $[_{V^0}$ V^0-COMP **NegV**^{0}]]]	$*!$		8	0	1		$*$
V.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]			8	2!	0		
VI. ☞	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 $[_{VP}$ NP$_{oj}$ $[_{V'}$ **NEG** V]]]]			8	1	1		
VII.	$[_{FP}$ NP$_{su}$ $[_{F'}$ F^0 **NEG** $[_{VP}$ V NP$_{oj}$]]]			8	2!	1		

In German, the ranking of Spine-R above Head-L ensures the optimal candidate to be head-final, provided that the higher-ranked constraint "V2" is respected. *Affix and ECONOMY are also high-ranked, crucially dominating both Abut-Neg(V-HD) and Neg-in-FP. The ranking produces the result that in a main clause negation will appear inside VP and not be in I^0, next (or at least closer) to the inflected verb in V2 position—head of CP in German. For concreteness, consider the two structures in (41).

(41)

		ECONOMY
a.	$[_{CP}$ NP$_{subj}$ $[_{C'}$ V$_{fin}$ $[_{IP}$ $[_{I'}$ Neg^0 $[_{VP}$ NP$_{obj}$]]]]]	9 violations
b.	$[_{CP}$ NP$_{subj}$ $[_{C'}$ V$_{fin}$ $[_{VP}$ NP$_{obj}$ $[_{V'}$ Neg^0]]]]	8 violations

In the structure in (41a), the finite verb is in C and the negative particle in I. This configuration would satisfy both Abut-Neg and Neg-in-FP. The structure in (41b), on the other hand, violates both these constraints: the finite verb is in C, and negation is in head-adjoined position under V$'$. In terms of economy, however, the structure in (41a) violates the ECONOMY constraint more severely. For convenience, we simply count the nodes that are shown in the brackets for the violation of ECONOMY: CP, NP$_{subj}$, C$'$, V$_{fin}$, IP, I$'$, Neg^0, VP, and NP$_{obj}$—the total of nine violations. In (41b), on the other hand, there are eight violations: CP, NP$_{subj}$, C$'$, V$_{fin}$, VP, NP$_{obj}$, and Neg^0. The ranking of these constraints as proposed thus yields the earlier example, repeated here in (42).

(42) $[_{CP}$ Warum $[_{C'}$ gefiel $[_S$ $[_{NP}$ unsere Lösung] $[_{NP}$ dem Peter] $[_V$ **nicht**]]]]?
why pleased our solution Peter not
'Why did our solution not please Peter?' (Haegeman 1995:168)

The example is a *wh*-question, where the *wh*-element *warum* 'why' occupies SpecCP, and the finite verb *gefiel* 'pleased' in C^0. Neg^0 adjoined to V, although the V head appears in C. As pointed out earlier, there is also empirical evidence that German lacks IP. This is another reason to eliminate a structure like that in (41a) where negation would be the only lexical item to occupy I.

Again, an important generalization that emerges from these data is that negation is associated with Infl, and the most straightforward syntactic expression is to be in I or in the domain of IP (e.g. under I$'$). If the language in question lacks I (and IP), then negation will be associated with some other verbal category, typically V. In languages like Japanese that lack any functional categories or languages like Bantu in which there is only a handful of lexical items that appear in C and everything else (e.g. tense, aspect, negation), both tense and negation are expressed as either suffixes or prefixes. For example in Japanese (and Turkish), both negation and tense are expressed as suffixes on the verb, as the earlier examples in (1)–(2) illustrate. In Bantu, both are expressed as prefixes, as illustrated earlier in (3). In Malagasy, illustrated earlier in (8), tense is expressed as a prefix on the verb, and negation is preverbal (expressed as a particle). These all point to the generalization that negation exhibits close affinity with tense (i.e. I or some other verbal category like V). Head-final languages where the sequence of functional elements and the main verb is V-I-C typically lack evidence for I or C, and negation is suffixal. In head-initial languages, the sequence of functional elements and the main verb is C-I-V. In languages like Bantu in which there is no evidence for Infl, negation is expressed as a prefix on the main verb along with tense; languages that exhibit IP (e.g. Swedish), the negative particle falls within the IP domain. In short, the form and placement of negation is generally predictable based on the presence/absence of I/IP and head-directionality.[6]

In Swedish, “V2” and Neg-in-FP are high-ranked. The ranking Head-L $\gg$ Spine-R yields head initial structure. As we saw in the tableau in (37), for present perfect clauses, either the ranking for English or the one for Swedish yields both their optimal candidates (V and VII). This is because even though English is not a V2 language, in present perfect, the auxiliary verb occupies I^0, rendering a V2-like structure and consequently satisfying the “V2” constraint as defined in (29a). These rankings, however, yield different results in simple past tense clauses and other clause types such as subordination and non-subject topicalization.

The tableau in (43) represents a Swedish simple past tense clause. Here we focus on Types V (English), VI (German), and VII (Swedish), as the others do not differ much from the earlier tableaux. Due to the absence of the V2 requirement and the ranking of *Lex-in-F above “V2” (not shown in the tableau) that bans a lexical head in functional head position, English employs *do* in simple past. This is interpreted as a violation of “V2”. It is possible to interpret *do* as being a lexical head filling functional head position, but this is not a desirable interpretation because the function of *do* which is basically to host tense information in the absence of an auxiliary verb precisely in cases where there is no auxiliary element to fill the functional head position. In Swedish, due to the high-ranking “V2” constraint (above *Lexi-in-F), the finite verb fills F^0. An example of this optimal structure is given in (44) from Swedish (Sells 2000, ex(18c)).

(43) Main clause (simple past)

		“V2”	Neg-in-FP	Head-L	Spine-R	ECONOMY	ABUT-NEG	*AFFIX
V.	$[_{FP}$ NP$_{su}$ $[_{F'}$ DO **NEG** $[_{VP}$ V NP$_{oj}$]]]	$*$		0	2	10		
VI.	$[_{FP}$ NP$_{su}$ $[_{F'}$ V$_{fin}$ $[_{VP}$ NP$_{oj}$ **NEG**]]]]		$*$	0	1	9	$*$	
VII. ☞	$[_{FP}$ NP$_{su}$ $[_{F'}$ V$_{fin}$ **NEG** $[_{VP}$ NP$_{oj}$]]]			0	2	9		

(44) Jag gav **inte** Elsa någonting.
I gave not Elsa anything

[6]The OT-LFG approach to clause structure pursued here also makes a claim that the presence/absence of functional projections in a given language is predictable in part based on the head-directionality of that language, as discussed earlier in section 3.

This ranking also eliminates candidate VI, the optimal structure for German, in which negation is not in FP. It is important to note that in German Spine-R will be ranked above Head-L and Neg-in-FP, as we saw in the tableau in (40): "V2" ... *Spine-R $\gg$ Head-L ... Neg-in-FP. High-ranking Spine-R prefers the structure in VI over the one in VII: VI violates Spine-R only once, but VII, placing Neg under F′ along with the leftmost V-head, violates it twice. The interaction of these constraints presented here thus formally illustrates how the position of negation correlates with the presence/absence of I and head-directionality.

The tableau in (45) represents a subordinate clause in Swedish. Here I only consider the two V2 languages in which the position of the finite verb differs from that in a matrix clause. These illustrate essentially the same point as that in the preceding tableau in (43): the interaction of Neg-in-FP and Spine-R is crucial here. The data that confirms the result of the constraint interaction in (45) is provided in (46) from Swedish (Sells 2000, ex(3c)).

(45) Subordinate clause

		"V2"	Neg-in-FP	Head-L	Spine-R	ECONOMY	ABUT-NEG	*AFFIX
VI.	[$_{C'}$ Comp [$_S$ NP$_{su}$ [$_{VP}$ [$_{V'}$ NP$_{oj}$ **NEG** V]]]]	*	*	0	2	11		
VII. ☞	[$_{C'}$ Comp [$_{FP}$ NP$_{su}$ [$_{F'}$ **NEG** [$_{VP}$ V NP$_{oj}$]]]	*		0	3	11		

(46) ... att jag **inte** [$_{VP}$ har gett boken till henne]
... that I not [$_{VP}$ have given the.book to her]

The tableau in (47) represents non-subject topicalization also in Swedish, again illustrating essentially the same point as the two preceding tableaux. Note additionally that in English, Abut-Neg ranks above Spine-R, and "V2" is low-ranked: Neg-FP $\gg$ Head-L $\gg$ Abut-Neg $\gg$ Spine-R $\gg$ ... $\gg$ "V2" (cf. (37)). This ranking would yield the structure in V to be optimal in (47).

(47) Nonsubject initial clause (simple past)

		"V2"	Neg-in-FP	Head-L	Spine-R	ECONOMY	ABUT-NEG	*AFFIX
V.	[$_{IP}$ NP$_{top}$ [$_{IP}$ NP$_{subj}$ [$_{F'}$ DO **NEG** [$_{VP}$ V NP$_{obj}$]]]]	*		0	3	13		
VI.	[$_{CP}$ NP$_{top}$ [$_{C'}$ V$_{fin}$ [$_S$ NP$_{subj}$ [$_{VP}$ NP$_{obj}$ **Neg**]]]]]		*	0	1	13	*	
VII. ☞	[$_{CP}$ NP$_{top}$ [$_{C'}$ V$_{fin}$ [$_{IP}$ NP$_{subj}$ [$_{I'}$ **NEG** [$_{VP}$ NP$_{obj}$]]]]]			0	2	13	*	

5 Conclusion

This paper has aimed at providing insight into the following areas: from the formal perspective within OT-LFG, I have proposed an additional formal mechanism that enables us to derive directionality parameters by interacting with independently motivated constraints on clausal skeleton and head-positioning.

This contributes to the on-going development in OT-LFG clausal syntax, which attempts to integrate universal properties of phrase structure into formal theory of clausal syntax. The present analysis of a syntactic typology of negation demonstrates that directionality parameters can be derived without losing Kayne's insight that phrase structure is fundamentally antisymmetric or recourse to unmotivated movement operations.

From the theoretical perspective, in particular with respect to the syntax of negation, the present analysis makes a claim that both form and placement of negation in a given language are in part predictable based on the head-directionality and the presence/absence of I/IP of that language—the generalization that is difficult to capture without seriously taking into account universal markedness and typological tendencies in formal theory of clausal syntax.

References

Bakker, Peter. 1995. Pidgins. In Jacques Arends, Pieter Muysken, and Norval Smith (eds.), *Pidins and Creoles. An Introduction*, 25–39. Amsterdam, Benjamins.

Berman, Judith. 2000. *Topics in the Clausal Syntax of German*. Doctoral dissertation, Universität Stuttgart, Stuttgart.

Bresnan, Joan. 1998. Pidgin genesis and optimality theory. In Miriam Butt and Tracy Holloway King (eds.), *On-line Proceedings of the LFG-98 Conference*, Stanford, CA. CSLI Publications. On-line at `http://csli-publications.stanford.edu/`.

Bresnan, Joan. 2000a. The emergence of the unmarked pronoun: Chicheŵa pronominals in Optimality Theory. In *Proceedings of the Berkeley Linguistics Society 23, 1997*. Berkeley, CA, Berkeley Linguistics Society. ROA-179-0297.

Bresnan, Joan. 2000b. Optimal syntax. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, Syntax and Acquisition*, 334–385. Oxford University Press.

Bresnan, Joan. 2001a. The emergence of the unmarked pronoun. *Optimality-theoretic Syntax*, ed. by G. Legendre, S. Vikner, and J. Grimshaw, The MIT Press.

Bresnan, Joan. 2001b. *Lexical Functional Syntax*. Oxford, Blackwell.

Bresnan, Joan, and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63(4), 741–782.

Butt, Miriam. 1994. Complex predicate scrambling in Urdu. In Miriam Butt, Tracy Holloway King, and Gillian Ramchand (eds.), *Theoretical Perspective on Word Order in South Asian Languages*, 67–90. Stanford, CA, CSLI Publications.

Cho, Young-Mee Y., and Peter Sells. 1995. A lexical account of inflectional suffixes in Korean. *Journal of East Asian Linguistics* 4 (2), 119–174.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford, CA, CSLI Publications.

Costa, João. 1998. *Word Order Variation: A Constraint-Based Approach*. The Netherlands, Holland Academic Graphics.

Dahl, Östen. 1979. Typology of sentence negation. *Linguistics* 17, 79–106.

Dryer, Matthew S. 1988. Universals of negative position. In Machael Hammond, Edith A. Moravcsik, and Jessica R. Wirth (eds.), *Studies in Syntactic Typology*, 93–124. Philadelphia/Amsterdam, John Benjamins.

Foley, William A. 1988. Language birth: the processes of pidginization and creolization. In Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, Vol. IV of *Language: The Socio-cultural Context*, 162–183. Cambridge, Cambridge University Press.

Grimshaw, Jane. 1997. Projection, heads, and Optimality. *Linguistic Inquiry* 28(3), 373–422.

Grimshaw, Jane. 2001. Economy of structure in OT. MS. Rutgers University. Available on-line via `http://ruccs.rutgers.edu/roa.html`.

Haegeman, Liliane. 1995. *The Syntax of Negation*. Cambridge, UK, Cambridge University Press.

Haider, Hubert. 1997. Precedence among predicates. *Journal of Comparative Germanic Linguistics* 1, 3–41.

Holm, John. 1989. *Pidgins and Creoles*. Vol. II of *Language Survey*. Cambridge, Cambridge University Press.

Holmberg, Anders. 1993. Two subject positions in IP in Mainland Scandinavian. *Working Papers in Scandinavian Syntax* 52, 29–41.

Holmberg, Anders, and Christer Platzack. 1988. On the role of inflection in Scandinavian Syntax. Working Papers in Scandinavian Syntax.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA, The MIT Press.

Kim, Johng-bok. 2000. *The Grammar of Negation*. Stanford, CA, CSLI Publications.

Kimenyi, Alexandre. 1980. *A Relational Grammar of Kinyarwanda*. Berkeley, CA, University of California Press.

McCarthy, John, and Alan S. Prince. 1993. Generalized alignment. In Geart Booji and Jaap van Marle (eds.), *Yearbook of Morphology*, 79–153. Dordrecht, Foris.

Mitchell, Erika. 1991. Evidence from Finnish for Pollock's theory of IP. *Linguistic Inquiry* 22, 373–379.

Morimoto, Yukiko. 2000. *Discourse Configurationality in Bantu Morphosyntax*. Doctoral dissertation, Stanford University, Stanford, CA.

Morimoto, Yukiko. 2001. Verb raising and phrase structure variation in OT. In Peter Sells (ed.), *Formal and Empirical Issues in Optimality-Theoretic Syntax*. Stanford, CA, CSLI Publications.

Payne, John R. 1985. Negation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, Vol. 1, 197–242. Cambridge, Cambridge University Press.

Rackowski, Andrea, and Lisa Travis. 2000. V-initial languages: X or XP movement and adverbial placement. In Andrew Carnie and Eithne Guilfoyle (eds.), *The Syntax of Verb Initial Languages*, 117–142. New York, Oxford University Press.

Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. Doctoral dissertation, Rutgers University, New Jersey.

Sells, Peter. 1995. Korean and Japanese morphology from a lexical perspective. *Linguistic Inquiry* 26, 277–325.

Sells, Peter. 1998. Optimality and economy of expression in Japanese and Korean. In Noriko Akatsuka et al. (ed.), *Japanese/Korean Linguistics*, Vol. 7, 499–514. CSLI, Stanford Linguistics Association.

Sells, Peter. 1999a. Constituent ordering as alignment. In Susumu Kuno et al (ed.), *Harvard Studies in Korean Linguistics*, Vol. 8, 546–560, Seoul. Hanshin.

Sells, Peter. 1999b. Order as generalized alignment. Handout of the talk given at Gothenburg University, Department of Swedish Language, May 19, 1999.

Sells, Peter. 2000. Negation in Swedish: Where it's not at. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG00 Conference*, Stanford, CA. CSLI Publications. On-line at `http://csli-publications.stanford.edu/`.

Sells, Peter (ed.). 2001a. *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford, CA, CSLI Publications.

Sells, Peter. 2001b. *Structure, Alignment and Optimality in Swedish*. Stanford, CA, CSLI Publications. Forthcoming.

Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, CA, University of California Press.

University of Düsseldorf, Department of Linguistics
`morimoto@phil-fak.uni-duesseldorf.de`

Categorial Grammar and Lexical-Functional Grammar*

Reinhard Muskens
Department of Linguistics
Tilburg University
The Netherlands

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong
Miriam Butt and Tracy Holloway King (Editors)
2001
CSLI Publications
http://csli-publications.stanford.edu/

*I wish to thank the participants of LFG2001, especially Mary Dalrymple and Ron Kaplan, for valuable feedback. All comments are very welcome.

Abstract

This paper introduces λ-grammar, a form of categorial grammar that has much in common with LFG. Like other forms of categorial grammar, λ-grammars are multi-dimensional and their components are combined in a strictly parallel fashion. Grammatical representations are combined with the help of *linear combinators*, closed pure λ-terms in which each abstractor binds exactly one variable. Mathematically this is equivalent to employing linear logic, in use in LFG for semantic composition, but the method seems more practicable.

While λ-grammars could be used to formalize many approaches to grammatical theory, they are certainly natural as a basis for the formalization of LFG. This leads to a theory I would like to call λ-LFG. In this paper it will be shown how the standard components of LFG can be set up in the framework. We will have descriptions of c-structure, descriptions of f-structure, and semantics. The difference between defining and constraining information will be explained in terms of entailment, and requirements on long-distance paths in f-structure will be explained in terms of entailment in the presence of a simple set of axioms.

1 Introduction

In this paper I want to discuss a version of Lexical-Functional Grammar (LFG) that is also a version of Categorial Grammar (CG). A convergence between the two frameworks has set in at least since Zeevat, Klein, and Calder (1986) proposed a 'sign-based' categorial grammar and Oehrle (1988) defined 'multi-dimensional' categorial grammars. The sign-based approach, now adopted by many categorial grammarians (e.g. Moortgat 1991; Morrill 1994; Moortgat 1997),[1] allows various grammatical representations to be combined in tandem, a move clearly reminiscent of LFG's multi-component architecture.

A second important step leading to further convergence of the two frameworks was Dalrymple et al.'s (1993) proposal to use Linear Logic[2] as a 'glue' between f-structure and semantics. The $\{\multimap, \otimes\}$ fragment of Intuitionistic Linear Logic that is used in this 'glue' approach is in fact identical with the undirected Lambek Calculus.[3] This still leaves a gap between the two frameworks, as most versions of

[1]Signs in Categorial Grammar are sequences of representations. For example, Zeevat, Klein, and Calder (1986) work with signs of the form ⟨phonology, syntactic category, semantics, order⟩. Each slot in the sequence may be associated with its own form of reasoning. This should be distinguished from the concept of a sign in Head-Driven Phrase Structure Grammar, where multiple levels of the grammar are represented in one attribute-value matrix.

[2]The original paper on Linear Logic is (Girard 1987). An attractive textbook (Troelstra 1992).

[3]For the original version of the Lambek Calculus, see Lambek 1958; for the undirected version

Categorial Grammar are directional and distinguish between categories A/B (seeking a B to the right) and $B\backslash A$ (seeking a B to the left). Linear Logic and the undirected Lambek Calculus collapse this to a single $B \multimap A$, a formula that consumes a B and then produces an A, irrespective of order. The motivation for using a directional system comes from a wish to treat word order directly on the level of the calculus. Grammatical formalisms that treat word order in a different way, as LFG does, can make do with nondirectionality.

In fact the treatment of word order can be factored out of the general calculus in CG as well. Consider the signs in (1), which are in the style of (Moortgat 1991; Morrill 1994) and are triples consisting of a string, a semantic term, and a directional category. Since (1a) has a type which seeks an *np* to the left, it can be combined with (1b). The combination consists of string concatenation in the first dimension (forming *likes John*), application in the second dimension (forming $\lambda x\lambda y.[like(y,x)](j)$, which reduces to $\lambda y.like(y, j)$), and a form of Modus Ponens in the third dimension. The result is as in (1c).

(1) a. likes : $\lambda x\lambda y.like(y,x) : (np\backslash s)/np$

b. John : $j : np$

c. likes John : $\lambda y.like(y, j) : np\backslash s$

The directionality of the types here codes the way in which strings can be concatenated, but, as was observed in (Oehrle 1994; Oehrle 1995), this information can be shifted to the syntactic dimension if we are willing to let syntactic representations be lambda terms, in analogy with semantic representations. (1a) then becomes (2a), whose first element is a lambda term over strings and whose last element is an undirected type. Combining (2a) with (1b), now using application in the first as well as in the second dimension, leads to (2b).

(2) a. λxλy. y likes x : $\lambda x\lambda y.like(y,x) : np \multimap (np \multimap s)$

b. λy. y likes John : $\lambda y.like(y, j) : np \multimap s$

There is much to be said for such a shift of word order information from the level of types to the syntactic dimension.[4] A priori it seems that once we have a syn-

Benthem 1986; Benthem 1988; Benthem 1991; for a survey of Categorial Grammar in the Lambek tradition Moortgat 1997.

[4]Interestingly, Curry (1961) already argues for this approach. Curry considers *functors*, which are expressions containing subscripted blanks, such as '$—_1$ is between $—_2$ and $—_3$' or '$—_1$ were eaten by the children'. Functors can apply to arguments and arguments are to be substituted for blanks in the order of the subscripts. Essentially then, although Curry does not explicitly mention this, functors are lambda terms over syntactic objects. For example, the first of the functors just mentioned can also be written '$\lambda x\lambda y\lambda z.x$ is between y and z'.

tactic dimension, word order should be handled there. But there are also empirical consequences, one positive, the other less so. Let us look at a consequence that is less than positive. One clear attraction of categorial grammars is the way in which (non-constituent) coordination is treated (see e.g. Steedman 1985; Dowty 1988; Moortgat 1988). The sign in (3a) can be formed from (1a) and (1b) with the help of hypothetical reasoning and can subsequently be coordinated with the similar sign in (3b). The result is as in (3c), which may be used to obtain (3d). Coordination is concatenation in the first dimension (with the coordinating word in between the coordinated elements) and either union or intersection in the semantic dimension, depending on whether the coordination is a disjunction or a conjunction. The types of the coordinated signs must be equal and the result will have the same type again.

(3) a. John likes : $\lambda x.\mathit{like}(j,x) : s/np$

b. Mary hates : $\lambda x.\mathit{hate}(m,x) : s/np$

c. John likes but Mary hates : $\lambda x.\mathit{like}(j,x) \wedge \mathit{hate}(m,x) : s/np$

d. John likes but Mary hates bananas : $\mathit{like}(j,b) \wedge \mathit{hate}(m,b) : s$

Treatments of coordination along these lines are not without their problems and usually result in overgeneration (see (Milward 1994; Moortgat 1997) for good discussions), but nevertheless seem to fare well compared to many other approaches. The analysis depends on a reduction of the syntactic dimension to *strings* and collapses when richer structures such as trees are taken to be basic. A move to representations such as the ones in (2) does not seem to be compatible with this analysis of coordination either.

On the other hand such representations immediately solve one of the problems of the standard Lambek Calculus. Since *John likes* can be analysed as an s/np, it is easy to obtain parses for sentences such as (4a) in Lambek's system. The trick is to categorize *who* as an item that searches an s/np to its right. *who* must also be categorized as searching for an $np\backslash s$ in view of (4b). But what to do if the gap is medial, as in (4c)?

(4) a. a woman who John likes

b. a woman who likes John

c. a woman who John likes enormously

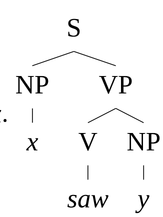
In order to deal with this problem clever extensions of the Lambek Calculus have been proposed (e.g. Morrill 1994). But these extensions are also considerable complications of the original idea. For the kind of representations in (2), on the other hand, a problem with medial gaps does not even arise: The relevant representations can be taken to be λx.John likes x, λx.x likes John, and λx.John likes x enormously, all of type $np \multimap s$. The relative pronoun can subcategorize for this type.[5]

(5) a. Everyone loves someone

b. Bill thinks someone kissed Carol

Similarly, it is well known that the Lambek Calculus can deal with scope ambiguities as long as the scope taking elements are in a peripheral position: (5a) will get its desired two readings. But as soon as a scope taking element is non-peripheral, difficulties may arise (Hendriks 1993; Morrill 1994; Moortgat 1997; Dalrymple et al. 1999). For example, (5b) will not obtain a reading in which *someone* takes scope over *thinks.* Again, operators (such as Moortgat's $\uparrow$) have been introduced that cleverly circumvent the problem (see Dalrymple et al. 1999 however) but also complicate the categorial machinery. Moving to an undirected calculus will make the problem disappear, as the periphery does not play any special role in such a system. I take this to be strong evidence that word order information should not be treated on the level of the type system but should be dealt with in a syntactic dimension. If it is represented in the type system this essentially syntactic information causes trouble in the semantics.

If Oehrle's move from simple strings to lambda-ed strings and from a directed to an undirected calculus is accepted then a third step narrowing the gap between CG and LFG is made. The same calculus now plays a central role in both grammatical frameworks. An obvious next step, made in (Oehrle 1999), is to bring the various components of the categorial grammar into line with the components that are recognized in LFG. Oehrle considers sequences such as the one in (6), consisting of a lambda-ed c-structure, a lambda-ed f-structure, a semantic term, and a type. Such signs are combined with the help of the undirected Lambek Calculus.

(6) $\lambda yx.$ 

$: \lambda yx. \begin{bmatrix} \text{PRED} & \text{`SEE'} \\ \text{TENSE} & \text{PAST} \\ \text{SUBJ} & x \\ \text{OBJ} & y \end{bmatrix} : \lambda yx.see(x,y) : np \multimap (np \multimap s)$

[5]The sign for the pronoun *who* could be taken to be λXλy.y who X(ε) : $\lambda X \lambda Y \lambda z.X(z) \wedge Y(z)$: $(np \multimap s) \multimap (n \multimap n)$, where ε is the empty string.

syntax:	$k, f : \nu$	$t, F : \nu t$	$T, \mathcal{F} : (\nu t)(\nu t)$	
semantics:	$x, y : e$	$i, j : s$	$p : st$	$P : e(st)$

Table 1: Typographical conventions for variables used in this paper. *Var* : *Type* means that *Var* (with or without subscripts or superscripts) always has type *Type*.

Oehrle's work will be my point of departure and should receive much praise, but I have a point of criticism as well. The terms in the first two dimensions of (6) are *structures*. More precisely, they are functions that take certain arguments and then give structures as a result. While I think it possible to flesh this out formally, I also think it is not very promising as a formalization of LFG. One of the important insights of (Kaplan and Bresnan 1982) was a separation, at least in the f-dimension, of *descriptions* and the structures satisfying them. This distinction, which is also at the heart of model theory, is lost if signs get a form as in (6).[6] As a consequence, it will not be possible to model *disjunctive* or *negative* constraints. It will also not be possible to formalize the difference between *defining* and *constraining* information, or to explain path constraints on f-structure arising out of long-distance dependencies. Such mechanisms are part and parcel of Lexical-Functional Grammar and in this paper I will show that they can easily be accounted for within a combined CG-LFG approach if a shift from structures to descriptions is made.

A second departure from Oehrle's proposal that I want to suggest involves a simplification of the technical machinery which now becomes possible. Thus far we have considered signs that needed to be combined with the help of some form of Lambek Calculus, but it is well-known (see Benthem 1986, for example) that proofs in the implicational fragment of the undirected Lambek Calculus (= Linear Logic) are isomorphic with *linear combinators*, closed pure λ-terms in which each abstractor binds exactly one free variable. This will make it possible to do away with proofs altogether and to just consider certain linear combinations of lexical elements.

2 λ-Grammars

To make concrete what I have in mind, I will define a small toy grammar in this section. In the next section some elaborations and revisions will be discussed. Consider the three signs in (7).

(7) a. john : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3) : john$

[6] A footnote in (Oehrle 1999) credits a referee with suggesting a move from structures to descriptions in the f-dimension.

b. $\mathsf{mary} : \lambda f.arc(f,cat,N) \wedge arc(f,num,sg) \wedge arc(f,pers,3) : mary$

c. $\lambda t_1 \lambda t_2.[t_2\ [\mathsf{loves}\ t_1]]$:
$\lambda F_1 \lambda F_2 \lambda f \exists f_1 f_2 [F_1(f_1) \wedge F_2(f_2) \wedge arc(f,cat,V) \wedge arc(f,tense,pres) \wedge$
$arc(f_1,cat,N) \wedge arc(f,obj,f_1) \wedge arc(f_2,cat,N) \wedge arc(f_2,num,sg) \wedge$
$arc(f_2,pers,3) \wedge arc(f,subj,f_2)]$:
$\lambda x \lambda y \lambda i.love(y,x,i)$

These signs each consist of a c-structure component, an f-structure component, and a semantic component. Expressions in sans serif in the c-structure terms are of type νt, and denote sets of nodes.[7] For example, john can be thought of as the set of nodes that are labeled 'John', whereas an expression such as $[\mathsf{loves\ mary}]$ can be thought of as the set of nodes k directly dominating a node k_1 labeled 'loves' and a node k_2 labeled 'Mary', with k_1 preceding k_2. More information about c-structure components will follow in the next section.

The f-components of our signs consist of λ-terms over the first order feature language of (Johnson 1991) and the semantics in the third component is in accordance with a streamlined form of Montague's (1973) theory. Constants *john* and *mary* are of type e and *love* is of type $e(e(st))$. Constants *cat, num, pers,* etc. are of a type a (attributes), while N, sg, 3, ... are of type ν (nodes).[8] More typing information is given in Table 1. For the moment, we consider a grammar with three dimensions, but in general the number of dimensions of a grammar is arbitrary (though fixed). The terms that we are interested in are all closed and we require that lexical elements have closed terms in each dimension.

Signs can be combined by means of *pointwise application.* In general, if $M = \langle M_1, \ldots, M_n \rangle$ and $N = \langle N_1, \ldots, N_n \rangle$ are sequences of λ-terms such that $M_i(N_i)$ is well-typed for each i, the pointwise application of M to N is just

$$\langle M_1(N_1), \ldots, M_n(N_n) \rangle\ .$$

Generalizing the notation for application, we denote this as $M(N)$. It is easily seen that the result of pointwise application of (7c) to (7a) equals (8a) modulo standard equivalences and that (8a)((7b)) reduces to (8b).

(8) a. $\lambda t_2.[t_2\ [\mathsf{loves\ john}]]$:
$\lambda F_2 \lambda f \exists f_1 f_2 [F_2(f_2) \wedge arc(f,cat,V) \wedge arc(f,tense,pres) \wedge arc(f_1,cat,N)$
$\wedge\, arc(f,obj,f_1) \wedge arc(f_2,cat,N) \wedge arc(f_2,num,sg) \wedge arc(f_2,pers,3) \wedge$
$arc(f,subj,f_2)]$:
$\lambda y \lambda i.love(y,john,i)$

[7] We drop the $\multimap$ in types, in conformity with the usage in semantics.

[8] For simplicity, we make no type distinction between tree nodes and feature nodes in this paper, but the conceptual distinction is important of course.

b. [mary [loves john]] :
$\lambda f \exists f_1 f_2[arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge$
$arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda i.love(mary, john, i)$

The three descriptions in sentential signs such as (8b) each denote a set in every possible model of the language; the first two sets of nodes (type νt), the third a set of possible worlds (a proposition, type st). The idea is that if the second set is non-empty in some model of the axioms in Johnson 1991 (see below), then any node satisfying the first description should be connected to the truth conditions expressed in the third element. The requirement that the second component should be satisfiable provides for a subcategorization mechanism. E.g., combining (8a) with a plural subject would have led to an f-description that can only denote the empty set.

In (9) and (10) some more lexical signs are given with two results of their possible combinations in (11).

(9) a. man : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3)$: man

b. woman : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3)$: $woman$

(10) a. $\lambda t \lambda T.T[$a $t]$:
$\lambda F \lambda \mathcal{F}.\mathcal{F}(\lambda f.F(f) \wedge arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P' P \lambda i \exists x[P'(x)(i) \wedge P(x)(i)]$

b. $\lambda t \lambda T.T[$every $t]$:
$\lambda F \lambda \mathcal{F}.\mathcal{F}(\lambda f.F(f) \wedge arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P' P \lambda i \forall x[P'(x)(i) \rightarrow P(x)(i)]$

(11) a. $\lambda T.T[$every man$]$:
$\lambda \mathcal{F}.\mathcal{F}(\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P \lambda i \forall x[man(x, i) \rightarrow P(x)(i)]$

b. $\lambda T.T[$a woman$]$:
$\lambda \mathcal{F}.\mathcal{F}(\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P \lambda i \exists x[woman(x, i) \wedge P(x)(i)]$

The terms that our signs consist of are typed, but it is expedient to type the signs themselves as well. Types for signs will be called *abstract types*. Abstract types in

abstract type	syntactic dimensions	semantic dimension
S	νt	st
NP	νt	e
N	νt	$e(st)$

Table 2: Concretizations of abstract types used in this paper.

this paper are built up from ground types S, NP and N with the help of implication, and thus have forms such as NP S, N((NP S)S), etc. A restriction on signs is that a sign of abstract type A should have a term of type A^i in its i-th dimension. The values of the function $.^i$ for ground types can be chosen on a per grammar basis and in this paper are as in Table 2. For complex types, the rule is that $(AB)^i = A^iB^i$. This means, for example, that $\text{NP}(\text{NP S})^1 = \text{NP}(\text{NP S})^2 = (\nu t)((\nu t)\nu t)$ and that $\text{NP}(\text{NP S})^3 = e(e(st))$. As a consequence, (7c) should be of type NP(NP S). Similarly, (7a) and (7b) can be taken to be of type NP, (8a) and (8b) are of types NP S and S respectively, etc. In general, if M has abstract type AB and N abstract type A, then the pointwise application $M(N)$ is defined and has type B.

Abstraction can also be lifted to the level of signs. Supposing that the variables in our logic have some fixed ordering and that the number of dimensions of the grammar under consideration is n, we define the k-th n-dimensional variable ξ of abstract type A as the sequence of variables $\langle \xi_1, \ldots, \xi_n \rangle$, where each ξ_i is the k-th variable of type A^i. The *pointwise abstraction* $\lambda\xi M$ is then defined as $\langle \lambda\xi_1 M_1, \ldots, \lambda\xi_n M_n \rangle$. A definition of *pointwise substitution* is left to the reader.

With the definitions of pointwise application, pointwise abstraction, and n-dimensional variable in place, we can consider complex terms built up with these constructions. (12a), for example, is the pointwise application of (11b) to the pointwise composition of (11a) and (7c). Here ζ is of type NP. (12a) can be expanded to (12b), where each dimension of a lexical sign is denoted with the help of an appropriate subscript (e.g. $(11b)_1$ is $\lambda T.T[\text{a woman}]$). The terms here can be reduced and the result is as in (12c), a sign coupling the c-description in its first dimension to one of its possible readings. The other reading is obtained from (12d), which reduces to (12e).

(12) a. $(11b)(\lambda\zeta.(11a)((7c)(\zeta)))$

b. $(11b)_1(\lambda\zeta_1.(11a)_1((7c)_1(\zeta_1)))$:
$(11b)_2(\lambda\zeta_2.(11a)_2((7c)_2(\zeta_2)))$:
$(11b)_3(\lambda\zeta_3.(11a)_3((7c)_3(\zeta_3)))$

c. $[[\text{every man}]\ [\text{loves}\ [\text{a woman}]]]$:
$\lambda f \exists f_1 f_2[arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge$

$arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda i \exists y[woman(y,i) \wedge \forall x[man(x,i) \rightarrow love(x,y,i)]]$

d. $(11a)(\lambda\zeta_2.(11b)(\lambda\zeta_1.(7c)(\zeta_1)(\zeta_2)))$

e. [[every man] [loves [a woman]]] :
$\lambda f \exists f_1 f_2[arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge$
$arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda i \forall x[man(x,i) \rightarrow \exists y[woman(y,i) \wedge love(x,y,i)]]$

Let us call terms such as (12a) and (12d), which are built up from lexical signs with the help of *n*-dimensional variables, pointwise application and abstraction, *n-terms*. It is worth to note that *n*-terms are subject to the laws of α, β, and η-conversion, i.e. reasoning with them is as usual. But clearly, not every *n*-term makes for an acceptable coupling between syntax and semantics. We restrict ourselves to *linear combinations* of lexical elements. These are *n*-terms that are closed and conform to the condition that every abstractor $\lambda\zeta$, with ζ an *n*-dimensional variable, binds exactly one free ζ. *n*-terms conforming to this condition are called *generated signs*.[9] Conditions such as the requirement that the third component of a generated sign must be satisfiable are *admissibility* conditions and a generated sign obeying them is called *admissable*.

Multidimensional grammars that are set up in the way sketched here, with λ-terms in each dimension of the grammar and linear combination as a generative device, will be called λ-*grammars*. If a λ-grammar is meant to be an alternative formalization of LFG insights, as the grammars in this paper will be, it is called a λ-LFG. For more information on λ-grammars, see (Muskens 2001a).

Since any *n*-term M obeys the usual laws of λ-conversion, it can be written in the form $C(L_1)\cdots(L_m)$, where $L_1,\ldots,L_m$ are lexical signs and C is an *n*-term that does not contain any lexical material. If M is closed, C is a multi-dimensional (and typed) variant of a *combinator* in the sense of (Curry and Feys 1958). In case M is a generated sign, C will correspond to a *linear* (or **BCI**) combinator. For example, (12a) can be rewritten as (13), with $\lambda Q_1 \lambda R \lambda Q_2.Q_1(\lambda\zeta.Q_2(R(\zeta)))$ playing the role of the linear combinator combining (11b), (7c), and (11a).

(13) $\lambda Q_1 \lambda R \lambda Q_2.Q_1(\lambda\zeta.Q_2(R(\zeta)))((11b))((7c))((11a))$

From the fact that linear combinators play an important underlying role we see that λ-grammars have obvious affinities not only with LFG and Lambek Categorial Grammar, but also with Combinatory Categorial Grammar (see e.g. Steedman

[9]Note that any linear combination of generated signs is itself a generated sign.

1996; Steedman 2000). But λ-grammars should be distinguished from standard categorial grammars in that they are non-directional and do not use derivations.

3 λ-LFG

The purpose of this section is twofold. First we need to fill in some details that were left open in the definition of our toy grammar. We will take a closer look at the c-description and f-description components. For the logic of the semantic component the reader is referred to the first chapters of (Muskens 1995). When the necessary details have been filled in, we will show how the descriptions approach that is taken in this paper allows for the incorporation of some further ideas that are central to LFG.

3.1 A Closer Look at C-descriptions and F-descriptions

3.1.1 C-descriptions

Terms such as $[\mathsf{loves}\ [\mathsf{a\ woman}]]$ in fact can be taken to be abbreviations of tree descriptions. We flesh this out by providing the ν domain with binary relations $\lhd^+$ (proper dominance), $\lhd$ (immediate dominance), and $\prec$ (precedence) and by imposing the necessary structure by means of axioms (see also Cornell 1994; Backofen et al. 1995; Muskens 2001b). Here we just adopt the requirements in (14). (14a) states that the relations $\lhd^+$ and $\prec$ are strict partial orders, while (14b) and (14c) impose *Inheritance*, (14d) requires *Rootedness* (r is a constant of type ν here), and (14e) defines $\lhd$ as an immediate dominance relation in terms of $\lhd^+$.[10] The last axiom excludes the possibility that leaf nodes have more than one label.

(14) a. $\lhd^+$ and $\prec$ are irreflexive and transitive

b. $\forall k_1k_2k_3\,[[k_1 \lhd^+ k_2 \wedge k_1 \prec k_3] \rightarrow k_2 \prec k_3]$

c. $\forall k_1k_2k_3\,[[k_1 \lhd^+ k_2 \wedge k_3 \prec k_1] \rightarrow k_3 \prec k_2]$

d. $\forall k\,[r \lhd^+ k \vee r = k]$

e. $\forall k_1k_2\,[k_1 \lhd k_2 \leftrightarrow \forall k_3\,[k_1 \lhd^+ k_3 \lhd^+ k_2 \rightarrow [k_3 = k_1 \vee k_3 = k_2]]]$

f. $\forall k \neg[\delta_1(k) \wedge \delta_2(k)]$, if δ_1 and δ_2 are distinct lexical labels

[10] These requirements in themselves do not suffice to axiomatize the notion of linguistic tree. For instance, the usual requirement of *Exhaustivity* ($\forall k_1k_2\,[k_1 \prec k_2 \vee k_2 \prec k_1 \vee k_1 \lhd^+ k_2 \vee k_2 \lhd^+ k_1 \vee k_1 = k_2]$) is not met. But the axioms are sufficient in the sense that the terms we generate will have a minimal model which is a tree.

Using the relations $\lhd$ and $\prec$, we define $[A_1 \cdots A_m]$ to be an abbreviation of (15).[11]

(15) $\lambda k \exists k_1 \dots k_m [A_1(k_1) \wedge \dots \wedge A_m(k_m) \wedge k \lhd k_1 \wedge \dots \wedge k \lhd k_m \wedge k_1 \prec \cdots \prec k_m]$

In other words, $[A_1 \cdots A_m]$ will denote the set of nodes with daughters $A_1, \dots, A_m$ (in that order), as expected. E.g., $[\mathsf{loves}\ [\mathsf{a\ woman}]]$ now is short for (16).

(16) $\lambda k \exists k_1 k_2 [\mathsf{loves}(k_1) \wedge k \lhd k_1 \wedge k \lhd k_2 \wedge k_1 \prec k_2 \wedge \exists k_3 k_4 [\mathsf{a}(k_3) \wedge \mathsf{woman}(k_4) \wedge k_2 \lhd k_3 \wedge k_2 \lhd k_4 \wedge k_3 \prec k_4]]$

From (16) we see that a minimality requirement is needed: As things stand the statement $[A_1 \cdots A_m](k)$ may hold in some model in which k has daughters other than $A_1, \dots, A_m$.[12] We exclude this possibility by our interpretation of generated signs. First, define the relation $\sqsubset$ between models for the tree language by letting $M' \sqsubset M$ if and only if (a) the domain of M' is a proper subset of the domain of M, (b) the root of M' is the root of M, and (c) elements of M' are in the proper dominance relation in M' if and only if they are in the proper dominance relation in M. (There is no similar constraint on precedence.) Now let $\mathcal{S}$ be a generated admissable sign with c-description component C and semantic component S. We say that a tree model M *expresses* S if M is a model of the axioms in (14) that satisfies $C(r)$ and no $M' \sqsubset M$ is both a model of the axioms in (14) and satisfies $C(r)$. The minimality condition with respect to $\sqsubset$ serves to rule out structures M that have more nodes than were intended.

Thus far, we have not put any category information into our c-descriptions. This could be done very easily by letting category labels such as AP, VP, ... be terms of type ν and by stipulating that $[_L A_1 \cdots A_m]$, where L is a category label, is short for $\lambda k(arc(k, cat, L) \wedge [A_1 \cdots A_m](k))$. However, we shall prefer treating major category information on a par with feature information in the rest of this paper.

Writing $[A_1 \cdots A_m]$ for (15) will be handy in circumstances where we want the grammar to prescribe the way in which the daughters of a given mother are ordered. But this is not always what we want. In many languages the order of daughters is not rigidly constrained. The non-configurational Warlpiri language (Simpson 1991), for example, has essentially free word order in nominal and verbal finite clauses, the only restriction being that 'auxiliaries' always take second position.[13]

[11] The use of square brackets in $[A_1 \cdots A_m]$ is special and should be distinguished from its normal use in terms.

[12] One way to rule out such undesired structures would be to strengthen the definition of $[A_1 \cdots A_m]$ with an extra condition of the form $\forall k'[k \lhd k' \to [k' = k_1 \vee \dots \vee k' = k_m]]$. However, such a strengthening would preclude the possibility of attaching extra daughters to a given node while constructing a description.

[13] AUX elements with a polysyllabic base can take first or second position (Simpson 1991). I will ignore this possibility for the sake of exposition.

In (17) it is illustrated how word order in the simple sentence *Ngarrka-ngku ka wawirri panti-rni* can vary, as long as the present tense 'auxiliary' is in second position. (The example is from (Hale 1983). We limit ourselves to three of the six possible permutations.)

(17) a. Ngarrka-ngku ka wawirri panti-rni
man-ERG PRES kangaroo(ABS) spear-NPST
The man is spearing the kangaroo

b. Wawirri ka ngarrka-ngku panti-rni

c. Panti-rni ka wawirri ngarrka-ngku

d. *Panti-rni wawirri ka ngarrka-ngku

This behaviour can be modeled by weakening the description in (15) and requiring pairwise disequalities instead of a series of precedences. Let us write $[A_1 \cdots A_m]^u$ for the following property:

(18) $\lambda k \exists k_1 \ldots k_m [A_1(k_1) \wedge \ldots \wedge A_m(k_m) \wedge k \lhd k_1 \wedge \ldots \wedge k \lhd k_m \wedge \bigwedge_{i \neq j} k_i \neq k_j]$

$[A_1 \cdots A_m]^u(k)$ will hold if k is the mother of daughters $A_1, \ldots, A_m$; but this time no ordering is prescribed.

We need a way to model the fact that AUX elements always take second position. Let us write A^{snd} for:

(19) $\lambda k[A(k) \wedge \exists k_1 k_2 [k_1 \lhd k \wedge k_1 \lhd k_2 \wedge k_2 \prec k \wedge \forall k_3 [[k_1 \lhd k_3 \wedge k_3 \prec k] \rightarrow k_2 = k_3]]]$

The term in (20) will now describe all grammatical variations of sentence (17) in the sense that any of its minimal models is an acceptable tree for one of these variations while each acceptable tree is a model.

(20) $[\text{Panti-rni wawirri ka}^{\text{snd}} \text{ ngarrka-ngku}]^u$

3.1.2 F-descriptions

The following three axioms are a direct adaptation from (Johnson 1991). The first puts a functionality requirement on the transition relation. The second embodies the constraint that atomic features have no further attributes (C_{val} stands for the set of constants denoting atomic features, such as *sg*, *past*, ...). And the third axiom schema gives constant-constant clashes by requiring that $past \neq sg$, $past \neq pres$, $V \neq N$, etc.

(21) a. $\forall a \forall f_1 f_2 f_3 [[arc(f_1, a, f_2) \wedge arc(f_1, a, f_3)] \rightarrow f_2 = f_3]$

b. $\forall a \forall f \neg arc(c, a, f)$, where $c \in C_{val}$

c. $c \neq c'$, for all syntactically distinct pairs $c, c' \in C_{val}$

For more information about these axioms the reader is referred to (Johnson 1991).

3.2 Incorporating More Ideas from LFG

3.2.1 Checking as Entailment

Thus far, information in the f-components of our signs has been of the *defining* kind, but LFG has always distinguished between defining and *constraining* equations (see Kaplan and Bresnan 1982). The first give positive information, the second an obligation to check whether certain information is present. For example, the defining equation $(\uparrow \text{INF}) = +$ sets the INF value of $\uparrow$ to $+$, whereas the constraining $(\uparrow \text{INF}) =_c +$ checks whether the INF value of $\uparrow$ is $+$ after all defining equations have been processed.

An obvious way to model this logically is to say that constraining information must be *entailed* by defining information. In (22) two of the signs in (7) are repeated, but this time there are *two* f-components. The first consists of defining information, the second of defining plus constraining material.

(22) a. john :
$\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3)$:
$\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3)$:
john

b. $\lambda t_1 \lambda t_2.[t_2 \text{ [loves } t_1]]$:
$\lambda F_1 \lambda F_2 \lambda f \exists f_1 f_2 [F_1(f_1) \wedge F_2(f_2) \wedge arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f, obj, f_1) \wedge arc(f, subj, f_2)]$:
$\lambda F_1 \lambda F_2 \lambda f \exists f_1 f_2 [F_1(f_1) \wedge F_2(f_2) \wedge arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge arc(f, subj, f_2)]$:
$\lambda x \lambda y \lambda i.love(y, x, i)$

Note that the third element of (22b) contains more material than the second. The extra statements are the ones that need to be checked. In order for this to work we need to revise the definition of *admissability* slightly. Suppose that some set of axioms such as the ones in (21) is given. A generated sign of type S is now defined to be admissable if (a) its first f-component has a non-empty denotation in some model of the axioms, and (b) the denotation of its first f-component is a subset of the denotation of its second f-component in all models of the axioms. The second condition is in fact an entailment requirement.[14] The reader will have no difficulty in seeing that agreement is now enforced. In 'John loves Mary' the required entailment holds, but only because the f-descriptions for 'John' and 'Mary' provide the material requested by 'loves'.

There is a lot of duplication in the four-dimensional signs in (22) and in practice it seems possible to use the notation in (23) where just one term is used, with subscripts *c* on some subformulas. This is just short for a four-dimensional sign. The second f-description is the one that is given (without the subscripts), but the first is obtained by deleting all subscripted material. The checking process now boils down to the requirement that subscripted material must become redundant under the usual rules of logic.

(23) a. john : $\lambda f.arc(f,cat,N) \wedge arc(f,num,sg) \wedge arc(f,pers,3)$: *john*

b. $\lambda t_1 \lambda t_2.[t_2$ [loves t_1]] :
$\lambda F_1 \lambda F_2 \lambda f \exists f_1 f_2 [F_1(f_1) \wedge F_2(f_2) \wedge arc(f,cat,V) \wedge arc(f,tense,pres) \wedge$
$arc(f_1,cat,N)_c \wedge arc(f,obj,f_1) \wedge arc(f_2,cat,N)_c \wedge arc(f_2,num,sg)_c \wedge$
$arc(f_2,pers,3)_c \wedge arc(f,subj,f_2)]$:
$\lambda x \lambda y \lambda i.love(y,x,i)$

The formalization of constraining information given here is straightforward and in the end we employ a notation that is very close to the standard one. But note that explaining the distinction between defining and constraining material as that between what is given and what must be derived essentially requires working on a level of descriptions. A structural approach, as the one in (Oehrle 1999), cannot explain the difference in this way.

3.2.2 The Long Distance: Checking F-paths

The technique developed in the previous section can also be used to characterize non-local dependencies by checking whether certain types of path exist in the functional domain. Standard LFG uses the device of *functional uncertainty* (Kaplan and Zaenen 1989) here, but it seems that the present approach can handle path checking without additional machinery. Consider Kaplan and Zaenen's rule for English topicalization in (24).

[14] It may be worthy of notice that while the treatment of agreement in (7) rested crucially on the axioms in (21), the new treatment in (22) does not in fact make use of these axioms. The situation is discussed and explained in (Johnson 1999). We like to remain agnostic about the question whether in the end axioms such as those in (21) will be needed.

(24) $S' \quad \rightarrow \quad \underset{\substack{(\uparrow\textsc{topic}) = \downarrow \\ (\uparrow\textsc{topic}) = (\uparrow\{\textsc{comp},\ \textsc{xcomp}\}^*(\textsc{gf} - \textsc{comp}))}}{\text{XP or } S'} \quad S$

The second annotation on this rule states that there must be a path from the f-structure connected to the mother S′ to that connected to the topicalized element and that the path should, top-down, start with zero or more COMP or XCOMP transitions, and then bottom out with one transition labeled by a grammatical function (SUBJ, OBJ, OBJ2,. . .) other than COMP. It is not stated how many open or closed complements are present, whence the uncertainty. The series of complements is called the *body* of the path, the final grammatical function its *bottom*.

Suppose we wanted to prove in our system that a certain path was a topicalisation path (for English). Then the following axioms would be needed.

(25) a. $\forall f f' f''[[body^{\mathrm{T}}(f, f'') \wedge bottom^{\mathrm{T}}(f'', f')] \rightarrow path^{\mathrm{T}}(f, f'')]$

b. $\forall f\, body^{\mathrm{T}}(f, f)$

c. $\forall f f' f''[body^{\mathrm{T}}(f, f') \wedge arc(f', comp, f'')] \rightarrow body^{\mathrm{T}}(f, f'')]$
$\forall f f' f''[body^{\mathrm{T}}(f, f') \wedge arc(f', xcomp, f'')] \rightarrow body^{\mathrm{T}}(f, f'')]$

d. $\forall f f'[arc(f, subj, f') \rightarrow bottom^{\mathrm{T}}(f, f')]$
$\forall f f'[arc(f, obj, f') \rightarrow bottom^{\mathrm{T}}(f, f')]$, etc., for all grammatical functions save *comp*.

The first two of these Horn statements are not particular to English and say that a topicalisation path must consist of a body and a bottom and that a body may be empty. The statements in (25c,d), on the other hand, give the particular form that bodies and bottoms in English topicalization paths must have.

In order to be able to give a treatment of topicalization using these axioms, we must at least have one complement taking verb at our disposal. In (26) a sign for *thinks* is given, very much along lines that are now familiar.[15] In combination with similar signs (26) can provide the arbitrarily long distances that some syntactic constructions can bridge.

(26) $\lambda t_1 \lambda t_2.[t_2\ [\text{thinks}\ t_1]]$:
$\lambda F_1 \lambda F_2 \lambda f \exists f_1 f_2[F_1(f_1) \wedge F_2(f_2) \wedge arc(f, cat, V) \wedge arc(f, tense, pres) \wedge$
$arc(f_1, cat, V)_c \wedge arc(f, comp, f_1) \wedge arc(f_2, cat, N)_c \wedge arc(f_2, num, sg)_c \wedge$
$arc(f_2, pers, 3)_c \wedge arc(f, subj, f_2)]$:
$\lambda p \lambda y \lambda i \forall j[B(y, j, i) \rightarrow p(j)]$

[15]The term in the semantic component of *thinks* gives a version of Hintikka's theory of belief as truth in all doxastic alternatives. Read $B(y, j, i)$ as 'world j is a doxastic alternative for y in world i'.

The sign in (27) gives a treatment of noun phrase topicalization. It has a form that requires it to be combined with (a) a sentence lacking a noun phrase which may be a generalized quantifier (type ((NP S)S)S), and (b) the lacking noun phrase (type (NP S)S). We use Z for variables of the first type, Q for those of the second. When applied to signs of such types (27) returns a sign of type S, so that its overall type is (((NP S)S)S)(((NP S)S)S).[16] The term in the first dimension of (27) preposes the syntactic material of its quantifier argument to the result of providing its first argument with a trace e. The term in the second dimension provides the Z argument with a new quantifier that essentially consists of the old one plus (a) the information that the f-structure connected with the NP is a topic of the f-structure of the result, and (b) the requirement that a topicalisation path must run from the latter to the former. The term in the semantic dimension merely copies the semantics of the Z argument.[17]

(27) $\lambda Z_1 \lambda Q_1.[Q_1(\lambda t.t) \;\; Z_1(\lambda T.T(\mathsf{e}))]$:
$\lambda Z_2 \lambda Q_2 \lambda f.$
$\quad Z_2(\lambda \mathcal{F}.\mathcal{F}(\lambda f'.Q_2(\lambda F.F)(f') \wedge arc(f, topic, f') \wedge path^{\mathrm{T}}(f, f')_c))(f)$:
$\lambda Z_3.Z_3$

As an example of how this works, consider the generated sign in (28a), which can be shown to reduce to the sign in (28b). Many requirements that have already been checked are omitted in (28b), but the crucial requirement $path^{\mathrm{T}}(f, f_1)_c$ is displayed. However, this statement can easily be derived using $arc(f, comp, f_3)$, $arc(f_3, obj, f_1)$, and the axioms in (25). It is clear that these path axioms will always allow the requirement to be satisfied as long as the defining information provides a path of the right kind.

(28) a. $(27)\Big(\lambda Q.(26)\Big(Q(\lambda \zeta.(23b)(\zeta)((7b)))\Big)\Big((7a)\Big)\Big)\Big((11b)\Big)$

b. $[[\mathsf{a\ woman}][\mathsf{john}[\mathsf{thinks}[\mathsf{mary}[\mathsf{loves\ e}]]]]]$:
$\lambda f \exists f_1 f_2 f_3 f_4[arc(f_1, cat, N) \wedge arc(f_1, num, sg) \wedge arc(f_1, pers, 3) \wedge$
$arc(f, topic, f_1) \wedge path^{\mathrm{T}}(f, f_1)_c \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge$
$arc(f_2, pers, 3) \wedge arc(f_3, cat, V) \wedge arc(f_3, tense, pres) \wedge arc(f_3, obj, f_1) \wedge$
$arc(f_3, subj, f_3) \wedge arc(f_4, cat, N) \wedge arc(f_4, num, sg) \wedge arc(f_4, pers, 3) \wedge$

[16] It seems we need to go this high in order to preserve scope possibilities. A simplified version of the sign makes do with type (NP S)(NP S):
$\lambda T \lambda t.[t \; T(\mathsf{e})] : \lambda \mathcal{F} \lambda F \lambda f.\mathcal{F}(\lambda f'.F(f') \wedge arc(f, topic, f') \wedge path^{\mathrm{T}}(f, f')_c)(f) : \lambda P.P$
This is simpler, but requires the NP to be quantified-in at the point of topicalization or higher.

[17] There are of course semantic and pragmatic consequences of topicalisation, but we ignore them here.

$$arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f, comp, f_3) \wedge arc(f, subj, f_4)] :$$
$$\lambda i \forall j[B(john, j, i) \to \exists x[woman(x, j) \wedge loves(mary, x, j)]]$$

We conclude that no extra mechanism is needed for path constraints. Path requirements are just constraining information in our approach and constraining information is modeled using entailment. But while the treatment of path requirements reduces to the treatment of constraints in general, there is still a computational difference between them and the simple agreement requirements that were met before. While the latter can be shown to be satisfied on the basis of simple properties of conjunctions, the former need reasoning in a Horn theory.

4 Conclusion

We have defined λ-grammars, a form of categorial grammar based on multidimensional signs that can be combined using linear combinators. We also hope to have shown that this form of grammar squares well with the set-up of LFG and many important ideas underlying that grammatical theory. Components of signs in λ-grammars are combined in a strictly parallel way. The linear combinators combining them provide a form of resource-sensitivity that is also present in the mathematically related linear logic. This takes care of the *Coherence* and *Completeness* requirements in LFG (approaches to LFG based on linear logic also take care of these requirements automatically). Since the various components of our grammar are terms, and since terms *describe* structure instead of providing it, we immediately reap some of the benefits of the descriptions approach to grammar. There is no difficulty with expressing negative or disjunctive information and it was shown that the distinction between defining and constraining information, important in LFG, can naturally be modeled as an entailment requirement. Constraints on long-distance paths can be modeled in essentially the same way, using a simple set of axioms characterizing acceptable paths.

The strictly parallel character has advantages, but it also has drawbacks. An advantage which the approach shares with other formalizations of LFG is that semantics need *not* be compositional with respect to surface structure. As I hope to show in a longer version of this paper, examples refuting surface compositionality, such as discontinuous adjective-noun combinations in Warlpiri and other non-configurational languages, pose no problem. A disadvantage of the strict parallelism in this paper is that the lack of communication between various components is total. Researchers in LFG have found it fruitful to mix statements about precedence and grammatical function (Bresnan 1995), but this is impossible in the present set-up. One way to allow a form of communication would be to let the various components share certain constants, but discussion about how this can be done in a sufficiently restricted way should also await the longer version of this paper.

References

Backofen, R., J. Rogers, and K. Vijay-Shankar (1995). A First-Order Axiomatization of the Theory of Finite Trees. *Journal of Logic, Language and Information 4*, 5–39.

Benthem, J. v. (1986). *Essays in Logical Semantics*. Dordrecht: Reidel.

Benthem, J. v. (1988). The Semantics of Variety in Categorial Grammar. In W. Buszkowski, W. Marciszewski, and J. v. Benthem (Eds.), *Categorial Grammar*, pp. 37–55. Amsterdam: John Benjamins.

Benthem, J. v. (1991). *Language in Action*. Amsterdam: North-Holland.

Bresnan, J. (1995). Linear Order, Syntactic Rank, and Empty Categories: On Weak Crossover. In M. Dalrymple, R. Kaplan, J. Maxwell III, and A. Zaenen (Eds.), *Formal Issues in Lexical-Functional Grammar*, Chapter 8, pp. 241–274. Stanford: Stanford University.

Cornell, T. (1994). On Determining the Consistency of Partial Descriptions of Trees. In *Proceedings of ACL-94*.

Curry, H. (1961). Some Logical Aspects of Grammatical Structure. In *Structure of Language and its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics*, pp. 56–68. AMS.

Curry, H. and R. Feys (1958). *Combinatory Logic*, Volume I. Amsterdam: North-Holland.

Dalrymple, M., J. Lamping, F. Pereira, and V. Saraswat (1999). Quantification, Anaphora, and Intensionality. In M. Dalrymple (Ed.), *Semantics and Syntax in Lexical Functional Grammar*, Chapter 2, pp. 39–89. Cambridge, MA: MIT Press.

Dalrymple, M., J. Lamping, and V. Saraswat (1993). LFG Semantics via Constraints. In *Proceedings of the Sixth Meeting of the European ACL*. European Chapter of the Association for Computational Linguistics.

Dowty, D. (1988). Type-raising, Functional Composition, and Non-Constituent Conjunction. In R. Oehrle, E. Bach, and D. Wheeler (Eds.), *Categorial Grammars and Natural Language Structures*, pp. 153–197. Dordrecht: Reidel.

Girard, J.-Y. (1987). Linear Logic. *Theoretical Computer Science 50*, 1–102.

Hale, K. (1983). Warlpiri and the Grammar of Non-configurational Languages. *Natural Language and Linguistic Theory 1*(1), 5–47.

Hendriks, H. (1993). *Studied Flexibility: Categories and Types in Syntax and Semantics*. Ph. D. thesis, University of Amsterdam.

Johnson, M. (1991). Logic and Feature Structures. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, Sydney, Australia.

Johnson, M. (1999). Type-driven Semantic Interpretation and Feature Dependencies in R-LFG. In M. Dalrymple (Ed.), *Semantics and Syntax in Lexical Functional Grammar*, Chapter 10, pp. 359–388. Cambridge, MA: MIT Press.

Kaplan, R. and J. Bresnan (1982). Lexical-Functional Grammar: a Formal System for Grammatical Representation. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, pp. 173–281. Cambridge, MA: The MIT Press.

Kaplan, R. and A. Zaenen (1989). Long-distance Dependencies, Constituent Structure, and Functional Uncertainty. In M. Baltin and A. Kroch (Eds.), *Alternative Conceptions of Phrase Structure*, pp. 17–42. Chicago: Chicago University Press.

Lambek, J. (1958). The Mathematics of Sentence Structure. *American Mathematical Monthly 65*, 154–170.

Milward, D. (1994). Non-Constituent Coordination: Theory and Practice. In *Proceedings of the 15th International Conference on Computational Linguistics, Coling 94*, Kyoto, Japan, pp. 935–941.

Montague, R. (1973). The Proper Treatment of Quantification in Ordinary English. In *Formal Philosophy*, pp. 247–270. New Haven: Yale University Press.

Moortgat, M. (1988). *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus*. Ph. D. thesis, University of Amsterdam.

Moortgat, M. (1991). Labelled Deductive Systems for Categorial Theorem Proving. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the Eighth Amsterdam Colloquium*, Amsterdam, pp. 403–423.

Moortgat, M. (1997). Categorial Type Logics. In J. v. Benthem and A. t. Meulen (Eds.), *Handbook of Logic and Language*, pp. 93–177. Elsevier.

Morrill, G. (1994). *Type Logical Grammar: Categorial Logic of Signs*. Dordrecht: Kluwer.

Muskens, R. (1995). *Meaning and Partiality*. Stanford: CSLI.

Muskens, R. (2001a). Language, Lambda's, and Logic. Submitted for publication.

Muskens, R. (2001b). Talking about Trees and Truth-conditions. *Journal of Logic, Language and Information 10*(4), 417–455.

Oehrle, R. (1988). Multi-Dimensional Compositional Functions as a Basis for Grammatical Analysis. In R. Oehrle, E. Bach, and D. Wheeler (Eds.), *Categorial Grammars and Natural Language Structures*, pp. 349–389. Dordrecht: Reidel.

Oehrle, R. (1994). Term-Labeled Categorial Type Systems. *Linguistics and Philosophy 17*, 633–678.

Oehrle, R. (1995). Some 3-Dimensional Systems of Labelled Deduction. *Bulletin of the IGPL 3*, 429–448.

Oehrle, R. (1999). LFG as Labeled Deduction. In M. Dalrymple (Ed.), *Semantics and Syntax in Lexical Functional Grammar*, Chapter 9, pp. 319–357. Cambridge, MA: MIT Press.

Simpson, J. (1991). *Warlpiri Morpho-Syntax: A Lexicalist Approach*, Volume 23 of *Studies in Natural Language and Linguistic Theory*. Dordrecht: Kluwer.

Steedman, M. (1985). Dependency and Coordination in the Grammar of Dutch and English. *Language 61*, 523–568.

Steedman, M. (1996). *Surface Structure and Interpretation*. MIT Press.

Steedman, M. (2000). *The Syntactic Process*. MIT Press.

Troelstra, A. (1992). *Lectures on Linear Logic*. Stanford: CSLI.

Zeevat, H., E. Klein, and J. Calder (1986). Unification Categorial Grammar. Manuscript.

On the Representation of French and Italian Clitics

Christoph Schwarze
University of Konstanz

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)\

2001

CSLI Publications
http://csli-publications.stanford.edu

0. Introduction

Romance clitics, and more precisely, clitic pronouns, have been a topic in generative linguistics since the seventies. For linguists like Perlmutter (1971) and Kayne (1975), their most striking property was their linear order, which is different from the order of the phrases that realize the same grammatical functions. The direct object, e.g., follows the verb when it is noun phrase, but precedes it when it is a clitic pronoun. Given the basic assumption of the generative theory of that period, accounts were proposed in terms of derivations from canonical abstract representations of the sentences containing them. In more recent work, the focus has shifted from how clitics how can be derived to how they are to be represented. New proposals on the treatment of Romance clitics have been made in the framework of optimality theory (OT, Grimshaw 1997) and of feature and exponence theories, also known under the label of Distributed Morphology (Monachesi 1999, 2000, Everett 2000, Luís & Sadler 2001).

Among these new approaches, the more radically innovative is the one known as Realizational or Distributed Morphology (DM). OT, in fact continues to consider clitics as morpho-syntactic entities, whereas DM denies them that status. Furthermore, Grimshaw's OT treatment of Italian clitics is limited in scope, since it does not comprise the whole inventory. [1] I will argue that the classical view, according to which Romance clitics are visible elements of c-structure is adequate and that there is no need to revise the classical LFG framework in order to be able to account for their grammatical properties. But under this assumption, two sets of descriptive facts still need to be accounted for, the restrictions and idiosyncrasies of clitic clusters, and the high ambiguity and syncretism of clitics.

The paper is organized as follows: In section 1, I will give a descriptive overview of those properties of Romance clitics that are a challenge to grammatical analysis. In section 2, I will discuss the DM approach to the analysis of Romance clitics. In section 3, I will show how notorious difficulties of representing Romance clitics can be resolved in the classical LFG framework.

1. The strange properties of Romance clitics

Romance clitics have a number of properties that are a challenge to grammatical analysis. They show mismatches between argument structure and constituent structure (1.1); they have a high degree of ambiguities, categorial as well as functional (1.2), and they show, in the case of Italian and Spanish, irregular phonological variation (1.3).

1.1 Mismatches between argument structure and constituent structure

Clitic clusters have an order that may differ from the order of the corresponding phrases. The default order of DIRECT and INDIRECT OBJECT (OBJECT II), when they are phrases, is "DIRECT OBJECT before INDIRECT OBJECT" (1). When the clitics are third person, the order is the same (2a), but it is opposite when the indirect object is first or second person (2b).

(1) La fée donne une bague au chevalier.
the fairy gives a ring-DO to-the knight-IO
'The fairy gives the knight a ring'.

(2) a. La fée le lui donne.
the fairy it-CLITIC-PERS3-DO him-CL-PERS3-IO gives
'The fairy gives it to him'.

b. La fée me la donne.
the fairy me-CL-PERS3-IO it-CLITIC-PERS3-DO gives
'The fairy gives it to me'.

There are asymmetries in distribution: the first and second person DIRECT OBJECT-clitics cannot form clusters with INDIRECT OBJECT-clitics (3), whereas the third person DIRECT OBJECT-clitics can (4):

(3) a. Elle *me lui présente.
she me-CL-DO-PERS1 him-CL-IO-PERS3 introduces

[1] Grimshaw 1997 considers neither *ne*, the partitive and "genitive" clitic, nor *ci* and *vi* in their spatial interpretation.

'She introduces me to him'   .

b. Elle *lui me présente.
she him-CL-IO-PERS3 me-CL-DO-PERS1 introduces
'She introduces me to him.

(4) Elle le lui présente.
she him-CL-DO-PERS3 her-CL-IO-PERS3 introduces
'She introduces him to her'.

Furthermore, the French first and second person DIRECT OBJECT and INDIRECT OBJECT clitics cannot be placed at the right-hand side of the verb unless they are followed by another clitic (5), whereas there is no such restriction for the third person clitics (6):

(5) a. *Aidez-me !
help me-CL-DO-PERS1
'Help me'!

b. Donnez-m'en !
give me-CL-IO-PERS1 from-it-CL-PARTITIVE
'Give me some'!

(6) Dis-le!
say it-CL-DO-PERS3
'Say it'!

In spoken French, one finds so-called portmanteau clusters, i.e. the merger of two grammatical functions into one form. This holds for SUBJECT and DIRECT (7) and for INDIRECT OBJECT and DIRECT OBJECT (8); the same phenomenon appears in Old French texts (9)-(10):

(7) [ʒezevy] *je les ai vus*
underlying lexical representation: /ʒez e vyz/
I-CL-SUBJ+CL-DO have seen
'I saw them'

(8) [ʒɥidire] *je le lui dirai*
underlying lexical representation: /ʒə yi dire/
I-CL-SUBJ him-CL-DO+IO will-say
'I will tell him this'

(9)

Prenez	Take
mon avoir, que vos la veez,	my belongings which you see there,
en cele male qui la pent ...	in that bag which is hanging there ...
Se ge muir portés la lou roi,	If I die, bring it to the king,
si dites que ge *li* envoi.	and say that I am sending *it to him*.

(*La Male Honte*, cit. Foulet 1958:148)

(10)	Di mei <...> u sunt voz dras ?	Tell me <...> where are your clothes?
	Dame, ceo ne dirai jeo pas	Lady, this I will say by no means.
	Kar si *jes* eüsse perduz	Because if I lost them
	E de ceo feusse aparceüz,	and would be discovered,
	Bisclavret sereie a tuz jurs.	I would be a werewolf forever.

(Marie de France, *Bisclavret*, 71ff; Rychner 1983:63)[2]

1.2 Ambiguities

From a descriptive point of view, the system of French clitics comprises twenty forms. Only eight of them (11) are unambiguous. The other twelve (12) are ambiguous in several respects:

(11) je, tu, ils, ce, la, lui, leur, les
(12) il, elle, ça, nous, vous, elles, on, me, te. se, y, en

The ambiguity is categorial for *elle, ça, nous, vous* and *elles*, i.e., these forms are clitics as well as noun phrases. This becomes visible most clearly in clitic doubling, as in (13) and (14):

(13) Elle, elle a compris.
she she-CL-SUBJ has understood
'She understood'.

(14) Ça, ça va pas.
that that-CL-SUBJ goes not
'That doesn't work'.

The ambiguity is a semantic one for the subject clitic *il*, which is a pronoun or an expletive (15).[3]

(15) a. Il chante. pronoun
'He is singing'.
b. Il pleut. expletive
'It is raining'.

Several clitics are ambiguous with respect to the grammatical function that they represent. Some (*me, te, se, nous, vous*) are DIRECT OBJECT as well as INDIRECT OBJECT; cf. (16):

(16) a. Elle me verra.
she me-CL-DO will-see

[2] Similar portmanteau clitics occur in Portuguese; cf. Luís & Sadler (2001:2).
[3] Other clitics may be expletive in idioms: *le* in *l'emporter* 'to gain the upper hand', *en* in *en vouloir à quelqu'un* 'to be angry with', *se* in *s'en aller* 'to leave', *y* in *ne pas y aller de main morte* 'to exaggerate'. I do not discuss the representation of idioms in this paper.

'She will see me'.

b. Elle me parlera.
she me-CL-IO will-speak
'She will speak to me'.

Y is an Oblique (17) or an Adjunct (18):

(17) J'y ai pensé.
I Y-CL-OBLIQUE-INANIMATE have thought
'I have thought of it'.

(18) Je l'y ai vu.
I le-CL-DO Y-CLITIC-SPATIAL-ADJUNCT have seen
'I saw him there'.

En is an Oblique (19) or a Partitive Modificator of the DIRECT OBJECT (20):

(19) Je leur en ai parlé.
I them-CL-IO of_it-CL-OBL have spoken
'I spoke to them about it'

(20) Elle en aime un autre.
she of_them-CL-PARTITIVE an other
'She loves someone else'

And *le* is not only a DIRECT OBJECT (21) but also a Complement of the copula (22):

(21) Je le verrai.
I him-CL-DO will-see
'I will see him'.

(22) Je le serai.
I it-CL-XCOMP will-be
'I will be'.

And the clitics of the first and second person are personal (23) as well as reflexive pronouns (24):

(23) Je te défendrai.
I you-CL-PERSONAL will-defend-PERS1
'I will defend you'.

(24) Tu te défendras.
You you-CL-REFLEXIVE will-defend-PERS2
'You will defend yourself'.

1.3 Irregular phonological variation

Some clitics show irregular phonological variation, i.e. alternations of phonological shape which are not due to postlexical phonology. In French, the vowel of the feminine singular object clitic *la*, just as the homophonous determiner, is elided before vowel (25), but there is no rule of *a*-elision (26):

(25)	Il l'aime 'He loves her'.	*Il la aime	Elision
(26)	Mon énergie 'My energy'.	*m'énergie	No elision

In Italian, the situation is similar (27):

(27)	a. Lui l'ama 'He loves her'.	?lui la ama	Elision
	b. Alta Italia 'Upper Italy'.	*alt'Italia	No elision

In Italian, *mi*, *ti*, *si* and *ci*, when they precede a direct object-clitic, lower /i/ to /e/ (28), a phenomenon which does not appear anywhere else in that language:

(28) a. *Mi* vedi?
'Can you see me'?

b. *Me* lo dici?
'Will you say it to me'?

And, finally, dissimulation is compulsory when two clitics are combined which have the same or almost the same phonological shape: in Italian, *si*+*si* appears as *ci si* (29), and in Spanish, *le*+*lo* is replaced with *se lo* (30):

(29) a. Si capisce.
SI-CL-SUBJ understands-SG
'One understands'.

b. Si capiscono.
SI-CL-RECIPROCAL understand-PL
'They understand each other'

c. Ci si capisce. *Si si capisce.
SI-CL-SUBJ SI-CL-RECIPROCAL understand-SG
'One understands each other'

(30) a. Le respondré mañana.
him-CL-IO will-answer-PERS1 tomorrow
'I will answer him tomorrow'

b. Lo digo.
it-CL-DO say-PERS1
'I say it'

c. Se lo digo. - *Le lo digo.

him-CL-IO it-CL-DO say-PERS1
'I say it to him'

2. Treating clitics as exponents

In the presence of such an amount of intricate descriptive facts, DM appears to be an ideal approach. This idea may seem to be paradoxical in a lexicalist framework such as LFG, given that in DM there is no such thing as a lexicon in the classical sense (Harley & Noyer 1999:3). But one could plausibly think of a syncretistic format, where only certain grammatical domains are represented in terms of mapping from feature bundles to phonological forms. One could consider that the Romance clitic systems are one of these domains. Syntax, then, does not have to account for the ambiguities and variations of the Romance clitic systems; it just has to provide feature configurations for clitics, identify their hosts, map them onto phonological strings, and put them in the correct linear order with respect to their host. In a sense, if such a treatment is chosen, it can actually be said that "there are no clitics" (Everett 2000).

Moreover, the treatment of Romance clitics as phonological exponents is no longer a mere program. Monachesi (1999) has shown how the Italian clitic system can elegantly be represented in HPSG on a DM base. All the problems mentioned above seem to be solved, included the treatment of clitic clusters. By treating them as complex objects at the level of features, and as unanalyzable strings at the level of surface expression, Monachesi (1999:76) not only captures their idiosyncratic restrictions, but also reduces their ambiguities to a large extent. In fact, the ambiguities which clitics show in isolation disappear in clitic clusters. Thus, those clitics that alternatively represent DIRECT and INDIRECT OBJECTS (*mi*, *ti*, *si* etc.), are unambiguous in clitic clusters, and *lo* is always a DIRECT OBJECT, never an XCOMP, when it appears in a cluster.[4]

[4] Another analysis of Romance, more precisely Portuguese, clitics, based on DM principles, is contained in Luís & Sadler 2001. The Portuguese clitics show an intricate behavior regarding their precedence relationship with and the identity of their hosts. If, unlike these authors, one treats morphological segments as constituents, then precedence relations are expressed by c-structure rules. This is doubtlessly feasible.

The lexicalist the representation of clitic clusters which I will propose in section 3 differs from Monachesi’s regarding the architecture of the grammar, but it seems to be logically equivalent if one looks at the relationship between the features and their realizations (“morphological segments“ in my terminology). In fact, Monachesi represents, e.g. It. *gli*, the masculine third person dative clitic pronoun, as (31):[5]

(31) [STEM … CLTS <dat, pers3, sg, masc>] →
[AFFIX [PHON <*gli*>]]

And she explains this notation with the following words: "It states that if there is a STEM with a CLTS list with one element which is a third person singular, masculine, dative NP, it must be realized as a clitic whose phonological form is *gli*." (Monachesi 1999:73).

A traditional LFG lexical entry would be (32):

(32) *gli*, CL
(↑ PRED)=PRO, (↑ GEN)=MAS, (↑ NUM)=SG, (↑ PERS)=3, (↑ PCASE)=A

This notation states that there is a morphological segment *gli*, which is a clitic and encodes the features pronoun, masculine, singular, third person, and dative case. The only difference, then, seems to consist in how constituency is treated in LFG and HPSG: in LFG, *gli* is some kind of constituent (labeled CL), whereas in Monachesi’s HPSG analysis, it is not a constituent of any kind, but just a phonological string.

Now the existence of elegant and descriptively satisfying analyses like Monachesi's do not necessarily prove the adequateness of the underlying theoretical claim. And, in fact, there are a number of objections to the DM approach, which I will address now, going from the general to the specific.

The most general objection is that there is no need to import DM into LFG, since the stratified representations of LFG make it possible to map feature structures onto word-forms which have no morphological structure, i.e., are not analyzable in c-structure. DM may well represent an important progress with respect to theories where form and function must be isomorphic, in such a way that unanalyzable word-forms must be derived

[5] I have omitted those elements of the representation which are not relevant to my argumentation and made some abbreviations more readable.

from underlying analyzable forms, as, e.g. *went* from *go+ed*. But LFG, just like DM, directly relates word-forms and feature configurations, no matter whether the word-forms have an internal constituency or not. In other words, the problem that is solved by DM, is also solved in LFG.

A more specific objection is that current presentations of the DM-model are formulated in such a way that it is not obvious what the phonology looks like. Taken at face value, DM excludes all kinds of Lexical Phonology, since there is no lexicon. One could of course explain that the items, which are spelled out by the realization rules of DM, are precisely the "lexical representations" to which postlexical phonology applies. But if affixes are not "visible" in any part of the grammar, one may lose important generalizations. As an example, take Italian thematic vowels, such as /i/ in [sen'tiːte]*sentite* 'you hear', [sen'tiːvo]*sentivo* 'I heard', [sen'tiːto]*sentito* 'heard (participle)'. The syllables in which they occur are candidates for stress assignment, in spite of their "light" linear structure. If the phonology does not know that they are thematic vowels, this factor of the stress assignment system is reduced to a huge list of arbitrary spell-outs.

If this may still be accepted in merely structural model of phonology, the situation becomes dramatic if one passes from the structural to the procedural point of view. In the FUL-model of speech perception (Preuss et al. in press), acoustic properties are transformed into phonological features and the phonological feature structures are mapped onto abstract lexical representations. The lexical representations of complex, analyzable words are strings of morphological segments, e.g. stems and affixes, and these segments are listed in the lexicon[6]. It seems hard, in fact, to see how phonological features could be mapped directly onto morpho-syntactic features.

Furthermore, there are certain descriptive facts that are hard to treat in terms of DM. One of them is the derivation of manner adverbs from adjectives in Italian. The general structure of these adverbs is (33):

(33) the feminine singular form of an adjective + *mente*

---

[6] Actually, Preuss et al. (in press) use the term "morpheme", but in their use that term does not imply the notion of Saussurian sign.

The Italian adjective has two distinct inflectional classes; cf. (34):

(34) The *o*-Class

	Sing.	Pl.
Masc.	***o***	***i***
Fem.	***a***	***e***

The *e*-Class

Sing.	Pl.
e	***i***

Hence, the feminine singular form of *o*-Class adjectives ends with -*a*, (e.g. *rapida*) and the singular form of *e*-Class adjectives ends with -*e* (e.g. *cortese*). Accordingly, one gets manner adverbs like the following:

(35) a. rapidamente 'fast'
b. cortesemente 'politely'

Now, how are forms like these derived in DM? The DM-type morphological representations of *rapida* and *cortese* are, respectively:

(36) rapid + [GEN = FEM, NUM = SG]
cortes + [NUM = SG]

Should *mente* be affixed to these representations, in such a way that one gets representations like the ones in (37)?

(37) rapid + [GEN =FEM, NUM = SG]+mente
cortes + [NUM = SG]+mente

Under this treatment the representations in (37) must then be mapped to phonology. But since there must also be a mapping relationship for the adjectives, there would be an unnecessary doubling.

Or should the adjectives first be realized by mapping them onto phonology, like in (38)?

(38) rapid + [GEN = FEM, NUM = SG] ⇒ /'rapida/
cortes + [NUM = SG] ⇒ /kor'teze/

The derivational suffix must then be adjoined to the spell-outs at the right-hand side of the double arrow. But this would create monsters, made up of a phonological representation and a morphological segment.

As a consequence, it will certainly be more adequate to regard the suffixes of Italian adjective inflection as visible morphological segments. And if derivational affixes are visible in the lexicon, why should other morphological segments not be visible as well?

There also is an argument specific to the data treated in the present paper. As has been pointed out above (3), certain clitic clusters which are licensed by argument structure, are ungrammatical. In a DM treatment, one has to rule out those feature configurations for which there are no phonological exponents. Monachesi (1999:80) postulates a "Clitic Realization Principle", which states that

> "All verbal hosts which have a nonempty CLTS list must satisfy one of the realizational constraints".

But these constraints are nothing else than lists of feature configurations paired with phonological forms. If, on the other hand, clitic clusters are stored in the lexicon as morphological segments to which feature configurations are associated, then the non-existence of some segments does not come as a surprise. Idiosyncratic restrictions on distribution are irregularities, comparable to those of defective inflectional paradigms. And were else than in the lexicon should irregularities be stored?

To conclude, it is not obvious that DM should be an advance with respect to LFG. Furthermore, it seems to have some undesirable consequences. Even if the question of how the syntax-morphology and the morphology-phonology interfaces are best modeled within the LFG framework is still open, DM does not seem to impose itself for the topic at hand, namely the treatment of Romance clitics. I will therefore maintain the hypothesis that clitics are morphological segments to which feature structures are associated, and I will try to show how some of the crucial descriptive facts can be accounted for without changing LFG.

3. Using the LFG framework for representing Romance clitics

I will now give a sketch of the c-structure analysis of Italian and French clitics (3.1) and make a proposition regarding the relationship between morphological case and grammatical function (3.2). I will then address three of the challenges which Romance clitics present for a treatment in terms of "visible" lexical items. Two of them concern functional ambiguities, namely ambiguities of grammatical function (3.3) and ambiguity between reflexivity and non-reflexivity (3.4). The third concerns the representation of clitic clusters (3.5). I will close with an outlook at the problems inherent in Italian impersonal *si* and French first plural *on* (3.6).

3.1 The c-structure analysis of Italian and French clitics

Since Grimshaw (1982), French clitics are treated as postlexical affixes to the verb. Even though she does not use this term, the idea is clearly expressed by the c-structure rule she proposes (Grimshaw 1982:90), and which I reproduce here as (39) in a simplified version:

(39) V' $\rightarrow$ (CL) (AUX) V

The question is, of course, whether this is a correct analysis. Clitics obviously resemble affixes. This holds especially for those cases in which the clitic is placed to the right-hand side of its host. Luís & Sadler (2001:4) show that, in Portuguese, the enclitic forms show morphological interaction with the verb. This also holds for Italian: If the verb form is an infinitive, the clitic is attached to its right-hand side, and the infinitive suffix /re/ loses its final /e/; cf. (40):

(40) fare 'to do' - farlo 'to do it'

This is a morphological alternation that also occurs in Italian word formation. In fact, adverb formation with *–mente*, which inactivates the inflectional features of the adjectival basis, deletes the corresponding morphological segment if the stem ends with /l/ or /r/; cf. (41):

(41) tale 'such a' talmente 'so'
regolare 'regular' regolarmente 'regularly'

But the fact that e-deletion is also phonologically conditioned, raises suspicion. Moreover, e-deletion also is required in phrases; cf. (42):

(42) dottore 'doctor' dottor Rossi 'Dr. Rossi'
potere 'to be able' per poter farlo 'in order to be able to do that'
bene 'well' ben tre volte 'three times'[7]

[7] *Three* pronounced with a stress, expression that three times is more than enough.

It might well be that, at least regarding Italian e-deletion, the interaction of the clitic with the verb takes places in phonology; this is an open question. As far as French is regarded, the verb form never seems to be affected by the presence of a clitic. I will conclude that there is no reason to deny that clitics are inserted postlexically. Their similarity with affixes can probably best be expressed, rather than at the c-structure level, by the representation of their functional properties.

I will thus follow Grimshaw's (1982) general idea[8]. The rule for Italian will be (43), if the sentence is declarative and the verb (auxiliary or full verb) is finite:

(43) V' → (CL1) (CL2) V

Notice that in this treatment, CL is thought to include clitic clusters as well as single clitics. CL1 is the negation particle *non*, which needs a verb as a host. The category CL2 functionally comprises objects, obliques and adjuncts. Unlike French, Italian has no subject clitics. The distinction, at the c-structure level, between two kinds of clitics is necessary because the negation particle always is at the left-hand side of the verb.[9]

For non-negated imperative sentences in the second person and for non-finite forms of the verb, the order is inverted:

(44) V' → (CL1) V (CL2)

For French, the situation is more complex, inasmuch as there is one more type of clitic forms: the subject clitic (CL1), which always precedes the negative particle *ne* (CL2), and which differs in distribution from the other

---

[8] For the representation of clitic clusters, Grimshaw (1982) makes the rule repeat the category CL, assigning each CL a position:

V' → (CL)1 (CL)2 (CL)3 (AUX) V

A different proposal for clitic clusters will proposed here. When, further on, I will use the notations like CL1, CL2, CL3, this will be meant to express not different positions, but different categories of clitic forms.

[9] Monachesi (2000) makes an analogous distinction for Romanian. According to her analysis, there are two kinds of clitics, those, which have word status, and those, which are affixes, and she derives the ordering of clitics from that difference of status. My own distinction is only motivated by facts of distribution.

pronominal clitics (CL3). The rule for French declarative sentences is (45), regardless of whether the verb is finite or not:

(45) V' $\rightarrow$ (CL1) (CL2) (CL3) V

In imperative sentences, there is no subject clitic, and the other pronominal clitics are placed at the right-hand side of the verb:

(46) V' $\rightarrow$ (CL2) V (CL3)

In French interrogative sentences[10] the clitic subject follows the verb; the rule is thus (47):

(47) V' $\rightarrow$ (CL2) (CL3) V (CL1)

These are the facts that justify the assumption of three clitic categories in French, as opposed to only one in Italian.

3.2 Morphological case and grammatical function

In Grimshaw 1982 French clitics are lexically assigned a case attribute with the values "accusative" and "dative"; Monachesi 1999 uses the same case labels for Italian, adding "locative" for *ci*, and one could use "genitive" for French *en* and Italian *ne*, and "nominative" for the French subject clitics. Grimshaw then maps cases onto grammatical function in the c-structure annotations.

It is not obvious, however, why one should have case features in the analysis of languages like French and Italian. It is true that, historically, the clitic forms are continuations of Latin case forms. But this is not a reason to postulate case as a feature in modern Romance. Since no other category than clitics have case, it is simpler to directly assign grammatical functions to clitics in the lexicon and to write entries such as (48) and (49) for French:

[10] In French, interrogation can also be realized by declarative sentences with an interrogative intonation or an interrogative pronoun.

For Italian, which has no subject clitics, (48) has no equivalent, and the category of (46) would just be CL.

(48) *je*, CL1
(↑ PRED) = 'PRO'
(↑ PERS) = 1
(↑ NUM) = SG
(↑ SUBJ)

(49) *la*, CL3
(↑ PRED) = 'PRO'
(↑ PERS) = 3
(↑ NUM) = SG
(↑ GEN) = FEM
(↑ OBJ)

This representation, in a model of grammar that has grammatical functions, eliminates a useless step of computation.

Moreover, there are two types of clitic doubling. The clitic may realize some grammatical function, and the constituent with which it is correferential, be a TOPIC. The clitic may also, instead of realizing a grammatical function, be just an agreement marker, as in Italian (50) and French (51).[11]

(50) Me la dai una mela?
me-CL-IO it-CL-OBJ_AGR an apple-OBj
'Do you give me an apple'?

(51) Il est sympa ce garçon.
he-CL-SUBJ_AGR is nice this boy-SUBJ
'That's a nice boy'.

This use of clitics as agreement markers can be represented more easily if the grammatical function of the clitic is directly encoded in the lexicon. In order to do this, all that it takes is to revise entries like (48) and (49) above in such a way that instead of grammatical functions, they encode agreement with grammatical functions; cf. (52), the revised formulation of (49):

(52) *la*, CL3
(↑ OBJ PRED) = 'PRO'
(↑ OBJ PERS) = 3
(↑ OBJ NUM) = SG
(↑ OBJ GEN) = FEM

[11] The criterion for the latter structure is intonation: when a clitic is used as an agreement marker, the typical TOPIC intonation of the related noun phrase is lacking. The clitics, which appear in this construction, seem to be restricted. Typically they are the object and, in French, the subject clitic. Whether other clitics are used as mere agreement markers is subject to further investigation.

This entry accounts for both, the use of the clitic as an agreement marker and as the realization of a grammatical function. It creates an object that is merged with the object of the verb if there is one, and alternatively becomes the object if there is no other object.[12]

3.3 Ambiguities of grammatical function

Some clitics show ambiguities with regard to the grammatical function they realize which can only be handled by lexical listing. This is the case, e.g., for French *y*. This clitic may be an oblique with PCASE A (53), a local OBLIQUE (54), and a local adjunct (55):

(53) J'y pense.
I it-CL-IO think
'I think of it.'

(54) J'y vais.
I there-CL-OBL go
'I go there.'

(55) J'y dors.
I there-CL-ADJUNCT sleep
'I sleep there.'

Three entries are needed for (53)-(55):

(56) *y*, CL3
(↑ OBL PRED) = 'PRO'
(↑ OBL PCASE) = A
(↑ OBL ANIMATE) = —

(57) *y*, CL3
(↑ OBL PRED) = 'PRO'
(↑ OBL LOC) = +

(58) *y*, CL3
(↑ ADJ PRED) = 'PRO'

[12] Halpern & Fontana (1994) make a distinction between two kinds of clitics, which they name X° and X^{max} clitics. Roughly speaking, X° clitics support inflection and agreement, whereas X^{max} clitics represent arguments. This distinction is not adequate for Romance, where clitic doubling is functionally ambiguous. Diachronically, clitics develop from argumental phrases. After phrases have turned into clitics, they may keep their status as arguments or lose it. The evolution of Romance clitics seems to have reached a stage where both kinds of functional status, argument and mere agreement marker, are synchronically present.

(↑ ADJ LOC) = +

But for other clitics, generalizations are possible, and they can be expressed via lexical rules. The first and second person clitics *me, te, nous, vous* are both, DIRECT OBJECT and INDIRECT OBJECT. The same holds for the reflexive *se*. This can be expressed by encoding only one of the functions and derive the other. The choice is arbitrary in principle, but the lexical rule can be formulated more simply if the encoded function is the INDIRECT OBJECT. The rule then reads as follows:

(59)	*x*, CL3 (↑ OBL PRED) = 'PRO' (↑ OBL PERS) = {1,2} (↑ OBL PCASE) = A	⇒	*x*, CL3 (↑ OBJ PRED) = 'PRO' (↑ OBJ PERS) = {1,2}

3.4 The ambiguity between reflexivity and non-reflexivity

For reflexive constructions, French and Italian, like other languages, e.g. German, have a formally unambiguous form only for the third person; in reflexive constructions in the first and second person, the personal pronoun is used.[13] But while, in German, the retrieval of first and second person reflexivity can be left to semantics, French and Italian need to make reflexivity explicit at the f-structure level for all persons for the sake of auxiliary selection in the compound tenses. In fact, the representation of sentences like French (60a) and Italian (61a) must make a feature "reflexive" available, analogously to the b-sentences:

(60) a. Tu t'es trompé.
you-CL-SUBJ-PERS2-SG you-CL-OBJ-PERS2-SG are misled
'You made a mistake.'

b. Il s'est trompé.
he-CL-SUBJ-PERS3-SG-MAS him-CL-REFL-OBJ-PERS3-SG is misled
'He made a mistake.'

(61) a. Mi sono fatto la barba.
me-CL-OBL-PERS1-SG am-PERS1-SG made the beard
'I shaved.'

b. Si è fatto la barba.
him-CL-REFL-OBL-PERS3-SG is-PERS3-SG made the beard

[13] This is probably due to the fact that semantic reflexivity can be derived from the correferentiality of the subject and the direct or indirect object if the subject is in the first or second person.

'He shaved.'

Therefore the personal pronouns of the first and second person must also be represented as reflexive pronouns by a feature, which I will write as "REFLEXIVE = +". One might think of expressing this ambiguity via a lexical rule, comparable to the one I proposed above for the DIRECT OBJECT-INDIRECT OBJECT ambiguity. But since the REFLEXIVE feature is needed for personal pronouns only in sentences with compound tenses, the ambiguity can be represented more adequately as a conditional constraint, i.e. as an inference drawn at f-structure. It can be formulated as (59):

(62) ***Reflexive inference***

SUBJ PERS = {1 ,2} & {OBJ, OBL} AGR = SUBJ AGR $\Rightarrow$ {OBJ, OBL} REFL = +

For illustration, consider sentence (63) and its f-structure (64), such as it is built up from the annotated c-structure tree:

(63) Je me l'explique.
I me-CL-OBL-PERS1 it-CL-OBJ-PERS3 explain
'I can see why this is so.'

(64)

$$\left[\begin{array}{ll}
\text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{'PRO'}\\ \text{PERS} & 1\\ \text{NUM} & \text{SG}\end{array}\right]\\[2ex]
\text{PRED} & \left[\begin{array}{l}\text{'EXPLIQUER <(SUBJ), (OBJ), (OBL)>'}\\ \text{(OBL PCASE) A}\end{array}\right]\\[2ex]
\text{OBJ} & \left[\begin{array}{ll}\text{PRED} & \text{'PRO'}\\ \text{PERS} & 3\\ \text{NUM} & \text{SG}\\ \text{GEN} & \text{MAS}\end{array}\right]\\[2ex]
\text{OBL} & \left[\begin{array}{ll}\text{PRED} & \text{'PRO'}\\ \text{PERS} & 1\\ \text{NUM} & \text{SG}\\ \text{PCASE} & \text{A}\end{array}\right]
\end{array}\right]$$

F-structure (64) meets the conditions for applying Reflexive Inference. When Reflexive Inference is applied, the resulting f-structure is (65). It differs from (64) insofar as the feature "REFL = +" is included in the functional description of the OBLIQUE.[14]

(65)

$$
\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PERS} & 1 \\ \text{NUM} & \text{SG} \end{bmatrix} \\
\\
\text{PRED} & \begin{bmatrix} \text{'EXPLIQUER <(SUBJ), (OBJ), (OBL)>} \\ \text{(OBL PCASE) A} \end{bmatrix} \\
\\
\text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEN} & \text{MAS} \end{bmatrix} \\
\\
\text{OBL} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PERS} & 1 \\ \text{NUM} & \text{SG} \\ \text{PCASE} & \text{A} \\ \textbf{REF}\text{ L} & \mathbf{+} \end{bmatrix}
\end{bmatrix}
$$

3.5 The representation of clitic clusters

As has been said above, a large amount of the difficulties of the French and Italian clitic systems can be avoided by representing clitic clusters as forms that are analyzed only at the f-structure level. The inventory of the items belonging to category CL2 in Italian and CL3 in French will therefore comprise, along with the simple clitics, the clitic clusters. We will thus account for the rigidity of linear order, for the restrictions on co-occurrence, for the merger of two arguments into one clitic and for phonological variation in the simplest way. The notation introduced above for representing simple clitics turns out to be most useful to keep the two sets of

[14] There is nothing, however, in (60), which triggers the application of Reflexive Inference. In a more explicit formulation it should be expressed that Reflexive Inference takes place only of the value of tense is the name of a compound tense.

features apart. As an example, consider the lexical representation of *me la* (66); merged arguments can be represented in exactly the same way (67):

(66) *me la*, CL3
(↑ OBL PRED) = 'PRO'
(↑ OBL PCASE) = A
(↑ OBL NUM) = SG
(↑ OBL PERS) = 1
(↑ OBJ PRED) = 'PRO'
(↑ OBL NUM) = SG
(↑ OBL GEN) = FEM
(↑ OBL PERS) = 3

(67) /yi/, CL3
(↑ OBJ PRED) = 'PRO'
(↑ OBJ NUM) = SG
(↑ OBJ GEN) = MAS
(↑ OBJ PERS) = 3
(↑ OBL PRED) = 'PRO'
(↑ OBL PCASE) = A
(↑ OBL NUM) = SG
(↑ OBL PERS) = 3

3.6 *Italian impersonal* si *and French first plural* on

In Monachesi (1999) impersonal *si* is considered a subject. If this is a correct analysis, which I will assume for the following discussion, then sentences like (68 b), are a problem for HPSG. Monachesi (1999:98) resolves it by a lexical rule, which removes the subject "from the subject list of the verb" and adds it "as a member of the CLTS list."

(68) a. Si mangia spesso le fragole.
SI eats often the strawberries
'One often eats strawberries.'

b. *La gente si mangia spesso le fragole
the people SI eats often the strawberries
'The people one eats often strawberries.'[15]

In an LFG analysis, this problem does not even appear: sentences like (68 b) are taken care of by the uniqueness constraint; whereas HPSG has to postulate an ad hoc rule.

But impersonal *si* presents still another problem: in sentences with an XCOMP, the XCOMP is in the plural, whereas the verb is in the singular:

[15] Monachesi's glossing and translation. Instead of *one*, *we* could also be used as a translation of impersonal *si*.

(69) Si è contenti.
SI is-SG happy-PL-MAS
'One is happy.'

Monachesi (1999:99ff) proposes a solution based upon a sophisticated theory of agreement, proposed by Pollard and Sag (1994) for French sentences like (70) and further developed by Kathol (1999).

(70) Vous êtes belle.
you-PL are-PL beautiful-SG
'You are beautiful.'

This theory uses indices in such a way that agreement constraints can be stated with respect to an index rather than between two inflected forms. In LFG, one might think of treating the problem in the following way (which I could not elaborate here): The agreement between the verb and its subject is standard. The XCOMP, however, has a (covert) subject of its own, which is only referentially identical with the subject of the sentence. This covert subject must be plural when impersonal *si* is the subject of the verb, and it can be both genders:

(71) Si è contente.
SI is-SG happy-PL-FEM
'One is happy.' (said of women)

An analogous solution is desirable for the French subject clitic *on*, which, in the spoken language, is semantically first person plural and agrees as such with its XCOMP (72) or Past Participle (73), but requires the third person singular for the verb:

(72) On est sympathiques.
ON is-SG nice-PL
'We are nice.'

(73) On est partis.
ON is-SG left-PL
'We left.'

The elaboration of a more expressive theory of agreement for LFG is a task for further research.

4. Clitic pre-emption

The phenomena that Sadler (1997:11) subsumes under the term of clitic pre-emption, concern an aspect of the grammar of clitics which is only rarely addressed in the literature. Regarding the languages treated in the present paper, the facts are the following.

In French, if an object or an oblique is realized by a pronoun, this pronoun must be a clitic; cf. (74):

(74) a. Il t'aime.
he CL-OBJ-PERS2-SG loves

b. *Il aime toi
he loves NP-OBJ-PERS2-SG

The non-clitic pronoun is possible only in clitic-doubling:

(75) Il t'aime, toi.
he CL-OBJ-PERS2-SG loves NP-OBJ-PERS2-SG
‚he loves you' (main stress on *you*)

Sadler (1997) observes similar facts for Welsh and explains them as the consequence of a general blocking principle. A similar solution may be formulated for French.

But regarding Italian, the situation is slightly different: The cliticity of the pronoun, rather than being necessary, is only the unmarked choice (75a). The non-clitic pronoun is not blocked, it is only marked; cf. (75b):

(75) a. Mi piace.
me-CL pleases
‚I like it'

b. A me piace.
to me-NP pleases
‚I like it' (stress on *I*)

The blocking and markedness effects that clitics may produce deserve to be carefully investigated, also with respect to other languages. But it is very likely that the underlying principles must refer to hierarchies in the lexicon. This is one more reason for a lexicalist treatment of clitics.

5. Conclusion

I hope to have shown that an adequate representation of French and Italian clitics is possible without changing the general LFG framework. It is true that there remain some points which need further reflection, and, which is more important, that the data are only from two languages. But the fact that such grammatical objects as clitics, which are on the borderline between syntax and morphology speaks in favor of the robustness of the framework. Given its week theoretical commitments, LFG makes it possible to develop theories without changing the overall framework, and it certainly is a wise research strategy to take advantage of such robustness not to modify the general framework unless it is necessary.

References

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Everett, Daniel L. 2000. "Why There Are no Clitics. On the Storage, Insertion, and Form of Φ-Features". in: Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.). *Lexical Specification and Insertion*. Amsterdam/Philadelphia: John Benjamins. pp. 91-114.

Foulet, Lucien. 1958. *Petite syntaxe de l'ancien français*. 3me éd. revue. Paris: Champion.

Grimshaw, Jane. 1982. "On the lexical representation of Romance reflexive clitics." In: J. Bresnan (ed.) *The mental representation of grammatical relations*. Cambridge, Massachusetts: MIT Press, pp. 87-148.

Grimshaw, Jane. 1997. "The best clitic: constraint conflict in morphosyntax." In: Liliane Haegeman (ed.) *Elements of grammar*. Dordrecht: Kluwer, pp. 169-196.

Halpern, Aaron and Josep M. Fontana. 1994. "X° and X^{max} Clitics". In: Erin Duncan et al. *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI.

Harley, Heidi and Rolf Noyer. 1999. "State-of-the-Art Article: Distributed Morphology". *GLOT* 4.4. pp. 3-9.

Kathol, Andreas. 1999. "Agreement and the syntax-morphology interface in HPSG." In: R. Levine and G. Green (eds.). *Readings in HPSG*. Cambridge: Cambridge University Press.

Kayne, Richard. 1975. *The Transformational Cycle*. The MIT Press.

Luís, Ana and Louisa Sadler. 2001. *Morphology and Markedness: Porttuguese Object clitics*. Handout of a paper given at the Department of Linguistics of the University of Konstanz on Tuesday, September 11, 2001.

Monachesi, Paola. 1999. *A lexical approach to Italian cliticization*. Stanford, California: CSLI Publications.

_________. 2000. “Clitic placement in the Romanian verbal complex”. In: Birgit Gerlach and Janet Grijzenhout (eds.). *Clitics in Phonology, Morphology and Syntax*. Amsterdam / Philadelphia: John Benjamins.

Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Pollard, C. and Sag, Ivan. 1994. *Head-Driven Phrase Structure Grammar*. Chicago - Stanford: University of Chicago Press and CSLI Publications.

Preuss, Letizia, Marslen-Wilson, William and Lahiri, Aditi. in press. "Abstractness in stem representation: Phonology-Morphology interaction in processing". *Journal of Memory and Language*.

Rychner, Jean (ed.). 1983. *Les Lais de Marie de France*. Paris: Champion.

Solving Headswitching Translation Cases in LFG-DOT

Andy Way

School of Computer Applications,
Dublin City University,
Dublin, Ireland.

Email: away@compapp.dcu.ie

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

It has been shown that LFG-MT (Kaplan *et al.*, 1989) has difficulties with Headswitching data (Sadler *et al.*, 1989, 1990; Sadler & Thompson, 1991). We revisit these arguments in this paper. Despite attempts at solving these problematic constructions using approaches based on linear logic (Van Genabith *et al.*, 1998) and restriction (Kaplan & Wedekind, 1993), we point out further problems which are introduced.

We then show how LFG-DOP (Bod & Kaplan, 1998) can be extended to serve as a novel hybrid model for MT, LFG-DOT (Way, 1999, 2001), which promises to improve upon the DOT model of translation (Poutsma 1998, 2000) as well as LFG-MT. LFG-DOT improves the robustness of LFG-MT through the use of the LFG-DOP *Discard* operator, which produces generalized fragments by discarding certain f-structure features. LFG-DOT can, therefore, deal with ill-formed or previously unseen input where LFG-MT cannot. Finally, we demonstrate that LFG-DOT can cope with such translational phenomena which prove problematic for other LFG-based models of translation.

1 Headswitching in LFG-MT

Kaplan *et al.* (1989) illustrate their LFG-MT proposal with the well-known headswitching case *venir de X* ⟷ *has just X-ed*, as in (1):

(1) The baby just fell ⟷ Le bébé vient de tomber.

They propose to deal with such problems in two ways. The first of these is as in (2):

(2) *just*: $(\uparrow \text{PRED}) = \text{`just}\langle(\uparrow \text{ARG})\rangle\text{'}$, $(\tau\uparrow \text{PRED}) = \text{venir}$, $(\tau\uparrow \text{XCOMP}) = \tau(\uparrow \text{ARG})$

That is, the XCOMP function of *venir* (in (1), *de tomber*) corresponds to the ARG function of *just* (in (1), *the baby fell*), as shown by the respective source and target f-structures in (3) and (4):

(3)
$$\left[\begin{array}{ll} \text{PRED} & \text{`just}\langle[\text{fall}]\rangle\text{'} \\ \text{ARG} & \left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{`baby'} \\ \text{SPEC} & \text{the}\end{array}\right] \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{`fall}\langle[\text{baby}]\rangle\text{'}\end{array}\right]\end{array}\right]$$

(4)
$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{`bébé'} \\ \text{SPEC} & \text{le}\end{array}\right] \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`venir}\langle[\text{bébé}],[\text{tomber}]\rangle\text{'} \\ \text{XCOMP} & \left[\begin{array}{ll} \text{DE} & + \\ \text{SUBJ} & [\;\;] \\ \text{PRED} & \text{`tomber}\langle[\text{bébé}]\rangle\text{'}\end{array}\right]\end{array}\right]$$

The second approach is where *just* is not treated as a head subcategorizing for an ARG, but as a 'normal' adverbial sentential modifier. Instead, headswitching occurs between source and target f-structures, as in (5):

(5)

$$\begin{array}{llll} \mathrm{S} \longrightarrow & \mathrm{NP} & \mathrm{ADVP} & \mathrm{VP} \\ & (\uparrow\mathrm{SUBJ}){=}\downarrow & (\uparrow\mathrm{SADJ}){=}\downarrow & \uparrow{=}\downarrow \\ & & (\tau\uparrow\mathrm{SADJ\ XCOMP}) = \tau\uparrow & \end{array}$$

$$\begin{array}{llll} \textit{just}\text{: ADV,} & (\uparrow \mathrm{PRED}) = \text{just} & \textit{fall}\text{: V,} & (\uparrow \mathrm{PRED}) = \text{fall} \\ & (\tau\uparrow \mathrm{PRED}) = \text{venir} & & (\tau\uparrow\mathrm{SUBJ}) = \tau(\uparrow\mathrm{SUBJ}) \\ & & & (\tau\uparrow \mathrm{PRED}) = \text{tomber} \end{array}$$

Here the τ annotation to ADVP states that the τ of the mother f-structure is the XCOMP of the τ of the SADJ slot. This set of equations (along with others of a more trivial nature) produces the f-structure (6):

(6)

$$\left[\begin{array}{ll} \mathrm{SUBJ} & \left[\begin{array}{ll} \mathrm{PRED} & \text{`baby'} \\ \mathrm{SPEC} & \text{the} \end{array}\right] \\ \mathrm{TENSE} & \mathrm{PAST} \\ \mathrm{PRED} & \text{`fall}\langle[\text{baby}]\rangle\text{'} \\ \mathrm{SADJ} & \left\{\left[\begin{array}{ll} \mathrm{PRED} & \text{`just'} \end{array}\right]\right\} \end{array}\right]$$

1.1 Embedded Cases of Headswitching

However, Sadler *et al.* (1989, 1990) show that neither approach is able to deal elegantly and straightforwardly with more complex cases of headswitching, as in (7):

(7) I think that the baby just fell ⟷Je pense que le bébé vient de tomber.

In (7), the headswitching phenomenon takes place in the sentential COMP, rather than in the main clause, as in (1). Here the structure in (3) must be a COMP to a PRED in a higher f-structure. Hence, the normal f-description on embedded S nodes ($\uparrow$ COMP = $\downarrow$) must be optional, and instead the structure in (3) must be unified to the root f-structure as the value of its COMP node. This can be handled by the disjunction in (8):

(8)

$$\begin{array}{lllll} \mathrm{VP} & \longrightarrow & \mathrm{V} & \text{that} & \mathrm{S} \\ & & & & \{(\uparrow\mathrm{COMP}){=}\downarrow,\ (\uparrow\mathrm{COMP\ ARG}){=}\downarrow\} \end{array}$$

We require this disjunction on embedded S nodes to include ($\uparrow$COMP ARG)=$\downarrow$ just in case they contain such a headswitching construction, as f-structure (9) shows:

(9)

$$\left[\begin{array}{ll} \mathrm{SUBJ} & \left[\begin{array}{ll} \mathrm{PRED} & \text{`I'} \end{array}\right] \\ \mathrm{PRED} & \text{`think}\langle[\text{i}],[\text{just}]\rangle\text{'} \\ \mathrm{COMP} & \left[\begin{array}{ll} \mathrm{PRED} & \text{`just}\langle[\text{fall}]\rangle\text{'} \\ \mathrm{ARG} & \left[\begin{array}{ll} \mathrm{SUBJ} & \left[\begin{array}{ll} \mathrm{PRED} & \text{`baby'} \\ \mathrm{SPEC} & \text{the} \end{array}\right] \\ \mathrm{TENSE} & \mathrm{PAST} \\ \mathrm{PRED} & \text{`fall}\langle[\text{baby}]\rangle\text{'} \end{array}\right] \end{array}\right] \end{array}\right]$$

Otherwise, structure (3) (rooted in *just*) is not connected to the higher COMP slot. Nevertheless, the solution proposed in (8) seems a little *ad hoc*, requiring a disjunction just in case the sentential COMP includes a headswitching case. We shall see in the next section that if such headswitching adverbs co-occur, then further disjuncts are required, unless these can be abbreviated by a functional uncertainty equation.

If we choose the second approach (5), where *just* is a sentential modifier, given that the headswitching is a τ operation, we require the lexical entry for *think* in (10):

(10) *think*: V, ($\tau \uparrow$ PRED) = penser, τ ($\uparrow$ SUBJ) = ($\tau \uparrow$ SUBJ), τ ($\uparrow$ COMP) = ($\tau \uparrow$ COMP)

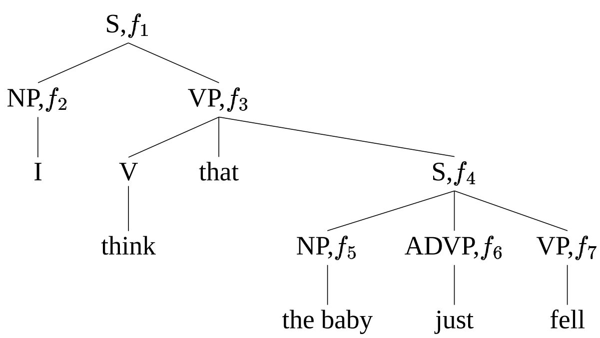
This specifies that τ of the mother f-structure's COMP slot is the COMP of the τ of the mother's f-structure. That is, both this argument, the COMP, and the SUBJ of *think* are to be translated straightforwardly. This is indeed the case in (11):

(11) I think that the baby fell $\longleftrightarrow$ Je pense que le bébé est tombé.

However, when the COMP includes a headswitching case, as in (7), we end up with a doubly rooted target f-structure because of a clash between the regular τ equation in the lexical entry for *think*, (10), and the structural τ equation on the ADVP in the (5), which requires the τ of the same piece of f-structure to be the XCOMP of the τ of the SADJ slot. One piece of f-structure is required to fill two inconsistent slots. We will now illustrate this in detail.

The c- and f-structures for the source sentence in (7), *I think that the baby just fell*, are shown in (12):

(12)



$$f_1, f_3\left[\begin{array}{ll}\text{SUBJ} & f_2\left[\begin{array}{ll}\text{PRED} & \text{‘I’}\end{array}\right] \\ \text{PRED} & \text{‘think}\langle\text{[i],[fall]}\rangle\text{’} \\ \text{COMP} & f_4, f_7\left[\begin{array}{ll}\text{SUBJ} & f_5\left[\begin{array}{ll}\text{PRED} & \text{‘baby’} \\ \text{SPEC} & \text{the}\end{array}\right] \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{‘fall}\langle\text{[baby]}\rangle\text{’} \\ \text{SADJ} & \left\{f_6\left[\begin{array}{ll}\text{PRED} & \text{‘just’}\end{array}\right]\right\}\end{array}\right]\end{array}\right]$$

The set of τ equations obtained are those in (13):

(13) From the lexical entry for *think*, (10):
τf_3 PRED = *penser*
τf_3 SUBJ = $((\tau f_3)$ SUBJ)
τf_3 COMP = $((\tau f_3)$ COMP)

From the rules and entries in (5):
τf_6 PRED = *venir*
τf_7 PRED = *tomber*
$\tau f_4 = (\tau(f_4$ SADJ) XCOMP)

From straightforward equations for the NPs:
τf_2 PRED = *je*
τf_5 PRED = *bébé*
τf_5 SPEC = *le*

Now, from the c-structure in (12), we see that f_3's COMP is f_4, so the equation τf_3 COMP = $((\tau f_3)$ COMP) can be altered quite simply to τf_4 = $((\tau f_3)$ COMP). By the same token, f_4's SADJ is f_6. We now have two equations which cannot both be solved with the result that a target f-structure is formed. The clash is shown in (14):

(14)
τ ($\uparrow$ COMP) = ($\tau \uparrow$ COMP) $\quad \tau f_4$ = $((\tau f_3)$ COMP)
$\tau \uparrow$ = τ ($\uparrow$ SADJ XCOMP) $\quad \tau f_4 = ((\tau f_6)$ XCOMP)

This results in the construction of the two partial target f-structures in (15):

(15)
$$\tau f_6\left[\begin{array}{ll}\text{SUBJ} & \tau f_5\left[\ \right] \\ \text{PRED} & \text{‘venir}\langle\text{[bébé],[tomber]}\rangle\text{’} \\ \text{XCOMP} & \tau f_4\left[\begin{array}{ll}\text{SUBJ} & \tau f_5\left[\begin{array}{ll}\text{PRED} & \text{‘bébé’} \\ \text{SPEC} & \text{le}\end{array}\right] \\ \text{PRED} & \text{‘tomber}\langle\text{[bébé]}\rangle\text{’}\end{array}\right]\end{array}\right] \qquad \tau f_3\left[\begin{array}{ll}\text{SUBJ} & \tau f_2\left[\begin{array}{ll}\text{PRED} & \text{‘Je’}\end{array}\right] \\ \text{PRED} & \text{‘penser}\langle\text{[je],[...]}\rangle\text{’} \\ \text{COMP} & \tau f_4\left[\ \right]\end{array}\right]$$

That is, τf_4 is required to be both the XCOMP of *venir* and the COMP of *penser* simultaneously, a conflict that needs to be resolved if a proper target f-structure is to be produced. As was done in (8), we might disjoin the problematic τ equation in (10) with an equation for the ‘special’ case in (16):

(16) $(\tau \uparrow \text{COMP}) = \tau\ (\uparrow \text{COMP SADJ})$

That is, τ of the mother at the COMP slot is τ of the mother's COMP SADJ slot. This is clearly undesirable, since it must be specified for every embedding verb.

1.2 Scoping of Multiple Adverbs

In a similar fashion, Sadler *et al.* (1989, 1990) show that the two approaches of Kaplan *et al.* cannot deal straightforwardly with other complex cases of headswitching involving scoping of multiple adverbs, as in (17):

(17) a. Jan zwemt toevallig graag $\longrightarrow$

b. John happens to like to swim

c. *John likes to happen to swim

That is, despite the fact that (17c) is a grammatical string of English, it is not the translation of (17a)—(17b) is. (17c) is the translation of *Jan zwemt graag toevallig*. Given that the word order differs, we have a different scoping of the adverbs, resulting in a different translation. Let us assume the S rule in (5), and the lexical entries in (18):

(18) *graag*: ADV, $(\uparrow \text{PRED}) = \text{graag}$, $(\tau\uparrow \text{PRED}) = \text{like}$

zwemmen: V, $(\uparrow \text{PRED}) = \text{zwemmen}$, $(\tau\uparrow \text{SUBJ}) = \tau(\uparrow \text{SUBJ})$, $(\tau\uparrow \text{PRED}) = \text{swim}$

Let us also assume the somewhat simpler sentence (19):

(19) Jan zwemt graag ($\longrightarrow$John likes to swim)

Given the S rule in (5), and the lexical entries in (18), the f-structure for (19) is (20):

$$(20)\quad \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`Jan'} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`zwemmen}\langle[\text{Jan}]\rangle\text{'} \\ \text{SADJ} & \left\{\begin{bmatrix} \text{PRED} & \text{`graag'} \end{bmatrix}\right\} \end{bmatrix}$$

This object is very similar to the f-structure in (6). The other possibility is that the lexical entry for *graag* is (21):

(21) *graag*: $(\uparrow \text{PRED}) = \text{`graag}\langle(\uparrow \text{ARG})\rangle\text{'}$, $(\tau\uparrow \text{PRED}) = \text{like}$, $(\tau\uparrow \text{XCOMP}) = \tau(\uparrow \text{ARG})$

In this case, the f-structure in (22) would be built for the Dutch sentence in (19):

$$(22)\quad \begin{bmatrix} \text{PRED} & \text{`graag}\langle[\text{zwemmen}]\rangle\text{'} \\ \text{ARG} & \begin{bmatrix} \text{SUBJ} & \text{`Jan'} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`zwemmen}\langle[\text{Jan}]\rangle\text{'} \end{bmatrix} \end{bmatrix}$$

Being an adverb of the same type, *toevallig* can occur in similar contexts as *graag*, as in (23):

(23) Jan zwemt toevallig ($\longrightarrow$John happens to swim)

As seen in (17), such adverbs can co-occur. To maintain this approach and produce the translation (17b), we require the f-structure in (24):

(24)
$$\left[\begin{array}{ll} \text{PRED} & \text{‘toevallig}\langle[\text{graag}]\rangle\text{’} \\ \text{ARG} & \left[\begin{array}{ll} \text{PRED} & \text{‘graag}\langle[\text{zwemmen}]\rangle\text{’} \\ \text{ARG} & \left[\begin{array}{ll} \text{SUBJ} & \text{‘Jan’} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{‘zwemmen}\langle[\text{Jan}]\rangle\text{’} \end{array}\right] \end{array}\right] \end{array}\right]$$

To associate the embedded S node with an f-structure which is the value of COMP ARG ARG of the mother S’s associated f-structure, we would need to add ($\uparrow$ COMP ARG ARG) = $\downarrow$ to the disjunction on S in (8). We may be able to simplify these equations using functional uncertainty as in ($\uparrow$ COMP ARG*) = $\downarrow$. Nevertheless, it remains very unnatural to annotate embedded S symbols to allow for the possibility that they *may* contain (any number of) subcategorizing ADVs of this type. It should be clear that cases like (25) are problematic in the same way as (7)-(16):

(25) Ik denk dat Jan toevallig graag zwemt.

The alternative f-structure corresponding to (17a) is (26):

(26)
$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘Jan’} \\ \text{NUM} & \text{SG} \end{array}\right] \\ \text{PRED} & \text{‘zwemmen}\langle(\uparrow\text{SUBJ})\rangle\text{’} \\ \text{TENSE} & \text{PRES} \\ \text{ADJUNCT} & \left\{\left[\begin{array}{ll} \text{PRED} & \text{‘toevallig’} \end{array}\right], \left[\begin{array}{ll} \text{PRED} & \text{‘graag’} \end{array}\right]\right\} \end{array}\right]$$

Given the original formulation of LFG, there is no way of producing the required embedding of *graag* (‘likingly’) under *toevallig* (‘by chance’), and not vice versa, without resorting to tuning: under both approaches outlined, changing our c-structure assumptions to deal with a difficult translation case necessitates the abandonment of modularity.

In general, this approach requires the tuning of f-structures, which is arguably as problematic as producing a sufficiently abstract representation for simple transfer in other systems. For example, since German and Dutch both have such adverbs, as shown in (27), it is possible to treat them as ‘normal’ adverbs and still produce adequate translations:

(27) a. NL: toevallig $\longrightarrow$DE: zufällig

b. NL: graag $\longrightarrow$DE: gerne

Hence the danger exists of producing different source language f-structures according to the target language requirements: the cases where the adverbs are top-level PREDs (such as (21), for instance) is appropriate for translation from Dutch to English, just because this involves the switching of heads. The alternative lexical form for the same adverb, (18), is required for translation from Dutch to German, where there is no headswitching for this example.

$$\text{a.} \left[\begin{array}{ll} \text{REL} & \text{‘toevallig’} \\ \text{ARG1} & \left[\begin{array}{ll} \text{REL} & \text{‘graag’} \\ \text{ARG1} & \left[\begin{array}{ll} \text{REL} & \text{‘zwemmen’} \\ \text{ARG1} & \text{‘Jan’} \end{array}\right] \end{array}\right] \end{array}\right] \quad \text{b.} \left[\begin{array}{ll} \text{REL} & \text{‘graag’} \\ \text{ARG1} & \left[\begin{array}{ll} \text{REL} & \text{‘toevallig’} \\ \text{ARG1} & \left[\begin{array}{ll} \text{REL} & \text{‘zwemmen’} \\ \text{ARG1} & \text{‘Jan’} \end{array}\right] \end{array}\right] \end{array}\right]$$

Figure 1: S-structures derived using Restriction

1.3 Other Solutions

1.3.1 Restriction

Given cases with adverbial modifiers such as (17), Kaplan & Wedekind (1993) attempt to solve them by the introduction of the notion of *restriction*, which seeks to overcome problems in mapping between flat syntactic f-structures to hierarchical semantic ones. The intuition is that in such cases semantic units correspond to subsets of functional information, and restricting the f-structure (in other words, removing *graag* and *toevallig* in turn from the adjunct set in (26)) enables (26) to be associated with the alternative s-structures in Figure 1. Kaplan & Wedekind (*op cit.*, p.199) define the restriction of an f-structure by an element of an element's set-value, as in (28):

(28) If *f* is an f-structure and *a* is an attribute:

$$f \setminus \langle a\ g \rangle = \begin{cases} f \setminus a & \text{if } (f\ a) - \{g\} = \emptyset \\ f \setminus a \bigcup \{\langle a,\ (f\ a) - \{g\} \rangle\} & \text{otherwise} \end{cases}$$

That is, the restriction of an f-structure *f* by a particular member of an attribute *a*'s set-value is the f-structure which results from deleting that member of the set, and also the attribute itself if an empty set results. We can illustrate how restriction works in (29), taking (26) as input:

(29)

$$f = \left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘Jan’} \\ \text{NUM} & \text{SG} \end{array}\right] \\ \text{PRED} & \text{‘zwemmen}\langle(\uparrow\text{SUBJ})\rangle\text{’} \\ \text{TENSE} & \text{PRES} \\ \text{ADJUNCT} & \left\{ \left[\begin{array}{ll} \text{PRED} & \text{‘toevallig’} \end{array}\right], \left[\begin{array}{ll} \text{PRED} & \text{‘graag’} \end{array}\right] \right\} \end{array}\right] \qquad g = \left[\begin{array}{ll} \text{PRED} & \text{‘toevallig’} \end{array}\right]$$

$$f \setminus \langle g \rangle = \left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘Jan’} \\ \text{NUM} & \text{SG} \end{array}\right] \\ \text{PRED} & \text{‘zwemmen}\langle(\uparrow\text{SUBJ})\rangle\text{’} \\ \text{TENSE} & \text{PRES} \\ \text{ADJUNCT} & \left\{ \left[\begin{array}{ll} \text{PRED} & \text{‘graag’} \end{array}\right] \right\} \end{array}\right]$$

We are now in a position to describe the semantic correspondences for sentences containing adverbs, such as (17a), using the restriction operator. Let *f* be the f-structure (26), *g* the f-structure corresponding to *graag* and *t* the f-structure corresponding to *toevallig*. We give in (30) the constraints necessary to map (26) into the s-structure (b) in Figure 1, where *graag* has wide scope:

(30) a. $(\sigma f\ REL) = (\sigma g\ REL)$

b. $(\sigma f\ ARG1) = \sigma[f \setminus \langle ADJUNCT\ g\rangle]$

c. $(\sigma[f \setminus \langle ADJUNCT\ g\rangle]\ REL) = (\sigma t\ REL)$

d. $(\sigma[f \setminus \langle ADJUNCT\ g\rangle]\ ARG1)$
$= \sigma[f \setminus \langle ADJUNCT\ g\ t\rangle]$
$= \sigma[f \setminus ADJUNCT]$

(30a,b) describe the outermost REL and ARG1 configuration in Figure 1b, and (30c,d) describe the next level of embedding at s-structure. These constraints allow f-structure subsumption relations to be mapped into the desired hierarchical s-structures. However, we note that the number of such constraints will grow in proportion with the size of the set of adjuncts. Kaplan & Wedekind (*op cit.*, p.200) give a rule which generates codescription constraints, as in (31):

(31) For *f* an f-structure, $g \in (f\ ADJUNCT)$, and g a sentence adverb,
$\sigma f = \sigma g$, and
$(\sigma f\ ARG1) = \sigma[f \setminus \langle ADJUNCT\ g\rangle]$

(31) allows each element to be selected non-deterministically from an adjunct set to contribute to the relation for the s-structure of the enclosing f-structure. Furthermore, the s-structure corresponding to the f-structure minus the selected member becomes the ARG1 of that relation. Given that the *restriction* operation applies non-deterministically to all members of a set, we get, as here, a correct and some incorrect structures: with respect to sentence (17), s-structure Figure 1a is correct whilst Figure 1b is incorrect. Although this may be seen as an improvement on the codescription approach, where the production of the correct f-structure could not be guaranteed, it nevertheless leaves something to be desired in that human intervention is necessary to manually select the optimal structure from the (possibly large) set of candidate solutions.

1.3.2 Linear Logic

Some cases of headswitching, notably adjuncts and embedded headswitching phenomena, have been tackled using linear logic to formalize transfer rules (Van Genabith *et al.*, 1998). Nevertheless, when it comes to headswitching cases the linear logic approach encounters some problems. Van Genabith *et al.* use the *like* $\longleftrightarrow$ *gerne* example in the sentence pair *Hans schwimmt gerne* $\longleftrightarrow$ *Hans likes swimming*. Given the appropriate f-structures, the source set of meaning constructors in (32) is derived:

(32)

$$\left\{\begin{array}{c} (f_2)_\sigma \rightsquigarrow hans \\ \forall X\,[(f_2)_\sigma \rightsquigarrow X \multimap (f_1)_\sigma \rightsquigarrow schwimmen(X)] \\ \forall P\,[(f_1)_\sigma \rightsquigarrow P \multimap (f_1)_\sigma \rightsquigarrow gerne(P)] \end{array}\right\} \vdash (f_1)_\sigma \rightsquigarrow gerne(schwimmen(hans))$$

The meaning constructor for *like* is (33):

(33)

$$like: \quad \forall X, P\,[((\uparrow SUBJ)_\sigma \rightsquigarrow X \otimes \forall Y\,((\uparrow XCOMP\ SUBJ)_\sigma \rightsquigarrow Y \multimap (\uparrow XCOMP)_\sigma \rightsquigarrow P(Y))) \multimap \uparrow_\sigma \rightsquigarrow like(X, P(X))]$$

This final term, *like(X,P(X))*, indicates that X is the subject of *like*, and that somewhere in its second argument, X re-enters as the XCOMP SUBJ. The complete set of meaning constructors for *Hans likes swimming* is given in (34):

(34)

$$\left\{ \begin{array}{c} (f_2)_\sigma \rightsquigarrow hans \\ \forall X, P\ [((f_2)_\sigma \rightsquigarrow X \otimes \forall Y\ ((f_2)_\sigma \rightsquigarrow Y \multimap (f_1)_\sigma \rightsquigarrow P(Y))) \multimap (f_3)_\sigma \rightsquigarrow like(X, P(X))] \\ \forall X\ [(f_2)_\sigma \rightsquigarrow X \multimap (f_1)_\sigma \rightsquigarrow swim(X)] \end{array} \right\}$$

$$\vdash (f_3)_\sigma \rightsquigarrow like(hans, swim(hans))$$

The transfer constructor *gerne* $\multimap_t$ *like* consumes the entire meaning constructor for *gerne* as there are no left-common prefixes in the meaning constructors, and the meaning constructor for *like* in (35) is derived:

(35)

$$\begin{array}{c} \forall F\ [\forall P(F_\sigma \rightsquigarrow P \multimap F_\sigma \rightsquigarrow gerne(P)) \\ \multimap_t \\ \forall X, P\ ((F\ SUBJ)_\sigma \rightsquigarrow X \otimes \forall Y\ ((F\ XCOMP\ SUBJ)_\sigma \rightsquigarrow Y \multimap F_\sigma \rightsquigarrow P(Y)) \multimap F_\sigma \rightsquigarrow like(X, P(X)))] \end{array}$$

Now the first problem with this approach can be seen. On the right-hand side of the $\multimap_t$, we observe that the meaning constructor has rewritten a node F rather than a node $F\ XCOMP$ to match $P(Y)$. Consequently, with the instantiated source meaning constructors and the transfer constructors, the equations in (36) are produced:

(36)

$$\textit{Source} \cup \left\{ \begin{array}{c} schwimmen \multimap_t swim \\ gerne \multimap_t like \end{array} \right\} \vdash_t$$

$$\left\{ \begin{array}{ll} 1. & (f_2)_\sigma \rightsquigarrow hans \\ 2. & \forall X\ [(f_2)_\sigma \rightsquigarrow X \multimap (f_1)_\sigma \rightsquigarrow swim(X)] \\ 3. & \forall X, P\ [((f_2)_\sigma \rightsquigarrow X \otimes \forall Y\ ((f_2)_\sigma \rightsquigarrow Y \multimap (f_1)_\sigma \rightsquigarrow P(Y))) \multimap (f_1)_\sigma \rightsquigarrow like(X, P(X))] \end{array} \right\} \vdash$$

$$(f_1)_\sigma \rightsquigarrow like(hans, swim(hans))$$

Comparison of the third target meaning constructor with (33) shows that the transfer operation has rewritten a single node $(f_1)_\sigma$ rather than accessing a complement node $(f_1\ XCOMP)_\sigma$ to match against $P(Y)$. Van Genabith *et al.* then give an example of embedded headswitching involving the *like* $\longleftrightarrow$ *gerne* case, and unsurprisingly the same fault is uncovered again.

Transfer should deliver exactly the set of meaning constructors as would be obtained by independent analysis of the target string. If it does not, target language generation from underspecified sets of target meaning constructors will not produce the required output, and the overall translation obtained will be wrong. Van Genabith *et al.* propose to rectify the problem by 'pushing down' the predicate-argument nucleus of verbs one (or more, as appropriate) levels, via functional uncertainty over XCOMP. Hence the transfer constructor *schwimmen* $\multimap_t$ *swim* gets amended to (37):

(37) *schwimmen* $\multimap_t$ *swim*:
$\forall F, X\ (F_\sigma \rightsquigarrow schwimmen(X) \multimap_t (F\ XCOMP*)_\sigma \rightsquigarrow swim(X))$

Furthermore, the transfer constructor *gerne* $\multimap_t$ *like* is also redefined given the knowledge that the predicate-argument structure and the corresponding semantic projector associated with the translation of the proposition in the scope of the source adjunct is 'pushed down' via functional uncertainty in (37) on the target side. Given these amendments, the set of target meaning constructors produced is identical to those in (33), as required.

However, there are a number of problems with this solution. It is clear that *any* verb plus associated arguments can occur as the complement of *like*. Consequently, the amended transfer constructor in (37) will have repercussions for *every* verbal translation relation. That is, all transfer constructors relating verbs will need to include such an equation, *just in case* it ever occurs as an XCOMP in an infinitival phrase of another verb. Furthermore, the addition of how to translate verbs such as *schwimmen* $\longleftrightarrow$ *swim* in the event of *swim* appearing as the complement of *like* has nothing to do with the translation relation in question at all: the translation of *schwimmen* as *swim* is a case of simple transfer. Information about *swim* as an XCOMP has nothing to do with *swim*, or *schwimmen*, but is an artefact of the *like* $\longleftrightarrow$ *gerne* case. Consequently it should be removed from the context of *schwimmen* $\longleftrightarrow$ *swim* and relocated in its proper place. If the approach is subsequently unable to deal with headswitching examples, then so be it, but at least the basic translation relations are kept intact and untainted by *ad hoc* information which does not belong there.

2 LFG-DOT: a new Model of Translation

While LFG's τ equations are in the main able to link exactly those source-target elements which are translations of each other, leading to elegant translation models such as that of Kaplan *et al.* (1989), we have described a number of cases, in particular embedded headswitching examples, where this machinery, and others, are unable to cope.

Way (2001) proposes the use of LFG-DOP (Bod & Kaplan, 1998) as the basis for an innovative MT system. LFG-DOP combines the syntactic representations of LFG with the statistical language modelling of DOP (Bod, 1998) to create a new, more powerful hybrid model of language processing. LFG-DOP representations consist of LFG $\langle c, f\rangle$ pairs with a mapping ϕ between them. The *Root* and *Frontier* decomposition operations of DOP operate on CF-PSG trees only, so these operations are adapted in LFG-DOP using the notion of ϕ-accessibility to stipulate exactly which c-structure nodes are linked to which f-structure fragments, thereby maintaining the fundamentals of c- and f-structure correspondence. A third, new decomposition operation, *Discard*, is introduced in LFG-DOP by which generalized fragments produced by *Root* and *Frontier* are created by freely deleting any combination of attribute-value pairs from an f-structure except those that are ϕ-linked to some remaining c-structure node, or that are governed by the local predicate. Fragments are combined together in two stages: c-structures are combined by leftmost substitution, as in DOP, subject to the matching of their nodes. F-structures corresponding to these nodes are then recursively unified, and the resulting f-structures are subjected to the grammaticality checks of LFG. LFG-DOP probability models are based on relative frequency: Bod & Kaplan (1998) give different possible definitions of competition sets from which sample derivations are chosen.

Way (2001) presents four models of translation which use LFG-DOP as their language models, but which differ with respect to how translations are obtained. We shall present these briefly here, and comment on their ability to handle more complex headswitching examples such as (7).

2.1 LFG-DOT1: Translation via τ

Given a source language LFG-DOP treebank, the model builds a target f-structure f' from a source c-structure c and f-structure f, the mapping between them LFG-DOP-ϕ, and the LFG translation equations τ. From this target f-structure f', a target string is generated via a target language LFG-DOP model, as in (38):

(38)
$$\begin{tikzcd} c \arrow[r, "\text{LFG-DOP-}\phi"] & f \arrow[d, "\tau"] \\ c' & f' \arrow[l, "\text{LFG-DOP-}\phi'"] \end{tikzcd}$$

While LFG-DOT1 contains source and target LFG-DOP language models, thereby adding robustness to LFG via the *Discard* operation, it maintains the use of τ equations to drive the translation component. Unsurprisingly, therefore, the problems of LFG-MT when confronted with certain headswitching data are maintained in LFG-DOT1, so we do not consider this model further here.

2.2 LFG-DOT2: Translation via τ and γ

LFG-DOT1 relates languages just at the level of f-structure (via τ), and so fails in the same way as LFG-MT with the headswitching data presented in section 1. Way (2001) demonstrates that the DOT2 system of translation (Poutsma, 2000) is able to handle embedded headswitching cases correctly.[1] Accordingly, the γ relation, which links source and target subtree fragments in DOT, is introduced into the LFG-DOT2 translation model. This requires integrated bilingual LFG-DOP corpora, where each node *n* in a source c-structure tree *c* is related both to its corresponding f-structure fragment *f* (via LFG-DOP-ϕ) and its corresponding c-structure node n' in a target c-structure tree c' (via γ). In addition, each f-structure fragment s in a source f-structure *f* is related to its corresponding language fragment s' in a target f-structure f', via τ, as shown in (39):

(39)
$$\begin{tikzcd} c \arrow[r, "\text{LFG-DOP-}\phi"] \arrow[d, "\gamma"'] & f \arrow[d, "\tau"] \\ c' & f' \arrow[l, "\text{LFG-DOP-}\phi'"] \end{tikzcd}$$

That is, the translation component consists of an integration of the γ probabilities with the τ mapping. Way (2001) describes how these two translation sources might best be combined, but this need not concern us here. Ultimately, despite the fact that LFG-DOT2 is a richer model than LFG-DOT1, it too is rejected as it continues to maintain τ equations which are incapable of ensuring that the correct translation is obtained in all cases of headswitching.

The presence of the f-structure information is required in order to allow *Discard* to run and thereby make LFG-DOT more robust than LFG-MT. However, *Discard* can operate whether the f-structures are linked via τ or not, so it would appear that the τ operation itself is not needed.[2] The next two LFG-DOT models, therefore, omit the τ operation and use the γ relation to produce translations.

[1] There are two remaining caveats: (i) in dealing with ill-formed translation pairs such as ⟨*John swim*, *Jan zwemmen*⟩, DOT2 considers such pairs as 'grammatical with respect to the corpus'. In LFG-DOT, such pairs can be handled only by removing certain atrribute-value pairs in the respective f-structures via *Discard* in order to permit them to be unified. Such derivations are considered 'ungrammatical', in line with our intuitive notion of well-formedness; (ii) DOT2 cannot handle such cases fully compositionally. With respect to this latter point, we shall see that LFG-DOT3 suffers in the same way, but LFG-DOT4 avoids the problem of limited compositionality.

[2] We leave for future work the question as to whether this approach is fruitful for languages which differ significantly at the level of surface structure, e.g. English and Warlpiri. In such cases, perhaps an LFG-DOT1 or LFG-DOT2 model may be better to relate translational equivalents at the level of f-structure rather than c-structure.

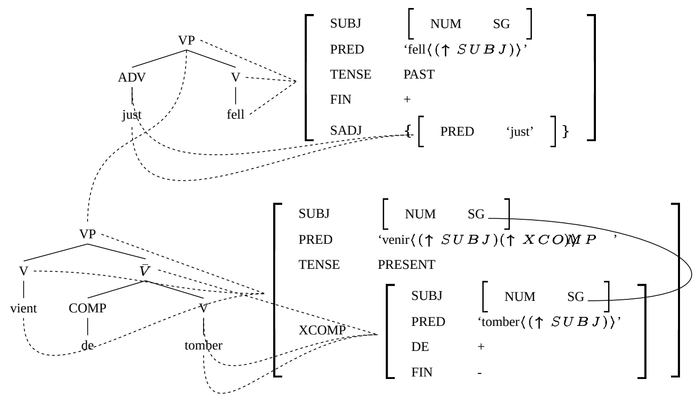


Figure 2: The *just* $\longleftrightarrow$ *venir de* case in LFG-DOT3

2.3 LFG-DOT Model 3: Translation via γ with Monolingual Filtering

The LFG-DOT3 translation model contains the DOT2 γ links between source and target c-structures, but with additional syntactic functional constraints which prevent the formation of certain ungrammatical structures (cf. note 1), thereby enabling correct translations to be output with appropriate probabilities. The f-structure information can be seen, therefore, as useful for monolingual disambiguation in both source and target sides. Ill-formed or unknown input is still processable by running *Discard* over the set of linked source and target $\langle c,$ LFG-DOP-ϕ, $f\rangle$ fragments. The LFG-DOT3 architecture is shown in (40): [3]

(40)
$$\begin{tikzcd} c \arrow[r, "\text{LFG-DOP-}\phi"] \arrow[d, "\gamma"'] & f \\ c' \arrow[r, "\text{LFG-DOP-}\phi'"'] & f' \end{tikzcd}$$

In order to exemplify how headswitching is dealt with in LFG-DOT3, let us consider the *just* $\longleftrightarrow$*venir de* case (1). In terms of LFG-DOT3, the translation relation is shown in Figure 2. The γ link between semantically equivalent elements in the source and target c-structures can be seen on the VP nodes. *fell* is not considered to be semantically equivalent to *tomber* owing to their different FIN(ite) values, added to the fact that *fell* has a TENSE value whilst *tomber* does not. Hence this translation fragment can only be reused by substituting this pair with associated singular NP subjects at the appropriate nodes in an S-linked fragment. In this respect, as with DOT2 (and LFG-DOT2), this LFG-DOT3 model continues to suffer from limited compositionality. We address this concern further in section 2.4 dealing with the LFG-DOT4 model, which has an extra level of processing called 'Extended Transfer'. In all other respects, LFG-DOT3 and LFG-DOT4 are the same models, so while we end up rejecting LFG-DOT3 in favour of LFG-DOT4, much of the ensuing discussion is relevant to our final choice of model, LFG-DOT4.

2.3.1 LFG-DOT3 and Embedded Headswitching

Having shown how LFG-DOT3 copes with a headswitching problem, we shall now investigate whether LFG-DOT3 can handle translation examples which LFG-MT finds problematic. We showed that in example (7), the default τ equations in the entry for *think* (10) clash with those on the structural rule for ADVP (5).

[3]Way (2001) proves that LFG-DOT2 and LFG-DOT3 are different models by showing that while τ links can be inferred from ϕ and γ links in the general case, this is not possible when more complex translation data is examined.

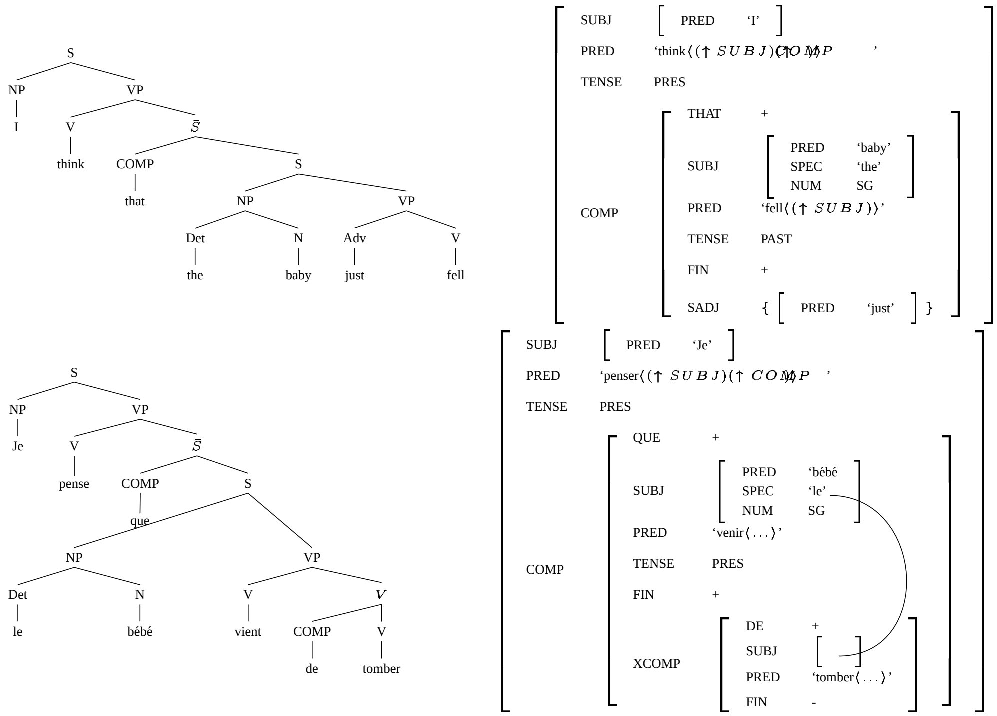
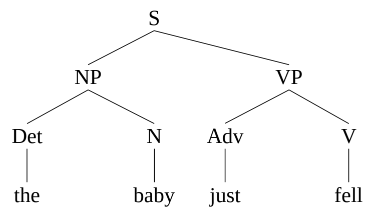


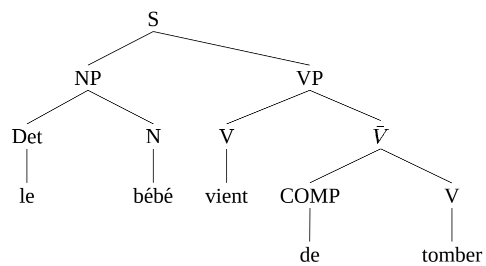
Figure 3: LFG-DOT3 representation for the embedded headswitching case *I think that the baby just fell* ⟷*Je pense que le bébé vient de tomber*

The LFG-DOT3 representations for the full trees in (7) are shown in Figure 3.[4] Source and target trees are linked at the topmost S, NP, V and VP levels, as well as at $\bar{S}$, COMP and embedded S levels. Given that each source fragment will be linked to its target counterpart with the same label, each ⟨*source*, *target*⟩ linked pair can be deleted to make the new linked fragment pair in (41):

[4]In this and some subsequent examples, ϕ and γ links are omitted for reasons of clarity.

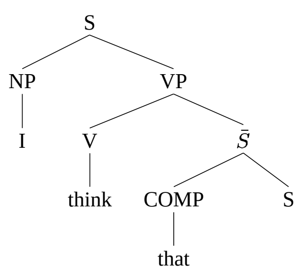
(41) 

$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘baby’} \\ \text{SPEC} & \text{‘the’} \\ \text{NUM} & \text{SG} \end{array}\right] \\ \text{PRED} & \text{‘fell}\langle(\uparrow SUBJ)\rangle\text{’} \\ \text{TENSE} & \text{PAST} \\ \text{FIN} & + \\ \text{SADJ} & \{\left[\begin{array}{ll} \text{PRED} & \text{‘just’} \end{array}\right]\} \end{array}\right]$$

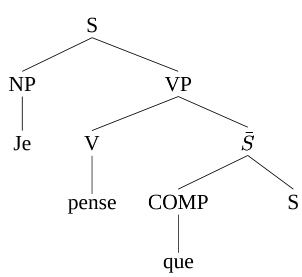


$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘bébé} \\ \text{SPEC} & \text{‘le’} \\ \text{NUM} & \text{SG} \end{array}\right] \\ \text{PRED} & \text{‘venir}\langle(\uparrow SUBJ)(\uparrow XCOMP)\rangle\text{’} \\ \text{TENSE} & \text{PRES} \\ \text{FIN} & + \\ \text{XCOMP} & \left[\begin{array}{ll} \text{DE} & + \\ \text{SUBJ} & [\ ] \\ \text{PRED} & \text{‘tomber}\langle(\uparrow SUBJ)\rangle\text{’} \\ \text{FIN} & - \end{array}\right] \end{array}\right]$$

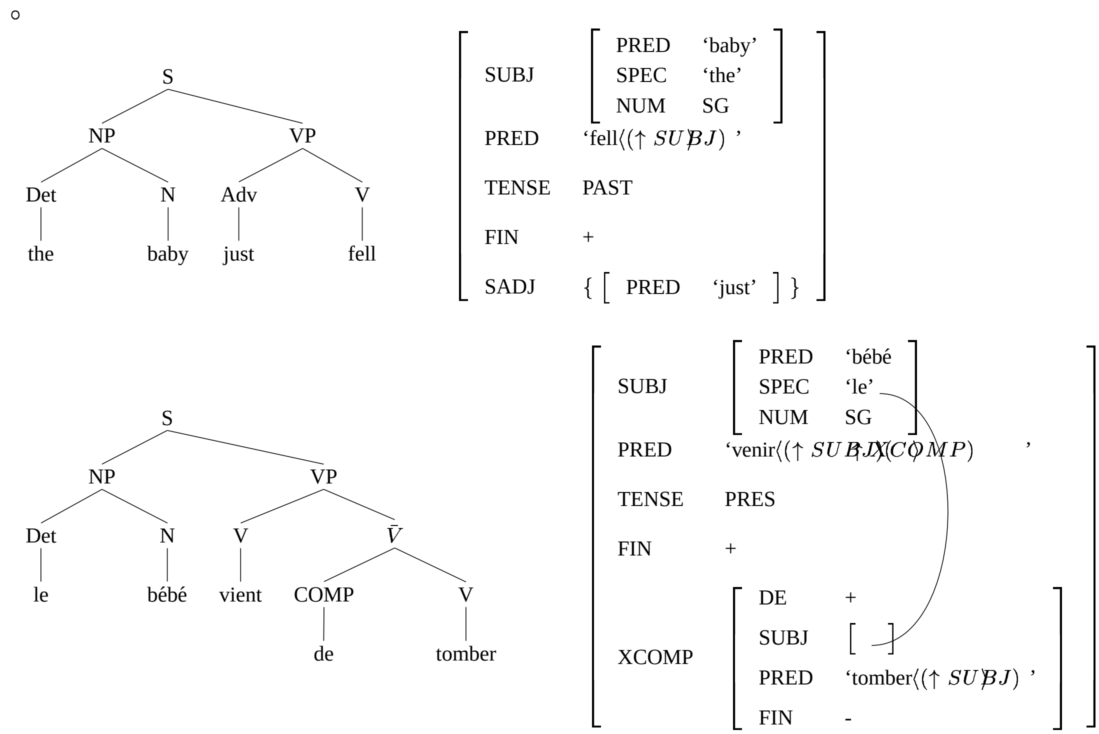
The trees in (41) are linked at S, NP, VP, DET and N levels. Once these are deleted, the remaining fragments are linked at embedded VP level, exactly as in Figure 2. That is, embedded headswitching cases in LFG-DOT3 are dealt with in exactly the same manner as non-embedded headswitching cases. Given that fragments such as Figure 2 and (41) exist, such complex cases can also be dealt with compositionally, as (42) illustrates:

(42) 

$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘I’} \end{array}\right] \\ \text{PRED} & \text{‘think}\langle(\uparrow SUBJ)(\uparrow COMP)\rangle\text{’} \\ \text{TENSE} & \text{PRES} \\ \text{COMP} & \left[\begin{array}{ll} \text{THAT} & + \end{array}\right] \end{array}\right]$$



$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘Je’} \end{array}\right] \\ \text{PRED} & \text{‘penser}\langle(\uparrow SUBJ)(\uparrow COMP)\rangle\text{’} \\ \text{TENSE} & \text{PRES} \\ \text{COMP} & \left[\begin{array}{ll} \text{QUE} & + \end{array}\right] \end{array}\right]$$



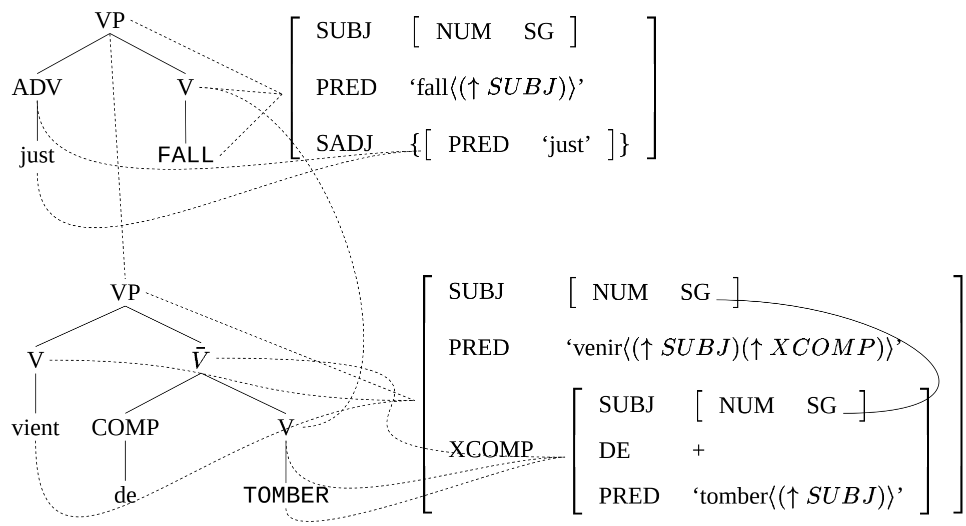
That is, the lower linked source and target sentence pair is substituted into the source and target S-nodes in the upper trees. At the same time, their f-structures are merged with the COMP f-structures of the source and target f-structures respectively. This is, of course, just one possible derivation of this translation. Others will be produced in the usual manner and their probabilities summed in order to derive probabilities for the translation as a whole with respect to the corpus.

2.4 LFG-DOT4: Translation via γ and 'Extended Transfer'

In the previous section, we observed that the outstanding problem with Model 3 is its retention of the DOT2 problem of limited compositionality. Returning to the *just* $\longleftrightarrow$ *venir de* headswitching case in Figure 2, we would like to be able to 'relax' some of the constraints in order to map $\langle fell, tomber \rangle$ to make these linked fragments more general, and hence more useful. In so doing, we would remove this problem of limited compositionality.

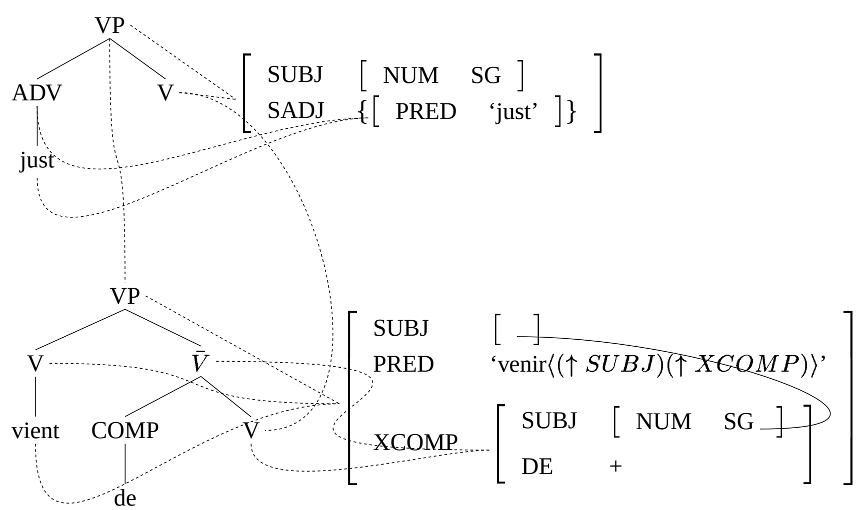
In LFG-DOT4, the basic translation relation is expressed by γ, as with LFG-DOT3. In LFG-DOT4, however, there is a second application of *Discard*, by which 'lemmatized' forms are arrived at on which 'extended transfer' can be performed. *Discard* relaxes constraints in order to produce a set of generalized fragments with the potential to deal with ill-formed or unknown input. Once the `TENSE` and `FIN` features have been relaxed on the lowest verbs in both fragments in Figure 2, they can be regarded as translationally equivalent. Given this, $\langle fell, tomber \rangle$ are linked and lemmatized, as in (43):

(43)



Now that ⟨FALL,TOMBER⟩ are linked, they can be deleted to produce the generalized form of the translation relation, namely (44):

(44)



If fragment pairs such as (44) prove subsequently to be of use in combining with other fragments, any resultant translation will be marked as ungrammatical with respect to the corpus, given that *Discard* was used in its derivation. Nevertheless, even if we restrict the impact of *Discard* on the probability space (cf. Way, 1999, 2001; Bod, 2000), in order to ensure that translations obtained via *Root* and *Frontier* are preferred over those derived via *Discard*, such translations will receive *some* probability, whereas the semi-compositional variants from which they were derived may not be able to produce *any* translation.

3 Contributions and Further Work

We described the two ways in which the original LFG model of translation (Kaplan *et al.*, 1989) attempted to cope with headswitching data. We summarized previous accounts (Sadler *et al.*, 1989, 1990) which showed that LFG-MT cannot cope with certain cases of headswitching. Other attempts at solving these cases using restriction (Kaplan & Wedekind, 1993) and linear logic (Van Genabith *et al.*, 1998) introduce further problems.

We then presented a number of new hybrid models of translation based on LFG-DOP. The first, LFG-DOT1, uses LFG-DOP for the source and target language models, but imports the τ equations from LFG-MT as the translation relation. Accordingly, therefore, it fails to cope with embedded headswitching and other complex data in the same way as the original LFG-MT model. Given this, we augmented LFG-DOT1 with the γ function from DOT2 to give an improved model of translation. Nevertheless, given that τ equations fail to derive the correct translation in all cases, subsequent LFG-DOT models omit the τ function and rely wholly on γ to express the translation relation.

LFG-DOT3 uses f-structure information purely for monolingual filtering. The presence of this functional information prevents the formation of certain ill-formed structures which can be produced in DOT2. LFG-DOT3, therefore, has a notion of grammaticality which is missing from DOT2. Importantly, this can be used to guide the probability models in the manner required. We showed that LFG-DOT3 copes with cases of embedded headswitching in exactly the same way as non-embedded headswitching examples. However, like DOT2, it suffers from the problem of limited compositionality, so that in some cases the minimal statement of the translation relation is impossible.

LFG-DOT4 adds an 'Extended Transfer' phase to LFG-DOT3 by producing lemmatized forms using a second application of *Discard*. This extension overcomes the problem of limited compositionality, enabling the statement of the translation relation in an intuitive, concise fashion.

More generally, we have shown:

1. that LFG-DOP can serve as a model of translation;

2. that contrary to previous perception, basic statements of translation relations can be stated at the level of c-structure, rather than at 'deeper' levels of linguistic analysis, provided that these trees are accompanied by syntactic (f-structure) information which acts as a monolingual filter on the structures produced.

As to future work, the models presented here need to be tested more thoroughly on large-scale LFG-DOT corpora, and the work of Frank *et al*. (2001, this volume) on producing such resources automatically seems promising in this regard. The work described here and in (Way, 2001) uses as its evaluation metric the ability to cope with 'hard' translation cases such as embedded headswitching. Nevertheless, as well as examining such complex translation phenomena, we need to investigate how well our models deal with simpler translation data, such as *Fido barks* sentences. Different probability models will also be evaluated (cf. Bonnema *et al.*, 2000), as will the possibility of pruning the search space, by cutting down the number of fragments produced (cf. Sima'an, 1999) in order to improve the efficiency of the models proposed.

References

[1] Bod, R. (1998): *Beyond Grammar: An Experience-Based Theory of Language*, CSLI Publications, Stanford, California.

[2] Bod, R. (2000): 'An Empirical Evaluation of LFG-DOP', in COLING: *Proceedings of the 19th International Conference on Computational Linguistics*, Saarbrücken, Nancy and Luxembourg.

[3] Bod, R. & R. Kaplan (1998): "A Probabilistic Corpus-Driven Model for Lexical-Functional Analysis", in COLING: Proceedings of the 17th International Conference on Computational Linguistics & 36th Conference of the Association for Computational Linguistics, Montreal, Canada, 1:145–151.

[4] Bonnema, R., P. Buying and R. Scha (2000): 'Parse Tree Probability in Data-Oriented Parsing', in *Proceedings of PACLING*, Mexico City.

[5] Frank, A., J. van Genabith and A. Way (2001): ‘Treebank vs. X-BAR based Automatic F-Structure Annotation’, in *Proceedings of 6th International Conference on Lexical Functional Grammar (LFG-2001)*, Hong Kong.

[6] Kaplan, R., K. Netter, J. Wedekind & A. Zaenen (1989): “Translation by Structural Correspondences”, in Fourth Conference of the European Chapter of the Association for Computational Linguistics, Manchester, pp.272–281.

[7] Kaplan, R. and J. Wedekind (1993): ‘Restriction and Correspondence-based Translation’, in *Sixth Conference of the European Chapter of the Association for Computational Linguistics*, Utrecht, The Netherlands, pp.193–202.

[8] Poutsma, A. (1998): “Data-Oriented Translation”, in Ninth Conference of Computational Linguistics In the Netherlands, Leuven, Belgium.

[9] Poutsma, A. (2000): *Data-Oriented Translation: Using the Data-Oriented Parsing framework for Machine Translation*, MSc thesis, University of Amsterdam, The Netherlands.

[10] Sadler, L., I. Crookston, D. Arnold and A. Way (1990) ‘LFG and Translation’, in *Third Conference on Theoretical and Methodological Issues in MT*, University of Texas, Austin, pp.121–130.

[11] Sadler, L., I. Crookston and A. Way (1989) ‘Co-description, projection, and ‘difficult’ translation’, Working Papers in Language Processing 8, Department of Language and Linguistics, University of Essex, Colchester.

[12] Sadler, L., and H. Thompson (1991): “Structural Non-correspondence in Translation”, in *Fifth European Conference on Computational Linguistics*, Berlin, Germany, pp.293–298.

[13] Sima’an, K. (1999): *Learning Efficient Disambiguation*, PhD Thesis, University of Utrecht, The Netherlands.

[14] Van Genabith, J., A. Frank and M. Dorna (1998): ‘Transfer Constructors’, in *Proceedings of LFG-98*, Brisbane, Australia, pp.190–205.

[15] Way, A. (1999): “A hybrid architechture for robust MT using LFG-DOP”, *Journal of Experimental and Theoretical Artificial Intelligence* **11**:441–471.

[16] Way, A. (2001): *LFG-DOT: A Hybrid Architechture for Robust MT*, PhD thesis, the University of Essex, Colchester, UK.

Complex-predicate Formation and Some Consequences in Japanese

Kenji Yokota
University of Tokyo

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)
2001

CSLI Publications
http://csli-publications.stanford.edu/

1. Introduction

The debate has not yet been settled as to whether Japanese morphological causative structures are 'bi-clausal' or 'mono-clausal'.[1] Within the framework of LFG, Ishikawa (1985) and Matsumoto (1996)[2] have claimed that the causatives are bi-clausal in f-structure (involving an XCOMP), and mono-clausal in c-structure.

In this paper I would like to argue that the recent LFG approach to complex predicates in various languages (cf. Alsina 1996, Butt 1995, etc) actually holds good in Japanese causatives and several other related constructions; i.e., these predicates exhibit the property of mono-clausality in f-structure, and bi-clausality in c-structure, contrary to Ishikawa's and Matsumoto's assumption. It is also

---

[1] Outside LFG framework, recent studies of Japanese causatives include Gunji (1999), Manning, et al.(1999) in HPSG, Harley (1995) in GB/MP.

[2] In Matsumoto (1996) he argued that in Japanese morphological causatives differ in their bi-clausal properties between coercive (*make*) causatives and permissive (*let*) causatives; the former are mono-clausal and the latter are bi-clausal at f-structure. However, as we will see shortly in the following sections, this distinction seems rather unmotivated, unnecessary in analyzing morphological causatives in the language.

argued that the determination of semantic scope should be treated in terms of c-structure as well as f-structure. We lastly argue for a claim that the causative verb -*(s)ase*[3] is certainly a verb, not an affix.

2. Evidence for functional mono-clausality

Japanese morphological causative verbs are formed by adding *sase* to a verb stem, as shown in (1). The causer is marked with the nominative case (-*ga*), and the causee is marked with the dative particle (-*ni*) or optionally the accusative case (-*o*) if the stem is intransitive.

(1) Hanako ga musume ni sono hon o yom-ase-ta.
Hanako Nom daughter Dat that book Acc read-Caus-Past
'Hanako made / let her daughter read that book.'

However, as will be discussed in Section 3, *sase* is certainly a verb, not a simple affix, if not a full-fledged verb.

2.1. Reexamination of Matsumoto (1998, 2000)

In his latest study, Matsumoto (1998, 2000) argues that there exist clear differences in f-structure between permissive causatives and coercive causatives; i.e., Though in Japanese both types of causative have mono-clausal c-structure, the coercive causative has mono-clausal f-structure and the permissive causative has bi-clausal f-structure. This assumption, however, is not correct. I will reexamine his data in order, and then alternatively propose that Japanese morphological causatives are definitely mono-clausal in f-structure, and try to argue for its bi-clausality in c-structure, though they apparently appear synthetically in the surface form (ex. *yom-ase-ta* 'read-Caus- Past' in (1)).

I agree with the assumption that causatives can be roughly divided into two types; coercive and permissive. I, however, assume that the Japanese causative verb *sase* has basically or intrinsically coercive reading, (i.e., mono-clausal in f-

---

[3] *Sase* is attached to a verb stem which ends with vowel, while *ase* is attached to a verb stem with consonant.

structure), though, of course, the inherent meaning of the base verb or pragmatic factors can sometimes decide which reading. Then, permissive reading often derives from the context, outside syntax. An adverbial element, for example, might be added to produce unambiguously either the permissive reading or the coercive reading, as exemplified in (2). In the sentences, -*o* and -*ni* are used as coercive and permissive, respectively.

(2) a. Taroo wa <u>iyagaru</u> musuko o / ??ni puuru de oyog-ase-ta
Taro Top be reluctant son Acc / Dat pool at. swim-Caus-Past
'Taro made / ??let his son, who is reluctant to do, swim in the pool.'

b. <u>Oyogi-tai to iu-node</u>, Taroo wa musuko ?o / ni puuru de.
swim-want Comp say-because, Taro Top son Acc / Dat pool
oyog-ase-ta
at swim-Caus-Past
'Taro ??made / let his son swim in the pool, because his son said he wanted to do so.'

The example (2) suggests that permissive reading itself need not be described in terms of syntax, which means Matsumoto's structural distinction of causatives between coercive and permissive is rather dubious. This is why I will not distinguish the syntactic structures of the two types of Japanese causatives.

The data are reexamined in what follows, which Mastumoto (2000) presents as strong evidence for the mono-clausality of the coercive causative and bi-clausality of the permissive causative in f-structure. Our analysis, however, will show that the supposedly crucial data for him are no longer strong evidence for functional bi-clausality.

2.1.1. Subject Honorification

According to Mastumoto, the best diagnostic for identifying grammatical subjects in Japanese is the test employing honorific marking, in which a certain marking on the verb indicates the speaker's sense of respect toward the grammatical subject of the verb.

In (3) a marking of *o*-V-*ni nar* ‘Hon-V-Cop become’is used as a test for the subjecthood of a causee NP.

(3) a. Permissive:

Daijin wa ooji ni sono kutsushita o *o*-**haki** *ni*
minister Top prince Dat the socks Acc Hon-put.on Cop *nar*-**ase** te ok-
are-mashi-ta.
beome-Caus leave-Hon-Pol-Past
‘The minister let the Prince (continue to) put the socks on his feet.’

b. Coercive:

??Karera wa muriyari ooji ni sono kutsushita o *o*-**haki**
they Top focibly prince Dat the socks Acc Hon-put.on *ni* *nar*-**ase**-rare-
mashi-ta.
Cop become-Caus-Hon-Pol-Past
‘They forcibly made the Prince put the socks on his feet.’
(Matsumoto 2000; 147)

The subject-honorified form can be followed by the causative, in which case it is the object, not the subject, that is honored. Matsumoto’s argument, based on the observation in (3), is as follows; The object *ooji* ‘prince’ is definitely a grammatical subject and so the permissive causative is bi-clausal at f-structure.

The test in (3), however, is not reliable in order to distinguish the two types of causative. First, in (3b) subject honorification marker -*(r)are* ‘Hon’ is placed on the main verb (i.e., on the causative verb in (3b)) to indicate the speaker’s respect toward the causer. This additional morpheme, as Matsumoto states, is added for pragmatic naturalness. This, however, seems to make the sentences in question unnecessarily more complicated, which might lead to the unacceptability of (3b). This does not mean that the marker to be put right after *sase* must be -*(r)are*. In fact, if the affix is removed, or if it is replaced by -*te o-oki-ni nari- mashi-ta*, which is another honorific pattern of -*te ok*, then the acceptability of the coercive pattern increases as shown in (4). Note that we can get coercive reading in both (4a) and (4b).

(4) coercive:

a. Daijin wa muriyari ooji ni sono kutsushita o *o*-**haki** .
minister Top forcibly prince Dat the socks Acc Hon-put.on
ni nar-**ase**-ta
Cop become-Caus-Past
'They forcibly made the Prince put the socks on his feet.'

b. Daijin wa muriyari ooji ni sono kutsushita o *o*-**haki**
minister Top forcibly prince Dat the socks Acc Hon-put.on
ni nar-**ase**-te o-oki-ni nari-mashi-ta.
Cop nar-Caus Hon-leave-Cop become-Pol-Past
'They forcibly made the Prince put the socks on his feet.'

Second, which is more important, in (3b) and elsewhere in Matsumoto's paper the permissive causative in his examples employ ***-te oku*** 'particle-leave,' which is placed on the causative verb so that the sentence in question will unambiguously produce permissive reading. I, however, assume this additional element is a two place control predicate like *-te hoshii* 'particle-want', *-te morau* 'particle-receive'[4], subcategorizing for VP. This raises a serious problem; in his data the (un)acceptability on the permissive causative are affected by the verb *-te oku*, which situation does not elucidate the contrastive nature of the two types of causatives. The construction of (3a) is different from that of (3b), which does not constitute a minimal pair any longer, and then the status of the two predicates in (3a) and (3b) are different. The verb *-te oku* seems to be responsible for the possible reading in (3a). From this reason throughout the paper when I discuss the permissive causative I will use adverbials indicating only permission like *(soo-)sitai-to iu-node* 'saying s/he wants to do so' or *iiyo to itte* 'saying yes you can', so that we can make fair judgments of sentences. The desirable sentence in our discussion would be as follows. Note that the acceptability is the same as (3a).

(4) Permissive:
Daijin wa ooji ni <u>soo-si-tai to iunode</u> sono kutsushita o *o*-haki *ni*

[4] See Matsumoto (1996; Chapter 3) for an analysis of particle *-te* + *hoshii / morau,* where he provides

nar-ase-ta / nar-ase-te o-oki-ni nari-mashi-ta .

'The minister let the prince, saying he wants to do so, put the socks on his feet.'

The honorification pattern exhibited in (4) will support the mono-clausality in f-structure of both the coercive and permissive causative in Japanese.

Thus, it is reasonable to consider that it is unnecessary to posit two different types of causative verb -*sase* as Matsumoto claims. If subject honorification has nothing to do with subjecthood in terms of grammatical subject but rather something to do with logical subject or pragmatic factor(s), re-examination of previous studies of causative constructions, which depends upon the test, might be needed.

2.1.2. Passivization

In LFG it is widely assumed that the arguments of both the causative verb and the base verb map onto grammatical functions (i.e., f-structure) of a single predicate in f-structure.[5] In (5) - (7) the direct object of the base verb can be long-distance passivized. This certainly confirms the mono-clausality of f-structure with Japanese causatives, as Matsumoto argues.

(5) Sono rinyuu-shoku wa mada dono akachan ni mo
the baby.food Top yet any baby Dat even
tabe-sase-rare-te i-nai.
eat-Caus-Pass Asp-Neg
'The baby food has not yet been given to feed any child.'
(Matsumoto 2000; 148)

(6) (Yotei yori ichi-jikan mo okurete yatto) kaikai ga

---

them with bi-clausality of f- and c-structures.

5 Regarding Japanese causatives, as many researchers have pointed out, not every direct object of the base verb can be a passive subject. I do not know exactly why. Probably, some pragmatic constraint for the (im)possibility of the passive construction need to be taken into account, which should be dealt with outside syntax. (cf.Takami & Kuno 1993)

schedule than one-hour as.much.as behind finally opening Nom John ni-yotte
gichoo ni sengen sase-rare-ta.
John by chairperson Dat announce Caus-Pass-Past
'(An hour behind the schedule, finally) the opening of the meeting
was made the chairperson to announce by John.'

(7) Taro ga wazato korob-ase-rare-ta.
Taro Nom on.purpose fall.down-Caus-Pass-Past
'Taro was caused, on purpose, to fall down.'

2.1.3. Pronominal Binding

Mastumoto does not use the reflexive *jibun* 'self' as a test for the subjecthood of a causee NP. (See Iida (1996) for non-grammatical factors involved in *jibun* binding.) Instead, he uses a pronoun *kare* 'he/his/him' as a test for distinguishing the two types of causatives. Consider (8).

(8) a. Taro$_i$ wa Jiro$_j$ ni sonomama kare$_{i,*j}$ / karejisin$_{*i,j}$ o bengo
Taro Top Jiro Dat as.it.is he/himself Acc defend
s-ase-te oi-ta.
do-Caus put-Past
'Taro let Jiro continue to defend him(self).'

b. Taro$_i$ wa Jiro$_j$ ni muriyari kare$_{??i,*j}$ / karejishin$_{i,j}$ o
Taro Top Jiro Dat forcibly he/himself Acc
bengo sase-ta.
. defend do-Caus-Past.
'Taro forcibly made Jiro defend him(self).'
(Matsumoto 2000; 148)

In LFG account, the pronoun *kare* must be referentially disjoint from its clausemate in f-structure. In causatives, the pronoun may be bound by the subject but must be disjoint in reference with the causee. (cf. Kitagawa 1986) However, I would like to argue that the pronominal *kare* -binding fact does not really support Matsumoto's assumption that the permissive causative has a bi-clausal f-struture and the coercive causative has a mono-clausal f-structure. With regard to (8) he argues the pronominal binding to the causee is not possible. The reflexive for *karejishin* 'himself', in contrast, must be bound within its clause.

As we have seen in 2.1.1, (8) does not constitute a minimal pair, either; in (8a) an additional word *-te oku* is placed on the causative verb *sase*. The problem is, it is highly likely that the verb might be responsible for the possible pronominal reading in (8a). The pronominal binding to the causee is actually allowed in both coercive and permissive reading as shown in (9).

(9) a. Taro$_i$ wa Jiro$_j$ ni iiyo to itte mina no maede kare$_{i,(?)j}$ /
Taro Top Jiro Dat saying OK everyone Gen front him /
karejisin$_{i,j}$ o hinan sase-ta.
himeself Acc criticize do.Caus-Past
'Taro let Jiro blame him(self) in front of others, saying OK.'

b. Taro$_i$ wa Jiro$_j$ ni muriyari mina no maede kare$_{i,(?)j}$ /
Taro Top Jiro Dat forcibly everyone Gen front him /
karejishin$_{i.j}$ o hinan sase-ta.
himself Acc criticize do.Caus-Past
. 'Taro forcibly made Jiro blame him(self), in front of others.'

In (9) the pronominal expression *kare* 'him' can be the causer (Taro) as well as the causee (Jiro). According to Matsumoto's assumption, this pronominal should exhibit behavior which is complementary to the reflexive *karejishin*, this, however, is not the case as indicated in (9).

The acceptability in (8) and (9) suggests that coreference relation, whether it is of *kare* 'him' or *karejishin* 'himself', cannot be simply determined in f-structure; it might also be involved in a-structure and c-structure, and sometimes pragmatic factor(s).[6] Thus, the pronominal binding test again does not clearly argue for the difference between mono-clausality and bi-clausality of f-structure.

2.1.4. Control

It is well-known that the subject of certain adverbial control clauses can be controlled by either grammatical or logical subject.

(10) a. John$_i$ wa sono ko$_j$ ni [PRO$_{?i,j}$ terebi o mi-nagara] sono kutsusita
John Top the child Dat [television Acc watch-while] the socks
o hak-ase-te oita.
Acc put.on-Caus left.
'John let the child put on the socks, watching TV.'

b. John$_i$ wa muriyari sono ko$_j$ ni [PRO$_{i,j}$ terebi o mi-nagara] sono
John Top forcibly the child Dat [television Acc watch-while] the
kutsushita o hak-ase-ta.
socks Acc put.on-Caus-Past
'John forcibly made the child put on the socks, watching TV.'
(Matsumoto 2000; 149)

Here too, the structure of the permissive causative that Matsumoto gives in his argument is different from that of the coercive causative. From this reason, I will continue to use an adverbial which provide an unambiguous meaning of permission. Then, compare (10) with (11).

(11) a. John$_i$ wa <u>iiyo to i-tte</u> musume$_j$ ni [PRO$_{i,j}$ terebi o mi-nagara]
John Top OK Comp say daughter Dat [TV Acc watch-while]
sono kutsushita o hak-ase-ta.
the socks Acc put.on-Caus-Past
'Saying 'yes you can', John let his daughter put on the socks,
watching TV.'

[6] See Bresnan (1995), Choi (1996) for a treatment of operator binding within Optimality Theory.

b. John$_i$ wa muriyari musume$_j$ ni [PRO$_{i,j}$ terebi o mi-nagara]
John Top forcibly daughter Dat [TV Acc watch-while]
sono kutsushita o hak-ase-ta.
the socks Acc put.on-Caus-Past
'John forcibly made his daughter put on the socks, watching TV.'

In (11a) and (11b), there exist no different behaviors with respect to the interpretation of PRO. Interestingly, when the control clause is placed at the beginning of the sentence, the interpretation differs greatly. In (12), PRO is interpreted as only John.

(12) [PRO$_{i,*j}$ Terebi o mi-nagara] John$_i$ wa iiyo to itte musume$_j$
[TV Acc watch-while] John Top saying OK Comp daughter
ni sono kutsushita o hak-ase-ta.
Dat the socks Acc put.on-Caus-Past
'Watching TV, John let his daughter put on the socks, saying 'yes you can'.'

The contrast observed between (11a) and (12) might pose quite serious a problem for the standard treatment of adverbials at f-structure, because f-structure does not tell us anything about word order.[7] (cf. Andrews 1983, Andres and Manning 1999, etc) If we want to get the proper semantic interpretation of the examples in (11) and (12), we need take c-structure into consideration in some way, because the semantic interpretation clearly reflects the phrase structure (c-structure). In (11a) and (11b) the adverbial clause is within VP, whereas in (12) outside VP. We now assume that c-structure as well as f-structure should be participated in the semantic interpretation.

Thus, the control test does not support the structural distinction of

[7] So do the examples involving adjuncts given in 2.1.5.

the permissive and coercive causatives. The same argument provided here is hold for the interpretation of adjuncts, which I will consider in the next section.

2.1.5. Adjunct Interpretation

Following Jackendoff (1972), Matsumoto argues that adverbials like *ooyorokobide* 'happily'or *shibushibu* 'reluctantly' can be interpreted with respect to either grammatical or logical subject when placed inside VP. Based on the assumption, Matsumoto (2000) presents (13).

(13) a. John wa sono ko ni sono kutsushita o <u>ooyorokobide</u>
John Top the child Dat the socks Acc happily
hak-ase-te oita.
put.on-Caus left
'John let the child put on his socks(,) happily.'

b. John wa muriyari sono ko ni sono kustushita o <u>shibushibu</u>
John Top forcibly the child Dat the socks Acc reluctantly
hak-ase-ta.
put.on-Caus-Past
'John forcibly made the child put on his socks(,) reluctantly.'
(Mastumoto 2000; 151)

Matsumoto (2000) actually focuses on the analysis of lexical *sase* causatives comparing with morphological causatives, so he was not considering or even mentioning the behavior exhibited in the_ permissive (13a) and the coercive (13b). It is not clear how these sentences including adjuncts can be treated properly, with Matsumoto's assumption that the permissive causatives have a bi-clausal f-structure, which has two PRED(icate)s, whereas the coercive causatives mono-clausal f-structure, which has a single PRED. But he is not able to account for the case where the position of an adverb is placed outside VP.[8]

(14) Ooyorokobide John wa sono ko ni sono kutsushita o hak-asete
happily John Top the child Dat the socks Acc put.on-Caus
oita.
left
'John happily let the child put on his socks. / *John let the child put on the socks happily.'

The following example also would support our claim that the ambiguity is more of a structural nature than a lexical nature.

(15) a. Taro ga Jiro ni sono hon o damatte yom-ase-ta. <ambiguous>
Taro Nom Jiro Dat the book Acc silently read-Caus-Past.
'Taro silently made / let Jiro read the book. / Taro made / let read the book silently.'

b. Damatte Taro ga Jiro ni sono hon o yom-ase-ta. <unambiguous>
silently Taro Nom Jiro Dat the book Acc read-Caus-Past.
'Taro silently made / let Jiro read the book. / *Taro made / let read the book silently'

In our analysis, Japanese causatives have mono-clausal f-structures. Throughout the discussion I adopt an important assumption concerning adjunct interpretation by Andrews and Manning (1999).

(16) Grammatical relation spreading:
"[I]f we regard complex predicates as a syntactically mono-clausal domain of grammatical relation spreading, the right notion is that adjuncts must scope

---

[8] There might be apparent counterexamples like (i).
(i) Yukkuri to Ken ga Naomi o suwar-ase-ta.
'Ken slowly made Naomi sit. / ??Ken slowly made Naomi sit slowly.'

Regarding (i), one might think that it is ambiguous; the adjunct *yukkurito* 'slowly' can be interpreted with respect to both Ken (causer) and Naomi (causee). However, the latter interpretation can be obtained only when some pause is put right after the adjunct. In such cases the adjunct clause may have been long-distance scrambled. If so, this does not affect the analysis proposed in this paper.

within the domain in which grammatical relations are shared."

(Andrews and Manning (1999; 55))

Given (16), in the c-structure if the two verbs form a complex predicate, then the complement VP should be the domain of grammatical relation spreading. In cases like this, semantic ambiguity will arise if an adverb appears within the VP domain.[9]

Examining the facts about the interpretation of adverbs (ex (13), (14) and (15)), I propose that c-structure as well as f-structure should participate in adverb interpretation.

2.1.5.1. The two types of syntactic compound verbs

As many researchers including Matsumoto point out, not all control / raising verbs exhibit the functional mono- clausality of complex predicates like *sase*. Following such researchers, I will divide syntactic compound verbs in Japanese into two types. According to Kageyama (1993, 1999), the primary motivation for establishing the two structures resides in the passivization of a whole (compound) verb and scope ambiguity facts.

The first are *Complex Predicates*, which are genuinely verb complex and exhibit functional mono-clausality, including -*sase-ru* 'Caus-Pres', -*hajimeru* 'begin', -*naosu* 'do again',-*oeru* 'finish', -*tukusu* 'exhaust, do thoroughly', etc.

(17) V1+*V2* = complex predicates:

a. Kare wa suupu o atatame-*naoshi*-ta

he Top soup Acc heat-do.again-Past

'He reheated the soup.'

---

[9] Technically, to handle possible adverb interpretations we will have to need combining the three assumptions;(i) the existence of a VP in Japanese, (ii) a requirement that the adverbs c-command their 'understood subjects' (experiencer of happiness (13a), displayer of reluctance (13b), etc), (iii) a mechanism such as the glue-logic scheme proposed by Andrews and Manning so that an adverb in a clause with a mono-clausality composed complex predicate can modify any level of the complex predication, provided that any other relevant constraints are satisfied. (Avery Andrews, p.c.)

b. suupu ga atatame-*naos*-are-ta.

soup Nom heat-do.again-Pass-Past

'The soup was reheated.'

As (17b) shows, in the construction the direct object of the base verb can be usually passivized. Consider then the adverb interpretation.

(18) Hanako ga biiru o nomi nagara tegami o kaki-*hajime*-.

Hanako Nom beer Acc drinking while letter Acc write-begin-

ta.

Past

'Hanako began writing a letter while drinking beer.'

(i) At some point during Hanako's drinking of beer, Hanako's writing of a letter was began. [V1]

(ii) Hanako began the action of drinking of beer and writing a letter simultaneously. [V1+V2]

(19) Kanojo wa teineini aisatsu shi-*naoshi*-ta.

she Top politely bow do-do.again-Past

(i) 'She made a deep bow again.' [V1+V2]

(ii) 'She had not made a deep bow, but this time she made a deep bow.' [V1]

The second are *XCOMP Control Predicates*, which do not form complex predicates and have an XCOMP complement, including -*sokoneru* 'miss', -*sobireru* 'miss the chance', -*akiru* 'become weary', -*tukeru* 'be accustomed to', -*kaneru* 'hesitate', etc.

(20) V1+*V2* = XCOMP control predicates:

a. Keeki o tabe-*sokone*-ta.

cake Acc eat-miss-Past

'(I) missed eating a cake.'

b. *Keeki ga tabe-*sokone*-rare-ta.

cake Nom eat-miss-Pass-Past

'A cake was missed eating.' (intended)

(20b) shows that the direct object of the base verb cannot be passivized in such constructions.

(21) Taro wa <u>biiru o nomi-nagara</u> sakana o tabe-*wasure*-ta.
Taro Top beer Acc drinking-while fish Acc eat-forget-Past
'Taro forgot eating some fish, while drinking beer.'
*(i) At some point during Taro's drinking of beer, Taro's eating of some fish was forgotten. [V1]
(ii) Taro forgot the action of drinking of beer and eating some fish simultaneously. [V1+V2]

(22) a. Shoogatsu-ni <u>kimono-de</u> dekake-*sokone*-ta.
New Year-at kimono-in go.out-miss-Past
(i) 'I missed going out in kimono at the New Year.' [V1+V2]
*(ii) 'The action of going out in kimono was not done at the New Year' (intended) [V1]

b. Kare ni sono shirase o <u>chokusetsu</u> shirase-*sobire*-ta.
he Dat the news Acc directly tell-fail-Past
(i) 'I failed to tell him the news directly.' [V1+V2]
*(ii) 'The action of telling him the news directly was not done.' (intended) [V1]

Although these predicates appear the same structures in c-structure, a complex predicate is mono-clausal in f-structure, while an XCOMP predicate is bi-clausal in f-structure. The information of both levels would participate in determining the semantic scope. So if a syntactic compound verb does not form a complex predicate, scope ambiguity involving adverbs would not arise even when some adverb is placed inside VP. Consequently, in such cases where grammatical relations are not shared, the adverb will modify unambiguously the whole VP; V1+V2.

2.1.5.2. *Shika-na(i)*

The causative pattern involving the *shika-na* 'only-Neg' construction will produce semantic ambiguity.

(23) Taro ga Hanako ni kyooiku terebi *shika* mi-sase-*nakat*-
Taro Nom Hanako Dat education channel only watch-Caus-Neg-
ta. [ambiguous]
Past
(i)Taro didn't cause Naomi to watch other things.
(cf. kyooiku terebi-*shika* [$_{VP}$ mi]-s-ase-*nakat*-ta)
(ii) Taro didn't cause Naomi to do other things.
(cf. [$_{VP}$ kyooiku terebi-*shika* mi]-s-ase-*nakat*-ta)

Note that this kind of ambiguity does not arise for the lexical ditransitive verb *mise* 'show', for which there can be no bi-clausal structure (cf. Gunji 1999).

(24) a. Ken ga Naomi ni kyooiku terebi-*shika* **mise**-*nakat*-ta.
Ken Nom Naomi Dat education channel-only show-Neg-Past
'Ken didn't show Naomi other things (than watching the channel).' [unambiguous]
b. Taro ga Saburo o <u>yukkuri</u> **osi-taosi**-ta.
Taro Nom Saburo Acc slowly push-topple-Past
'Taro pushed and toppled Saburo, slowly.' [unamiguous]
c. Ookami ga niwatori ni gaburi to **kami-tsui**-ta.
wolf-Nom Nom chicken Dat mimetics bite-attach-Past
'The wolf bit at the chicken.' [unambiguous]

I have argued that the ambiguity is more of a structural nature than a lexical nature.

2.2. More evidence for functional mono-clausality

So far we have seen that all the diagnostic tests given by Matsumoto do not support the f-structure differences between the permissive and coercive causative. Rather, it is reasonable to say that both are uniformly treated as functional mono-clausal,which means that the causative verb *sase* certainly constitute a complex predicate with the base verb at syntax.

In this section, we additionally present another syntactic evidence for the mono-clausality of Japanese causatives. The further evidence comes from double *o* constraint, and *shika-na(i)* construction.

2.2.1. Double-*o* constraint[10]

I will argue that the 'deep' double-*o* constraint can be used as a test to identify functional mono-clausality. It has been pointed out in the literature that there are two types of double *o* constraint; the 'deep' double *o* constraint and the 'surface' one. The former is shown as (25), and the latter is shown as (26), respectively

(25) a. Taro ga kooen o arui-ta.
Taro Nom park Acc walk-Past
'Taro walked in the park.'

b. Taro ga Jiro *o / ni kooen o aruk-ase-ta.
Taro Nom Jiro Acc / Dat park Acc walk-Caus-Past
'Taro made Jiro walk in the park.'

(26) a. Taro ga hon o yon-da.
Taro Nom book read-Past
'Taro read the book. '

b. Taro ga Jiro *o / ni hon o yom-ase-ta.
Taro Nom Jiro Acc / Dat book Acc read-Caus-Past
'Taro made Jiro read the book.'

Although (25b) and (26b) appear to show no difference, a difference, as Kuroda (1978) notes, emerges if one forms a so-called pseudo cleft sentences. In these

[10] See Kuroda (1978), Poser (1983) and Sells (1989) for detailed discussion on the two different double-

pseudo-cleft sentences, the NP that immediately precedes *da* 'is' the focus of the sentence. Notice the contrast between (27) and (28).

(27) Taro ga Jiro o aruk-ase-ta no wa kooen (o) da
Taro Nom Jiro Acc walk-Caus-Past Gen Top park (Acc) is
'It is the park that Taro made Jiro walk in.'

(28) *Taro ga Jiro o yom-ase-ta no wa hon (o) da.
Taro Nom Jiro Acc read-Caus-Past Gen Top book (Acc) is
'It is a book that Taro made Jiro read.'

The contrast indicates that although the double-*o* constraint equally rules out (25b) and (26b), as we observed, the NPs marked with *o* are different in nature. In (25b) both NPs are not objects, while in (26b) both NPs are objects. The ungrammaticality of (25b) is accounted for if we regard the verb *yom-ase-ta* 'read-Cause-Past' in (26b) as a single predicate. Keeping this in mind consider the causative cases.

If Japanese morphological causatives were functionally bi-clausal, then the two objects would appear in two different clauses in f-structure, and so the deep double *o* constraint, which checks the number of direct objects in a single clause, does not rule out such a case. To handle this problem, Matsumoto argues that the permissive causative does not rule out the occurrence of two objects, while the coercive causative does.

(29) a. ??Taro wa Hanako o sono-mama hon o yom-asete oita.
Taro Top Hanako Acc as.it.is book Acc read-Caus left
'Taro let Hanako continue to read a book.'

b. *Taro wa Hanako o muriyari sono hon o yom-ase-ta.
Taro Top Hanako Acc forcibly the book Acc read-Caus-Past
'Taro forcibly made Hanako read the book.'
(Matsumoto 1996; 151)

---

o constraints.

It should be noted that in (29a) *Hanako o* ‘Hanako Acc’ might be an argument of *te-oita* ‘particle-left’, which I believe is a control verb. Given that the double-*o* constraint applies only within a single clause, there should be nothing strange about (29a) better than (29b), because the NPs might belong to different clauses.

Thus, for the present purposes, (29a) should be look like (30a) below. The sentence to be compared with (29b) as the permissive causative would be that in (30a), which is as bad as the coercive causative (29b). If we omit *sono hon o* ‘the book Acc’ as in (30b), then the sentence will be bad. This cannot be accounted for in terms of f-structure.

(30) a. *Taro wa Hanako o iiyo to itte sono hon o yom-ase-ta.
Taro Top Hanako Acc OK Comp say book Acc read-Caus-Past
‘Taro let Hanako read the book, saying, ‘Yes you can.’. ’

b. Taro wa Hanako o iiyo to itte *(sono hon o) yom-ase-ta.
‘Taro let Hanako read the book, saying, ‘Yes you can.’. ’

Mastumoto also notes that some speakers find that the replacement of *o* by *mo* more significantly improves double-*o* permissive causative sentences, indicating that what is involved here is the “surface” double-*o* constraint, while this is not the case with coercive causatives, as given in (31).

(31) a. (?)Taro wa Hanako mo sono-mama sono hon o yom-ase-te oita.
Taro Top Hanako too as.it.is the book Acc read-Caus left
‘Taro let Hanako continue to read that book, too.’

b. * Taro wa Hanako mo muriyari sono hon o yom-ase-ta.
Taro Top Hanako too forcibly the book Acc read-Caus-Past
‘Taro forcibly made Hanako read that book.’
(Matsumoto 1996; 151)

In the next example, however, the object *sono hon* ‘that book’ in (31) has been topicalized, so that it does not take *o*. (32a) is as bad as (32b). The double-*o* constraint is violated, which suggests that the deep double-*o* constraint is really involved in the causatives.

(32) a. *Sono hon wa Taro ga Hanako mo iiyo to itte yom-ase- ta.
the book Top Taro Nom Hanako too OK Comp say read-Caus-Past
'That book, Taro let Hanako read it, too.'

b. *Sono hon wa Taro ga Hanako mo muriyari yom-ase-ta.
the book Top Taro Nom Hanako too forcibly read-Caus-Past
'That book, Taro forcibly made Hanako read it, too.'

In this subsection I have argued that both coercive and permissive causatives are in fact subject to the 'deep' double *o* constraint, which prohibits more than two objects occurring in a single clause.

2.2.2. *Shika-na(i)* construction

There is another piece of evidence for the mono-clausal structure of the causatives in Japanese. I will argue that the *shika-na(i)* 'only-Neg' sequence, serves as a diagnostic test for functional mono-clausality. It has been argued that *shika* can only be added to a phrase that is in the same clause as its associated negative marker (the Locality Condition)[11]. Consider the contrast between sentences like (33a) and (33b).

(33) a. Taro wa Tokyo e *shika* ik-a*na*kat-ta.
Taro Top Tokyo Goal only go-Neg-Past
'Taro went to Tokyo only.'

b. *Hanako wa [Taro ga Tokyo e *shika* itta] to iw- a*nakat*-ta.
Hanako Top Taro Nom Tokyo Goal only went Comp say- Neg-Past
'Hanako said that Taro went to Tokyo only' (intented)

[11] See Muraki (1978), Kato (1985, 1993) and Sells (1996) for detailed discussion on the *sika-na(i)* construction, and references therein.

Next consider the following example.

(34) a. Hanako wa musume ni hoka no hon o yomi-tai-to
Hanko Top daughter Dat other Gen book Acc read-want-Comp
itta noni sono hon-*sika* yom-ase-*nakat*-ta.
said though the book-only read-Caus-Neg-Past
'Hanako made her daughter read only the book, though her daughter said she wanted to read another book. '

b. Hanako wa musume ni hoka no hon wa yomi-tak-unai
Hanako Top daughter Dat other Gen book Top read-want-not
to iu node sono hon-*sika* yom-ase-*nakat*-ta.
Comp say because book-only read-Caus-Neg-Past
'Hanako let her daughter read only the book, for her daughter said she didn't want to read another book.'

If we distinguish between a permissive causative and a coercive one in f-structure, then it is difficult to account for (35). The grammaticality of (35a) and (35b) suggests that both *shika* and *na(i)* are within the same clause at its surface structure. Mastumoto explains (35), based on the assumption that the clause-mate condition on *shika* --- *na(i)* must be met at the surface structure, i.e., c-structure. As a piece of evidence, he gives the following example.

(35) a. John wa [PRO BLS ni *shika* ik-a*naka*-tta koto] ga aru.
John Top [BLS Goal only go-Neg-Past Comp] Nom have
'John has the experience of going only to BLS (among many conferences in a year).'

b. ??BLS ni *shika* John wa [PRO ik-a*naka*-tta koto] ga aru.
BLS Goal only John Top [go-Neg-Past Comp] Nom have
'John has the experience of going only to BLS (among many conferences in a year).' (intented)

c. BLS ni *shika* Jon wa [PRO i-tta] koto ga *na*i.
BLS Goal John Top [go-Past] Comp Nom have.Neg
'John does not have the experience of going to conferences other than BLS.'
(Mastumoto 1996; 38)

Regarding (35b), however, my informants all judge the sentence as grammatical. Matsumoto's assumption that the distribution of *shika* --- *na(i)* is sensitive to c-structure configuration fails to account for sentences like (36), where *shika* --- *na(i)* is not within the same clause.(cited from Sells (1996))

(36) a. Taro wa [Hanako ni Tanaka sensee o syookai-*shika*
Taro Top [Hanko Dat Tanaka teacher Acc introduce-only
suru] tumori-wa *na*i.
do intend-Foc Neg.
'Taro intends only to introduce Professor Tanaka to Hanako.'

b. Boku wa [Hanako *shika* soko ni iru to] omow-a*nak*-atta.
I Top [Hanako only there in be Comp] think-Neg-Past
'I thought that only Hanako was there.'

c. Gakko de *shika* John wa [Bill ga benkyoo shi-*na*i to] omotte-iru.
school at only John Top [Bill ga study do-Neg to] think-Pres
'John thinks that Bill studies at school only.'

Furthermore, the following involving long-distance scrambling of adjuncts might be counterexamples to Matsumoto's view (cited from Sugisaki (2001; 387- 388)).

(37) a. Mary ga [John ga yukkurito booru o nageta to] itta.
Mary Nom John Nom slowly ball Acc threw Comp said
'Mary said that John slowly threw a ball.'

b. Yukkurito Mary ga [John ga booru o nageta to] itta.
slowly Mary Nom John Nom ball Acc threw Comp said

'*Mary said that John slowly threw a ball.'
'Mary said slowly that John threw a ball.'

c. Yukkurito *shika* Mary ga [John ga booru o nage-*nak*-atta
slowly only Mary Nom John Nom ball Acc throw-Neg-Past
to] itta.
Comp said
'Mary said that John only slowly threw a ball.'

(38) a. Mary ga [John ga kyuuni naki-dashita to] itta.
Mary Nom John Nom suddenly cry-began Comp said
'Mary said that John suddenly started crying.'

b. Kyuuni Mary ga [John ga naki-dashita to] itta.
suddenly Mary Nom John Nom cry-began Comp said
'*Mary said that John suddenly started crying.'
'Mary suddenly said that John started crying.'

c. Kyuuni *shika* Mary ga [John ga naki-dasa-*nak*-atta to]
suddenly only Mary Nom John Nom cry-begin-Neg-Past Comp
itta.
said
'Mary said that John only suddenly started crying.'

(39) a. Mary ga [John ga nikai Susan ni kisu-shita to] itta.
Mary Nom John Nom twice Susan Dat kiss-did Comp said
'Mary said that John kissed Susan twice.'

b. Nikai Mary ga [John ga Susan ni kisu-shita to] itta.
twice Mary Nom John Nom Susan Dat kiss-did Comp said
'*Mary said that John kissed Susan twice.' '
'Mary said twice that John kissed Susan.'

c. Nikai *shika* Mary ga [John ga Susan ni kisu-shi-*nak*-atta
twice only Mary Nom John Nom Susan Dat kiss-do-Neg-Past
to] itta.
Comp said
'Mary said that John kissed Susan only twice.'

3. Revisit of c-structure of Japanese causatives

As argued by Kitagawa (1986), Manning et al. (1999), among others, a sequence like *kak-ase* 'write-Caus' is a phonological unit. It then should be noted that it behaves just like a single word as far as phonology is concerned.

But the discussion in Section 2 has suggested that such a sequence would be inconsistent if it is indeed a single word. In this last section I would like to argue that the c-structure of Japanese causatives is actually 'bi-clausal', not mono-clausal as has been traditionally assumed in LFG (cf. Ishikawa 1985, Matumoto 1996). What I would like to pursue here is that the caustive morpheme -*sase* can appear as an independent word. (cf. Kuroda 1981, 1990) The evidence for it comes from separability of the predicate itself.

3.1. Separability

As Kuroda (1981, 1990) argues negative morpheme *na*- and certain focus particles such as *wa*- , *mo*- and *sae*- can intervene between a base stem and a bare causative morpheme -*sase*. Further, Kuno (1983) discuss the separability of the base verb and the causative verb in a honorific form *o-V ni nar* as follows.

(40) Karera ga ooji ni sono kutushita o *o*-haki *ni* sae / mademo
they Nom prince Dat the socks Acc Hon-put.on even
*na-r*ase-ta.
become-Caus-Past
'They made the prince even put on the sox.'

(40) indicates that one of the honorific forms '*o-V-ni-nar*' is not one word formed at lexicon, but formed at syntax; hence bi-clausal at

c-structure in our context. Note that the subjecthood is not exhibited any longer by the honorific marking test, as we have observed in Section 2.1.1. Thus, f-structure of (40) does not need to be bi-clausal. Consider the following contrasts between (41a) and (41b), (42a) and (42b).

(41) a. Boku wa Jon ni pan o tabe-*wa*-sase-ta ga ...
I Top John Dat bread Acc eat Foc Caus-Past but
'I did make John eat the bread, but ... '

b. John wa pan o tabe *wa* shi-ta ga ...
John Top bread Acc eat do-past but
'John did eat the bread, but ... '

(42) a. John wa Biru ni tabako o suwa-*naku*-sase-ta.
John Top Bill Dat tobacco Acc smoke-Neg Caus?-Past
'John made Bill not be able to smoke cigarettes.'

b. *Biru ga tobacco o suwa-*naku* sita.
'Bill Nom tobacco Acc smoke-Neg did'
'Bill didn't smoke cigarettes.'

There might be two possible accounts for (41) and (42). One, which I believe correct, takes *sase* as an independent verb (cf. Kuroda 1981, 1990). If one assumes that (41a) is derived from (41b), then s/he can equally expect (42a) to be derived from (42b). However, this is not the case as (42b) shows its ungrammaticality. This indicates that the c-structure of the causative in (42b) has a bi- clausal structure. The other is based on the assumption that *sase* in (41a) or (42a) is not an independent causative verb, but the form that results from adding *sase* to the verb stem s-'do' + (s)ase _ sase.[12] However, as Kuroda notes, few lexicalists taking this view give convincing arguments

[12] See Miyagawa (1989), Manning, et al. (1999), among others. See also Kato (1985).

against the former Kuroda's syntactic account (ex. (41) and (42), (45)).

3.1.1. Particle intervention and *do*-support in Japanese:

'Do' in the constructions like (43a) is different from do in the light verb constructions in that its existence is dependent on focus particles (cf. Poser (1992), Sells (1998)).

(43) a. tabe-*wa*-shi-ta *tabe-shi-ta
eat-Top-do-Past eat-do-Past

b. shokuji-*wa/o*-shi-ta shokuji-shi-ta
eating-Top / Acc-do-Past eating-do-Past

We here assume that the scope-bearing particles like *wa*, *mo*, *sae* can attach to the fronted VP, not IP (or TP). (cf. Nishiyama & Cho 1998)

(44) a. John ga compuuta o kai-wa / mo / sae shita.
John Nom computer Acc buy -at.least / also / even did
'John at least / also / even brought a computer.'

b. *John ga compuuta o kat-ta-wa / mo / sae shita.
John Nom computer Acc buy-Past-at.least / also / even did
'John at least / also / even brought a computer.'

Thus, it is reasonable to assume that a causative verb *sase* takes a VP complement and the focus particles usually attach to the VP.

(45) John ga Bill ni compuuta o kawase-wa / mo / sae sase-ta.
John Nom Bill Dat computer Acc buy-at least / also / even Caus-Past
'John made Bill at least / also / even buy a computer.'

There are clear syntactic differences between auxiliaries and verbs if one follows the assumption that *do*-insertion salvages from affix hopping (cf. Lasnik 1981, Chomsky 1995). The dummy verb *su* 'do' appears in a tensed clause, i.e., IP (or TP) for supporting the tense marker *ta* 'Past'. In (45), where *sase* clearly takes a VP complement, therefore, it is not necessary for the dummy verb *su* 'do' to get inserted within the VP.

4. Summary

Examining Japanese morphological causatives in detail, I have argued that they are certainly formed at syntax, not at lexicon, through the complex-predicate formation in the sense of Alsina (1996); i.e., characteristically mono-clausal at f-structure. I have also argued that the c-structure is not mono-clausal, but bi-clausal, contrary to Matsumoto and Ishikawa. The present analysis can correctly capture the facts involving morphological causatives in Japanese.

References

Alsina, Alex. 1996. *The Role of Argument Strcture in Grammar*. CSLI Publications.

Andrews, Avery D. and Christopher D. Manning. 1999. *Complex Predicates and Information Spreading in LFG*. CSLI Publications.

Bresnan, Joan and Sam A. Mchombo. 1995. "The Lexical Integrity Principle: Evidence from Bantu," *NLLT* 13, 181-254.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. CSLI Publications.

Gunji, Takao 1999. "On Lexicalist Treatments of Japanese Causatives," in Levine, R. and Green, G. eds (1999), 119-160.

Iida, Masayo 1996. *Context and Binding Japanese*. CSLI Publications.

Ishikawa, Akira. 1985. *Complex Predicates and Lexical Operators in Japanese*. PhD. diss. Stanford University.

Kuno, Susumu 1983. *Shin nihon bunpoo kenkyuu* [The New Study of Japanese Grammar]. Taishukan Publishers.

Kuroda, Shige-Yuki 1990. "Shieki no jodooshi no Jiritusei ni tsuite [On the independence of the causative auxiliary]", In Kunihiro Tetsuya Kyooju Kanreki Taikan Kinen Ronbunshuu Henshuu-iinkai, ed., *Bunpoo to imi no aida* [Between Grammar and Semantics], 93-104. Tokyo: Kuroshio.

Levine, Robert and Green, Georgia (1999) *Studies in Contemporary Phrase Structure Grammar*. Cambridge University Press.

Manning, Christopher D., and Ivan A Sag, and Masayo Iida. 1999. "The Lexical Integrity of Japanese Causatives," in Levine, R. and Green, G. eds. (1999), 39-79.

Matsumoto, Yo. 1996. *Complex Predicates in Japanese*. CSLI Publications and Kuroshio Publishers.

Matsumoto, Yo 1998. "A Reexamination of the Cross-linguistic Parametrization of Causative Predicates: Japanese Perspectives", in *Proceedings of the LFG98 Conference*, ed. M. Butt and T. H. King. CSLI On-line Publications.

Matsumoto, Yo. 2000. "On the Crosslinguistic Parametrization of Causative Predicates: Implication from Japanese and Other Languages", in *Argument Realization*, ed. M. Butt and T. H. King. 135-169. CSLI Publications.

Nishiyama, Kunio and Cho, Eun. 1998. "Predicate cleft constructions in Japanese and Korean: The role of dummy verbs in TP / VP-preposing", *Japanese / Korean Linguistics 7*

Sells, Peter. 1994. "The projection of phrase structure and argument structure in Japanese", ms.

Sugisaki, Koji. 2001. "Scrambling of adjuncts and last resort", *Japanese / Korean Linguistics 10*

LFG and the Analysis of Chinese

Workshop for LFG 2001

Organized by

Adams Bodomo and Luke Kang Kwong

The 6th International Conference on Lexical-Functional Grammar

University of Hong Kong

27th June 2001

Table of contents

1. Introduction by Adams Bodomo

The aim of this workshop is to explore ways in which some aspects of the structure of Chinese may be analyzed in LFG or related constraint-based grammar formalisms. I will first point out some salient features of the Chinese language and raise some possible questions and implications that these might have for LFG. I will then summarize various papers in the workshop, pointing to what issues that are being discussed and what solutions that are proposed.

2. The structure of Chinese

Many works on the Chinese language from Chao (1968), through Li and Thompson (1981) to Huang (1984, 1989, 1991) and beyond have observed that Chinese (including its dialects like Cantonese, Mandarin, Hakka, Minnan, etc) has some very unique structural properties from the perspective of languages like English, French, Italian, German, and Norwegian. Like Italian and Spanish, it is a pro-drop language, but unlike these languages, it is a pro-drop language exhibiting little verbal morphology. Unlike English it is a topic prominent rather than subject prominent language (Li and Thompson 1976). Like Norwegian and other Scandinavian languages it permits long-distance binding (see work by Pan and Hu in this workshop). Cantonese also exhibits considerable complexity in verb complementation (Bodomo and Lee in this workshop), and it has a more flexible word order than English. These and other features show that the structure of Chinese poses some challenges to linguistic description and theory. These properties have attracted the attention of many Chinese linguists working in various grammatical frameworks (e.g. Huang, J. 1984, 1989, 1991, Huang, C. 1989, Huang C. and K. Chen 1989).

3. Issues and questions

Each paper in this workshop takes up one or more aspects of this structure of Chinese. A first issue is that given the situation where Chinese is a pro-drop language and yet does not have much verbal inflection, like many other pro-drop languages, the question as to how grammatical functions can be specified forcefully comes to the fore. In particular, the subject-condition, a universal constraint on sentence structure, comes under threat given such a language type. This is a question for which answers are sought in this workshop.

The phenomena of long distance anaphora demand a restatement of classical principles of binding. This is a relevant issue with respect to the structure of Chinese reflexives. What constraints are needed to predict the correct occurrences of complex reflexives in Chinese? This is another question for which answers are being sought in the workshop. Another important issue that comes up from the short introduction to the structure of Chinese as given above is that the language has a fairly intricate system of verbal complementation. Some of the questions that arise for such a system are what grammatical functions are needed to capture all the nuances of verbal complementation and how this should be done. Beyond these questions of theory and description we need to bring in issues of formal representation and computation. An obvious question would be how powerful the theoretical formalism should be in order to capture the intricacies of the Chinese language.

4. The power of LFG and other constraint-based approaches

Does LFG, along with other constraint-based formalisms, have answers to these issues and questions? LFG, designed as an elaborate linguistic theory but with strong psychological and computational reality, should have the resources to handle these issues of theory and computation with respect to the Chinese language. Two aspects that need brief mention are its commitment to finer-grained functional categorization, and the development of a strong and more powerful constraint-based system in what has come to be known as OT-LFG.

5. The various papers

Each of the four papers in this workshop raises an issue of theoretical importance and then employs, in one way or the other, one or many of the LFG and other constraint-based grammar resources to address the issue.

In the paper titled ***The Subject Condition in Cantonese*** Luke Kang Kwong, Adams Bodomo, and Owen Nancarrow of the University of Hong Kong take up of the issue of grammatical function specification, particularly the theoretically relevant concept of subject condition and seek answers to how one can apply this to Cantonese, a Yue dialect of Chinese as spoken in Hong Kong, which like other dialects of Chinese manifests issues of pro-drop in the absence of overt morphological inflections. The phenomenon of pro-drop is very productive in Cantonese, as shown in the following sentences.

(1) A: *Nei5* *jam2-gan2mat1 je5?*
2.SG drink-ASP what thing
'What are you drinking?'

B: *Jam2-gan2* *seoi2*
drink-ASP water
'(I'm) drinking water.'

After a discussion of various aspect of subjecthood in Cantonese the paper proposes that rather than morphological function specification one needs functional mapping principles as provided in the LFG architecture and pragmatic-discourse criteria to specify subject functions in Chinese and thus salvage the subject condition in this language.

In a paper titled ***An Optimality-Theoretic Account of Mandarin Complex Reflexive*** 'ta ziji' (s/he-self) Pan Haihua and Hu Jianhua of the City University of Hong Kong take up the issue of long distance binding in Mandarin, as shown in (2):

(2) *Johni shuo naben shu hai-le ta-zijii*
say that book hurt-PERF he-self
'Jogn said that that book hurt himself'

Their proposal lies in the exploitation of the relatively new notion of hard and soft constaints to set up a ranking that can account for the facts of complex reflexive binding in Mandarin.

Sun Maosong of Tsinghua University, Beijing in his article ***LFG for Chinese: Issues of Representation and Computation*** assesses the power of the LFG formalism with respect of Chinese computing. According to him, LFG is quite powerful in describing linguistic constructions of Chinese which are of relative sophistication as shown in the Mandarin sentence (3):

(3)	*Zhang-san*	*fang4*	*gou3*	*yao3*	*si3*	*le*	*Li-si*
	person1	send	dog	bite	die	AUX	person2
	N1	V1	N2	V2	V3	AUX	N3
	'*Zhang-san sent the dog to bite Li-si, and Li-si died.*'						

In the author's opinion however, LFG, as a computational formalism, is still not strong enough for computing Chinese. This evaluation of the LFG formalism should raise interesting points for discussion during the workshop.

In the last paper titled ***On the Function COMP in Cantonese*** Adams Bodomo and Sophia Lee of the University of Hong Kong examine phenomena of verbal complementation in Cantonese. In particular, the work focuses on the grammatical function, COMP, a controversial member of the taxonomy of functions in the framework of LFG. There is considerable amount of controversy as to whether clausal complements of the type in (4) should be specified as holding the OBJ or COMP function.

(4)	*Ngo5*	*zi1*	[*keoi5*	*hai6*	*hok6saang1*]
	1.SG	knows	3.SG	be	student
	'I knows that s/he is a student.'				

Indeed questions are asked about the need for COMP as a grammatical function. The paper addresses these issues and proposes that COMPs exist in Cantonese and that sentential complements can function as OBJs and COMPs in the language, and one should thus consider Cantonese as a *mixed language* (Dalrymple and Loedrup 2000).

6. Conclusion:

It is hoped that these approaches to Chinese grammar taken up here open up new ways and directions for pursuing the study of Chinese grammar.

References

Bodomo, A. B. and Y. M. Lee. 2001. On the function COMP in Cantonese. This workshop

Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. California: University of California Press.

Dalrymple, M. and H. Lødrup. 2000. The Grammatical Functions of Complement Clauses. In Butt and King (ed.), Proceedings of the LFG00 Conference, University of California, Berkeley: CSLI Publications. http://www-csli.stanford.edu/publications/

Huang, C. 1989. A Brief Introduction to the Lexical-Functional Grammar. (In Chinese). *Proceedings of ROCLING II & R.O.C. Computational Linguistics Workshops II (1989)*, pp.73-94.

Huang, C. and K. Chen. (eds.) 1989. *Proceedings of ROCLING II & R.O.C. Computational Linguistics Workshops II (1989)*. Taipei: Academia Sinica.

Huang, J. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 531-574.

Huang, J. 1989. Pro-drop in Chinese: a generalized control theory. In Jaeggli, O. and Safir, K. (eds.), *The Null Subject Parameter*.

Huang, J. 1991. Remarks on the status of the null object. In Robert Freidin (ed.), *Principles and Parameters in Comparative Grammar.* Cambridge: MIT Press. pp56-76.

Li, C. and S. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. California: University of California Press.

Pan, H. and Jianhua Hu. 2001. An Optimality-Theoretic Account of Mandarin Complex Reflexive 'ta-ziji' (s/he-self), Workshop on LFG and the Analysis of Chinese, LFG2001.

The Subject Condition in Cantonese

K. K. Luke, A. Bodomo and O. Nancarrow
Department of Linguistics, University of Hong Kong

Proceedings of LFG01 Conference
University of Hong Kong, Hong Kong
Mirriam Butt and Tracy Holloway King (Editors)
2001
CSLI Publication
http://csli-publications.stanford.edu/

Abstract

This paper discusses the subject condition with data from Cantonese. We show that it is not possible to identify subjects in this language based on morphosyntactic criteria alone, and in order to maintain the subject condition in Cantonese, a pro-drop language, one has to determine other ways of identifying subjects. We propose some ways in which subjects can be identified using the lexical mapping theory, augmented by pragmatic discourse criteria.

1. Introduction

The grammatical notion of subject occupies a key position in most linguistic theories. In configurational approaches, the concept, subject, is defined either as [Spec IP] (or some other functional projections) (Haegeman, 1994) or generated in adjunction to VP (Stowell 1981, Chomsky 1986 and Manzini 1983). In LFG, subjects and other grammatical functions are regarded as primitives and represented in functional structures. The subject condition stipulates a default subject for every clause predicated by a finite verb. There are several statements of this condition in the literature. In Bresnan (2001:311) it is stated thus: "Every predicator must have a subject." The status of the subject condition as a grammatical universal has been a matter of some controversy (see, for example, Alsina 1996, Bresnan and Zaenen 1990, Bresnan and Kanerva 1989, and Berman 1999). In this paper, we examine the subject condition with data from Cantonese. Cantonese, like all the other Chinese dialects, is a pro-drop language. However, unlike other pro-drop languages such as Italian, implicit subjects cannot be retrieved through verbal morphology, as will be shown in the next section.

The issue then is how to identify subjects in sentences where they do not have overt expressions. In this paper, we put forward a set of criteria with which some kinds of implicit subjects could legitimately be recovered. We propose that subjects should still be represented at f-structure in consonance with LFG approaches but that instead of achieving functional specification solely at the level of morphosyntax, subjects in Chinese should also be identifiable at a pragmatic-discourse level. This proposal is in line with Bresnan's (2001:98) characterization of the subject as having "…the unique property of being both an argument function and a (grammaticalized) discourse function."

The paper will be structured as follows. In section 2, we introduce the subject condition and discuss on issues surrounding the topic in the literature. In section 3, we focus on the structure of Cantonese, especially its status as a pro-drop language and take up issues of functional specification. Section 4 gives a further focus on Cantonese data. We first outline different types of sentences showing differing subject occurrences in Cantonese and go on to illustrate how to retrieve subjects from a text using discourse-pragmatic information. In section 5, we sketch a formal (LFG) analysis in which the subject condition can be maintained in Cantonese by functional mapping principles. Section 6 is a brief outline showing that discourse-pragmatic criteria are necessary to identify the full range of subject features such as person and number in Cantonese. These mapping principles and discourse pragmatic criteria for subject identification are seen as measures towards satisfying the subject condition in Cantonese and, by extension, other Chinese dialects.

2. The subject condition

The subject condition is a constraint on sentences, and it has been stated in many ways in the literature. In most grammatical frameworks there is an asymmetry between subjects which can have only one member and other grammatical categories like objects which can have more than one members. Simply put the constraint requires that every predicator have a subject (Bresnan 2001). This condition is more elaborately stated in Alsina (1996: 20):

(1) Subject condition:
An f-structure with propositional content must include a subject (as one of its grammatical functions) and no f-structure may include more than one subject.

A major issue with the subject condition is its acceptance as a universal condition in LFG. Obviously some languages do not have an overt c-structure subject even with sentences headed by finite verbs. Such so called pro-drop languages like Italian and Spanish in many instances allow the referential properties of the covert subject to be retrieved by features such as the verbal inflection as shown below for Italian and Spanish:

Italian:
(2) a. *pro ho telefonato*
have.1.SG telephoned
'I have telephoned'

b. *Gianni ha parlato*
Gianni has.3.SG spoken
'Gianni has spoken.'

c. *pro ha parlato*
has.3.SG spoken
'He has spoken.'

Spanish:

(3) a. *(yo) como comida*
1.SG eat.1SG food
'I eat food.'

b. *(nosotros) comemos comida*
1.PL eat.1.PL food
'We eat food.'

However there are languages that allow covert c-structure subjects whose referential properties can hardly be inferred from verbal inflection or other morphological manifestations. Such a language would pose problems for the universality of the subject condition. Cantonese is such an exemplar. We shall look more closely at the properties of Cantonese as a pro-drop language and the consequences such a structure has for functional specification and the applicability of the subject condition in this language.

3. Pro-drop in Cantonese and Functional Specification

In this section we review the concept of pro-drop and illustrate it with Cantonese sentences in section 3.1. In section 3.2., we discuss the notion of functional specification in LFG and use it to show how subjects can be retrieved from sentences with covert subjects.

3.1. Cantonese as a pro-drop language

Cantonese, like all Chinese dialects, has some very unique structural properties from the perspective of languages like English, French, Italian, German, and Norwegian. Like Italian and Spanish, it is a pro-drop language, but unlike these languages, it is a pro-drop language exhibiting little verbal morphology. Unlike English it is a topic prominent rather than subject prominent language (Li and Thompson 1976). Like Norwegian and other Scandinavian languages it permits long-distance binding (Pan and Hu in 2001 workshop). Cantonese also exhibits considerable complexity in verb complementation (Bodomo and Lee in 2001 workshop), and it has a more flexible word order than English. These and other features show that Cantonese and other dialects of Chinese pose some challenges to linguistic description and theory. These properties have attracted the attention of many Chinese linguists working in various grammatical frameworks (Huang 1984, 1989, 1991, Hu and Pan 2000) While all these are interesting in themselves our focus here will be on subjecthood conditions of Cantonese.

The phenomenon of pro-drop is very productive in Cantonese, as shown in the following sentences.

(4) A: *nei5 jam2-gan2 mat1 je5?*
2.SG drink-ASP what thing
'What are you drinking?'
B: *jam2-gan2 seoi2*
drink-ASP water
'(I'm) drinking water.'

(5) (Talking about dogs)
wui5-m4-wui5 beng6 gaa3
Will-not-will ill PART
'Would (they = the dogs) get ill?

(6) *ji1gaa1 lok6-gan2 jyu5*
now fall-ASP rain
'(It's) raining now.'

(7) A: *teng1gong2 seoi6si2 hou2 leng3 wo3*
Hear Switzerland very pretty PART
'(I) hear that Switzerland is very pretty'
B: *tai2 fung1ging2 lo1 hai6*
See scenery PART only
'(It's) only (good for) sight-seeing'

In (4)-(7), it is not possible to determine from the verb forms the identity of the implicit subjects. In (4), the subject pronoun can be retrieved from the immediate speech context. In (5), the subject is understood to be the current topic of the conversation. (6), like all other meteorological sentences in the language, does not come with an expletive subject. Sentence (7) can be understood as either containing a zero-subject of generic reference (like 'on' in French) or having 'sight-seeing' as subject. There are also topic-comment sentences where the initial NP is not an argument of the main verb but is nevertheless related to it pragmatically or through the discourse context.

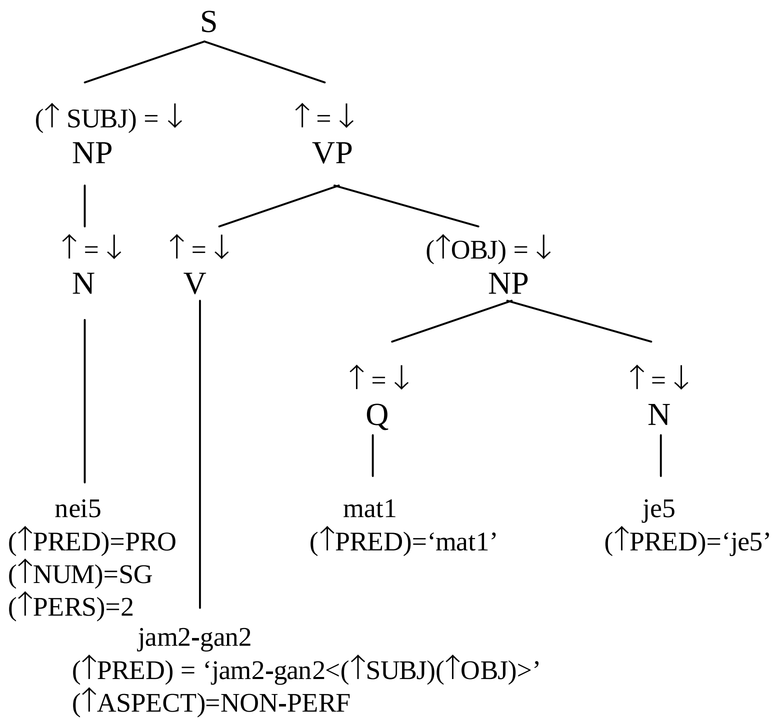
3.2. Functional specification

These structural features of Cantonese in which there is little verbal morphology and yet there is the possibility of pronoun drop involving various grammatical functions introduce an issue of function specification. Function specification is quite an important issue in LFG. In the framework, two main types of function specification are recognized (Berman 1999), structural and morphological function specification, where grammatical functions are defined or specified in terms of the structural positions in which they occur in the sentence in the former; and where grammatical functions are specified by the help of verbal and nominal morphology such as case, and other types of feature specifications in the latter.

Languages vary with respect to the choice of the two. In this paper we observe that the two are not enough with respect to Cantonese and thus propose a third kind of functional specification: discourse-pragmatic function specification, where grammatical functions can be retrieved from the discourse-pragmatic context.

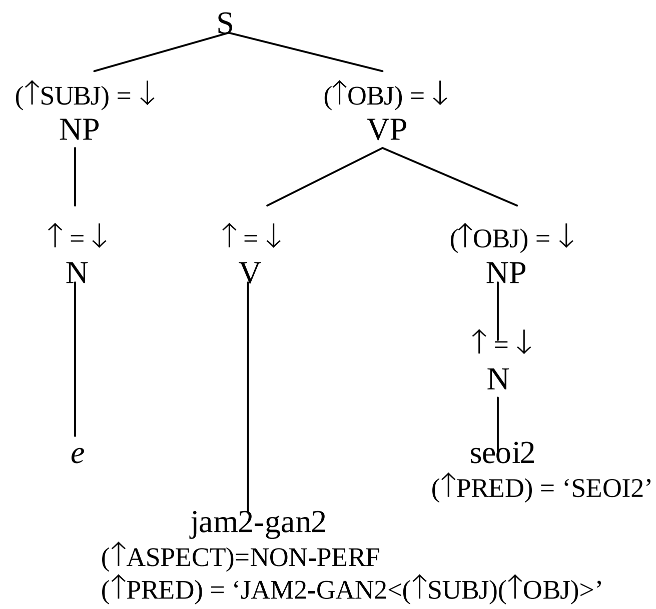
We illustrate function specification and other aspects of the structure of Cantonese by the following c- and f-structure diagrams of the two sentences in (8):

(8) C- and f-structures of *nei5 jam2-gan2 mat1 je5*



$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{'PRO'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 2 \end{array}\right] \\ \text{ASPECT} & \text{NON-PERF} \\ \text{PRED} & \text{'jam2-gan2} <(\uparrow\text{SUBJ})(\uparrow\text{OBJ})>\text{'} \\ \text{OBJ} & \left[\begin{array}{ll} \text{Q} & \left[\begin{array}{ll} \text{PRED} & \text{MAT1} \end{array}\right] \\ \text{PRED} & \text{'je5'} \end{array}\right] \end{array}\right]$$

(9) C- and f-structures of *jam2-gan2 seoi2*



$$\left[\begin{array}{ll}
\text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{'PRO'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 1 \end{array}\right] \\
\text{ASPECT} & \text{NON-PERF} \\
\text{PRED} & \text{'JAM2-GAN2<(}\uparrow\text{SUBJ)(}\uparrow\text{OBJ)>'} \\
\text{OBJ} & \left[\begin{array}{ll} \text{PRED} & \text{'SEOI2'} \end{array}\right]
\end{array}\right]$$

From these diagrams, we notice that Cantonese, being a language with scant inflectional morphology, belongs to the type of languages that opts for structural functional specification. In a simple declarative sentence the subject occupies a preverbal position while an object (for those predicates that subcategorise for it) occur postverbally. This is shown in the c-structure in (8). The c-structure in (9) contains a phonologically unexpressed subject whose referential and agreement features cannot be retrieved morphologically and syntactically. That is where discourse-pragmatic functional specification comes in. From the context we know it is the individual who is being asked the question in (4) that points to the subject of the sentence. Hence functional specification, including number and person features as shown in the c-and f-structures in (9) are obtained from the discourse situation.

We will return to the issue of function specification from an LFG perspective, but for now we will look at subjecthood and how to describe it in more detail in the next section.

4. More on subjecthood in Cantonese

In this section, first, a list of four types of sentences showing differing subject expressions in 4.1. is given. An illustration of how to retrieve subjects from a text using discourse-pragmatic information is then provided in 4.2.

4.1. Types of subjects in Cantonese sentences

As in Mandarin Chinese and other Chinese dialects, subjects cannot usually be readily identified in Cantonese. The main kinds of cases are:

(a) Sentences which clearly and explicitly have a subject.

(10) *ngo5 lou5gung1 ci3ci3 faan1lei4 dou1 haak1 saai3 gam2*
my husband every-time come-back also tanned like-this
'My husband is tanned every time he comes back.'

(11) *nei5 jau6 m4 hai6 sau3ji1*
you also not be vet
'You are not a vet.'

(b) Sentences which clearly do not have subjects

(12) *zan1 gaa4*
true FP
'Is that true?'

(13) *mat1 je5 waa2*
what
'What?'

(c) Sentences which have an initial constituent which is either an NP which is not an argument of the verb, or a verb or adjective (and therefore cannot possibly be an argument)

(14) *zik1hai6 keoi5 mat1 dou1 laat6 gaa3*
that-is it what all hot FP
'Everything was hot there.'

(15) *siu1 m4 sai2 cin2*
roast not need money
'The roasting was free.'

(d) Sentences in which it is not clear what the subject is

(16) *di1 hang4lei5 baai2zo2 hai2 go3 gaa2 soeng6bin1*
PL luggage place at M shelf above
'The luggage was on the shelf.'

(17) *di1 min6baau1 sik6dou3 ngo5 hou2 baau2*
PL bread eat-to me very full
'The bread made me full.'

The most difficult cases are those where the initial NP of a sentence is the patient of the verb, but the form of the verb is clearly not passive.

While he uses the term 'subject' in his works, Chao (1968) implies that it may not have any significance other than as a convenient term. In talking about topic-comment being the grammatical meaning of subject-predicate in Chinese, he stresses that the relationship between subject and predicate can be quite variable. A sentence is fine "so long as there is some relationship of topic and comment between subject and predicate." He goes on to say:

"For example, in *zheijian shi zao fabiao le*, 'This matter has long been published', we are translating *fabiao* by passive verb form 'has been published', but in the Chinese there is no marker for received action (*bei* 'by', '-ed' would not be appropriate here), and a closer structural translation would be: '(As for) this matter, (one) has long published (it),' Again in: *Zhe gua chizhe hen tian*. 'This melon eating very sweet, --- taste very sweet' seems to be an active verb used passively, but a nearer rendering of the structure of the sentence is: 'This melon, (when one is) eating (it), is very sweet.' All such renderings in English, however, are limited by the exigencies of English grammar requiring a clear actor-action relation, at least in the grammatical sense, thus entailing a number of parenthetical devices which never were in the Chinese, which simply said: 'This matter has long been ago published'. 'This melon eating very sweet'." (1968:70)

Chao (1968) illustrates the point further with his famous example:

Ni	*jiu*	*xie*	*ta*	*touche*	*de*	*shiqing*
You	then	write	his	stole-car	GE	matter

The sentence is ambiguous between:

(1) You just write about the accident of his/her stealing the car.

(2) You just write about the accident of his/her bicycle being stolen.

In English, not every sentence has a c-structure subject either. An examination of English text suggests that about 1 or 2 in every 10 sentences in ordinary conversations do not have a subject. For some sentence types, it is usually possible to retrieve missing subjects. For other sentence types there are simply no subjects. On balance, subject is a useful notion in English.

4.2. Illustrating discourse-pragmatic phenomena for Cantonese subjecthood

In this section, an illustration of how subjects and topics are retrieved from the discourse-pragmatic context is given by going through a short extract of a conversation in some detail.

The passage is taken from a recording of a naturally occurring conversation. The participants C and M are talking about a trip which M has recently made to her home village in Guangdong, for the special purpose of eating lychees fresh from the trees. M's home village is famous for its lychees.

We go through the passage sentence by sentence. For each sentence we provide four lines of information, as follows:

Line 1: romanisation of the sentence
Line 2: word for word gloss
Line 3: literal translation (staying close to the syntax of Cantonese sentence)
Line 4: free translation

Each sentence is numbered along the right margin.

C: *do1-m4-do1 lai6zi1 sik6 aa3?* (1)
lots_of-not-lots_of lychee eat SP
Lots-or not lychees eat?
'Were there lots of lychees to eat?'

M: *waa1, do1dou3 fei1hei2.* (2)
wow, lots_until fly_up
wow, such a lot that fly (i.e. such a lot that one can fly)
'Wow, there was an awful lot of lychee to eat that one can't possibly imagine how much'

sik6dou3 ngo5 baau2 aa3 zan1hai6 (3)
eat_until I full SP really
eat until I full, really
'I ate until I was full, actually'

waa1, faan1dou3heoi3 aa3, (4a)
wow return_arrive PP
wow return
'Wow, when I got to the village,'

nei5 aa3-saam1-suk1gung1 go2dou6 sik6 go2di1 lai6zi1 aa3, (4b)
you NamePfx-three-grand_uncle there eat those lychee PP
the lychee eat at your third grand uncle's place,
'the lychees that we ate at your third grand uncle's place,'

go2-po1 Gwai3mei2 gam3 daai6po1 (4c)
that-CL Gwaimei so big-CL
that Gwaimei (the name of a kind of lychee) tree so big
'the Gwaimei tree was so tall.'

Sei3-go3 yan4 sik6saai3 jat1-po1 Gwai3mei2. (5)
four-CL people eat-all one-CL Gwaimei
Four people eat all of one Gwaimei
'The four of us ate a whole Gwaimei tree.'

C: *hai6 aa4?* (6)
yes SP
yes?
'Really?'

M: *sai1-m4-sai1lei6 aa3?* (7)
great-not-great SP
Great or not?
'Isn't that something?'

C: *dim2gaai2 m4 daai3 di1 faan1lai4 aa3?* (8)
why not bring some back SP
why not bring some back?
'Why haven't you brought some lychees back?'

M: *daai3-zo2 faan1lai4 laa1* (9)
bring-PERF return SP
brought back
'I have brought some back'

sik6-dak1-saai3 me1? (10)
eat-can-all SP
Can eat all?
'You think one can eat it all up?'

nei5 sik5-dak1 hou2 do1 me1? (11)
you eat-can very lots SP
You can eat a lot, you think?
'You think you can eat a lot?'

nei5 jau6 heoi3-zo2 Jat6bun2. (12)
you also go-PERF Japan
you also gone to Japan
'And you had gone to Japan'.

C: *o5* (13)
oh
I see
'I see.'

M: *zing6faan1 hou2 do1* (14)
remain very much
left lots
'Lots of lychees were left.'

hau6mei1 bei2-saai3 jan4 sik6 zi1maa3 (15)
afterwards give-all people eat SP
Afterwards just given all to others eat
'We gave them all to others to eat afterwards/ They were all given to others to eat afterwards'.

C: *jau6 m4 lau4-faan1 gei2 lap1 bei2 ngo5 si3haa2, gam3 daai6-lap1* (16)
and not save a-few CL give I/me try so big-CL
And didn't save a few for me try, so big
'And you didn't save a few pieces for me to try – such big lychees'.

M: *sik6-saai3 laa3!* (17)
eat-all SP
All eat!
'All have been eaten/ We have eaten them all'

go2-di1 hai6 aa3-Jing1 go3 lou5gung1 maai5 faan1lai4 bei2 ngo5dei6 zi1maa3 (18)
that-PL be NamePfx-Jing CL husband buy return give we/us SP
Those are Jing's husband buy back give to us only
'Those were bought by Jing's husband for us.'

Our analysis of each of these sentences is presented in the following table:

1	*do1-m4-do1 lai6zi1 sik6 aa3?* lots_of-not-lots_of lychee eat SP Lots-or not lychees eat? 'Were there lots of lychees to eat?'	The eating of the lychees is understood to have been done by M by reference to the discourse context
2	*waa1, do1dou3 fei1hei2.* wow, lots_until fly_up wow, such a lot that fly 'Wow, there wasasuch a lot of lychees to eat that one can't possibly imagine how much.'	(a) 'Such a lot' is understood to be referring to lychees by virtue of the topic of the conversation up to this point. (b) The subject of 'fly' is understood to be anyone, i.e. generic reference
3	*sik6dou3 ngo5 baau2 aa3 zan1hai6* eat_until I full SP really eat until I full, really 'I ate (the lychees) until I was full/ The lychees were such that I fed on them and was full'	The topic is lychee. It is clearly what was eaten. It may be subject or object of the sentence.

4	*waa1, faan1dou3heoi3 aa3,* wow return_arrive PP wow return Wow, when I got to the village,’	The ‘returning’ is understood to be done by the speaker (and possibly her relatives) and the destination is understood to be her home village from the discourse context.
5	*sei3-go3 yan4 sik6saai3 jat1-po1* four-CL people eat-all one-CL *Gwai3mei2.* Gwaimei Four people eat all of one Gwaimei ‘The four of us ate a whole Gwaimei tree.’	The subject is explicitly ‘the four of us’.
6	*hai6 aa4?* yes SP yes? ‘Really?’	No subject. A ‘non-propositional sentence’?
7	*sai1-m4-sai1lei6 aa3?* great-not-great SP Great or not? ‘Isn’t that something?’ [Subject: nil; Topic: the fact that 4 people finished eating a whole tree of lychees]	In “isn’t x something?’, x is understood from the speaker’s previous turn to mean ‘the fact that the four of us finished eating a whole tree of lychees’.
8	*dim2gaai2 m4 daai3 di1 faan1lai4* Why not bring some back *aa3?* SP Why not bring some back? ‘Why haven’t you brought some lychees back?’	The subject is understood to be ‘you’ from the discourse context; the object is understood to be ‘lychees’, which is the current topic.
9	*daai3-zo2 faan1lai4 laa1* bring-PERF return SP brought back ‘I have brought some back’	The subject is understood to be ‘I’; the object ‘lychee’
10	*si6k-dak1-saai3 me1?* eat-can-all SP Can eat all? ‘You think one can eat it all up?’	The one who eats lychees is understood to be ‘you’, as this is a rhetorical question. The object is lychees, the topic.
11	*nei5 sik6-dak1 hou2 do1 me1?* you eat-can very lots SP You can eat a lot, you think? ‘You think you can eat a lot?’ [Subject: you; Topic: lychees]	(same as 10)
12	*nei5 jau6 heoi3-zo2 Jat6bun2.* you also go-PERF Japan you also gone to Japan ‘And you had gone to Japan’.	The subject is explicitly ‘you’, the object ‘Japan’.

13	*O5* Oh Oh 'I see.'	No subject. 'non-propositional sentence'?
14	*zing6faan1 hou2 do1* remain very much left lots 'Lots of lychees were left.'	what is left is understood to be lychees, the topic.
15	*hau6mei1 bei2-saai3 jan4 sik6* Afterwards give-all people eat *zi1maa3* SP Afterwards just given all to others eat 'We gave them all to others to eat afterwards/ They were all given to others to eat afterwards'	The 'giver' is understood to be the speaker, the object of give 'lychees', and the recipient is explicitly 'others'.
16	*jau6 m4 lau4-faan1 gei2 lap1 bei2* and not save-FAAN a-few CL give *ngo5 si3haa2, gam3 daai6-lap1* I/me try so big-CL And didn't save a few for me try, so big 'And you didn't save a few pieces for me to try – such big lychees'.	The initiator of 'save' is understood to be M, by reference to the discourse context. The object is lychee. The recipient is explicitly 'me'.
17	*sik6-saai3 laa3!* eat-all SP All eat! 'All have been eaten/ We have eaten them all'	A classical sentence which actually turns up in this conversation. There is no explicit subject or object. The one(s) who did the eating is probably the speaker (M), or a group including her (or some other people). What was eaten is clearly the lychees. The sentence could be an abbreviated form of either 'Lychee eat all (i.e. all eaten)' or 'I/We ate all the lychees'.
18	*go2-di1 hai6 aa3-Jing1 go3* that-PL be NamePfx-Jing CL *lou5gung1 maai5 faan1lai4 bei2* husband buy return give *ngo5dei6 zi1maa3* we/us SP Those are Jing's husband buy back give to us only 'Those were bought by Jing's husband for us.'	With the verb 'be', the subject is explicitly 'those'.

There is nothing unusual or peculiar in terms of syntactic structure about the sentences found in this passage. Most Cantonese conversations are like this. It can be seen from the above that subjects and objects are more often than not left implicit in the Cantonese sentence, only to be recovered from the discourse-pragmatic context.

5. Analysis and formalization: From A-Stucture to Syntactic Functions

It this section we return to an LFG analysis and attempt to show how the subject condition can be maintained in Chinese, given the complex cases of possibly subjectless sentences in Cantonese that have been outlined in the previous section. It is clear from the above data on Cantonese subjecthood that there are sentences in which there is no overt c-structure subject. However, though the subject condition as stated by Alsina (1996) is a condition on f-structure, we want to claim that it is possible to pursue a structural analysis of function specification. This is an important issue in discussions about the falsifiability of syntactic principles, as of all other scientific principles. The analysis requires us i. to posit an empty category *pro* and then ii. use functional mapping principles.

Consider the sentence in (4) repeated below as (18) for convenience:

(18)	*e*	*jam2-gan2*	*seoi2*
		drink-ASP	water
	'(I'm) drinking water.'		

For those theories that define functions configurationally we would say the empty category, *e*, is a subject position since it is [Spec IP], following the External Projection Principle (the requirement that all sentences have subjects), an extension of the Projection Principle (the requirement that argument structure or lexical properties of words be projected in syntax).

The second approach, which is compatible with functional specification within LFG is to rely on the idea of using mapping principles from argument structure to functional structure as a way of satisfying the subcategorization requirements of a predicate. As Alsina (1996:45) indicates: "Any theory has to guarantee that the subcategorization requirements of a predicate are satisfied, namely, that the syntactic structure include all grammatical functions required by the predicate and no spurious ones." To accomplish this various types of mapping principles have been developed.

We adopt those proposed in Bresnan (2001:311) to specify the functions of the Cantonese pro-drop construction in (4/18).

Bresnan (2001) proposes the following mapping principles for the purposes of specifying grammatical functions of the arguments of a predicate:

(19) Mapping Principles

a. Subject roles:

(i) $\hat{\theta}$ is mapped onto SUBJ when initial in the a-structure;

[-*o*]

otherwise

(ii) $\hat{\theta}$ is mapped onto SUBJ

[-*r*]

b. Other roles are mapped onto the lowest compatible function in the partial ordering : SUBJ>OBJ, OBL_θ>OBJ_θ

So in the Cantonese lexicon then, the following lexical entry for the verbal predicate *jam2-gan2* would obtain:

(20) *jam2-gan2* V (↑ASP)=NON-PERF
(↑PRED)='JAM<(↑SUBJ)(↑OBJ)>'

From this lexical entry it is clear that the verbal predicate *jam2* is a transitive verb that subcategories for a subject and an object.

To specify these structures, especially in situations where there is the absence of morphological marking, one must look for correspondences between levels of representation such as a- and f-structure.

An a-structure consists of a predicator with its argument roles, along with an order that represents the relative prominence of the roles and a syntactic classification of each role indicated by a feature (Bresnan 2001). The a-structure of our verbal predicate, *jam2* would thus be as in (21):

(21) *jam2* < x y >
[-*o*] [-*r*]

The features $\pm o$ and $\pm r$ refer to (non)objective and (un)restricted syntactic functions respectively

Relative prominence is encoded by the following Thematic Hierarchy (Bresnan 2001):

(22) Thematic Hierarchy:
Agent>beneficiary>experiencer/goal>instrument>patient/theme>locative

To specify the grammatical functions, including the subject which is our concern here, we need to apply the above mapping rules to the argument structure of *jam2-gan2* the verbal predicate that heads our target sentence:

(23) pro *jam2-gan2 seoi2*

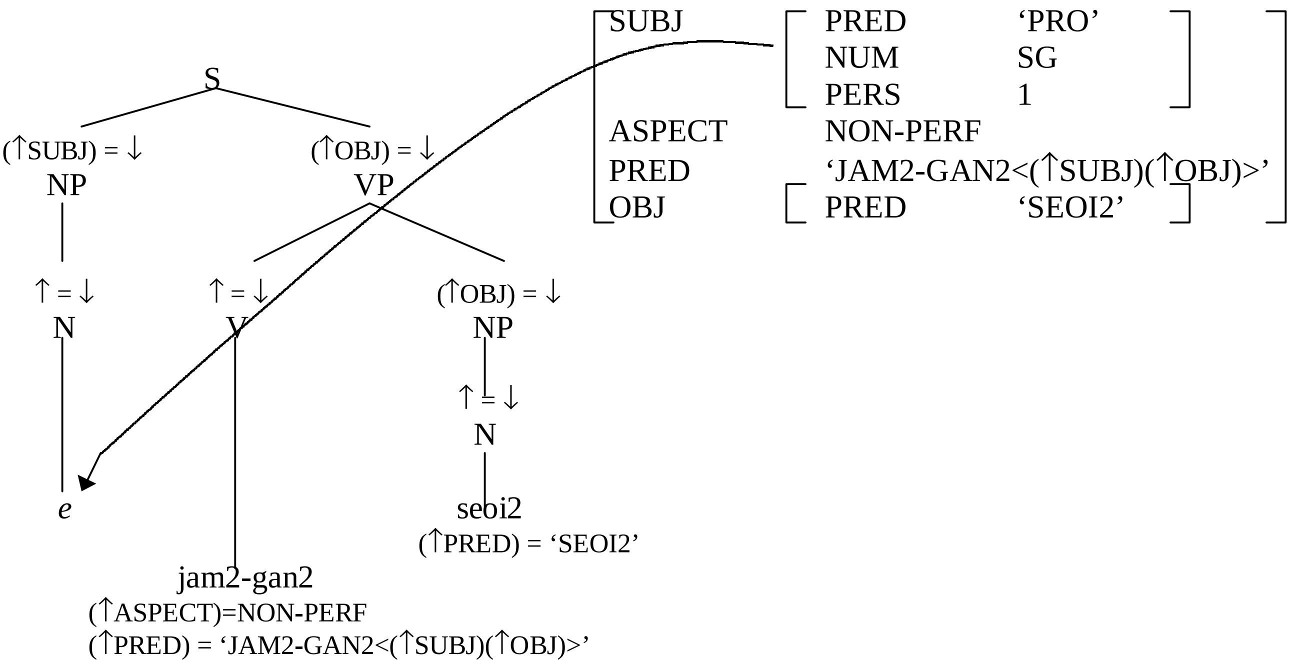
(24)

a-structure	*jam2*	<	x	y	>
			[-o]	[-r]	
			|	|	
f-structure			SUBJ	OBJ	

The most prominent argument $\hat{\theta}$, the agent, is mapped to the SUBJ function since it is initial in the a-structure. By principle (b), the other role is mapped onto OBJ according to the partial ordering SUBJ>OBJ, $OBL_\theta > OBJ_\theta$ in (19). This gives us a complete functional specification for the f-structure of our target sentence.

We have argued above that the subject condition be maintained even in overtly subjectless Cantonese sentences like *jam2 seoi2* if we posit a covert pronoun or a pro-drop scenario and then apply mapping principles, aligning a-structure with f-structure.

This pro will then have the following specification as (25):



With this analysis we are able to defend the subject condition in (1) which we repeat in (26) for convenience:

(26) Subject condition:
An f-structure with propositional content must include a subject (as one of its grammatical functions) and no f-structure may include more than one subject.

Notice from the above that the subject function can be specified and assigned to pro, but that we cannot fully determine the full range of person and number features of the pronoun. For these reasons it is impossible to eliminate the empty category by the economy principle (Bresnan 2001).

To address this and other issues of a fuller function specification we propose to augment the above defence of the subject condition by functional mapping principles with a discourse-pragmatic analysis.

6. Discourse-pragmatics and the subject condition in Cantonese

Our analysis so far is that the subject condition can be maintained in Cantonese by treating subjectless c-structure constructions as having a pro-drop node and applying functional mapping principles to license subjecthood.

Not all the full range of subject features can be captured in such a way. As there are no subject-verb agreement or other morphological evidence in Cantonese, mapping principles alone are not enough in specifying the full range of subject features. We believe that we can appeal to discourse-pragmatic criteria in the case of Cantonese to address the issue.

Indeed Bresnan (2001) indicates that the subject has the unique property of being both an argument function and a grammaticalized discourse function. Let us take a look at the following sentences repeated from (4)-(7) below in (27)-(30) for convenience:

(27) A: *nei5 jam2-gan2 mat1 je5?*
2.SG drink-ASP what thing
'What are you drinking?'
B: *jam2-gan2 seoi2*
drink-ASP water
'(I'm) drinking water.'

(28) (Talking about dogs)
wui5-m4-wui5 beng6 gaa3
Will-not-will ill PART
'Would (they = the dogs) get ill?

(29) *ji1gaa1 lok6-gan2 jyu5*
now fall-ASP rain
'(It's) raining now.'

(30) A: *teng1gong2 seoi6si2 hou2 leng3 wo3*
Hear Switzerland very pretty PART
'(I) hear that Switzerland is very pretty'
B: *tai2 fung1ging2 lo1 hai6*
See scenery PART only
'(It's) only (good for) sight-seeing'

In (27)-(30), it is not possible to determine from the verb forms the identity of the implicit subjects. In (27), the subject pronoun can be retrieved from the immediate speech context.

In this case it is 1st person singular pronoun since only such a pronoun can serve as a response to the previous sentence *nei5*…. In (28), the subject is understood to be the current topic of the conversation, dogs, in which case we are dealing with 3rd person plural pronouns or a full NP in plural form. (29), like all other meteorological sentences in the language, does not come with an expletive subject. Sentence (30) can be understood as either containing a zero-subject of generic reference (like 'on' in French) or having 'sight-seeing' as subject. There are also topic-comment sentences where the initial NP is not an argument of the main verb but is nevertheless related to it pragmatically or through the discourse context.

So from the above we see that we have used discourse-pragmatic criteria such as:

(i) immediate speech context

(ii) current topic of conversation, and

(iii) metereological and other expletive subject situations which need no person and number specification anyway

to identify and fully specify semantic and grammatical features of subjects in subjectless sentences.

7. Conclusion

We have shown in this paper that the subject condition can be maintained in Cantonese, not by morphological function specification, but by functional mapping principles and discourse-pragmatic considerations. Unlike other subject pro-drop languages such as Italian and Spanish or a language like German with a relatively rich inflectional morphology where subject-verb agreement features can specify the subject (Berman 1999), Cantonese is relatively unique in being a pro-drop language with only scant inflectional morphology (Bodomo 2000, Luke 2001, Bodomo and Lee 2001). We propose that to satisfy the subject condition in Cantonese and thus maintain the universality of this constraint, one has to appeal to functional mapping principles and discourse-pragmatic function specification. We have also provided quite a large variety of Cantonese sentences illustrating various types of subjects, and illustrating a differentiation between subjects and topics, and how to handle more complex conversational data in the form of text. We believe that with simple cases of pro-drop sentences in Cantonese, the subject condition can be maintained and accounted for and thus defended as a universal condition on sentence structure.

8. References:

Alsina, A. 1996. *The Role of Argument Structure in Grammar: Evidence from Romance*. California: CSLI Publications.

Berman, J. 1999. Does German satisfy the subject condition? In Butt and King (1999). Proceedings of LFG conference 1999. Stanford University: CSLI Publication.

Bodomo, A. B. 2000. Expressing the benefactive in Dagaare and Cantonese. 22nd West African Languages Congress, 15th -22nd August 2000. University of Ghana, Accra, Ghana.

Bodomo, A. B. and Y. M. Lee. 2001. On the function COMP in Cantonese. Workshop on LFG approaches to Chinese, LFG2001.

Bresnan, J. 2001. *Lexical-Functional Syntax*. Blackwell Publishers.

Bresnan, J. and Kanerva, J. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* **20**:1-50.

Bresnan, J. and Zaenen, A. 1990. Deep Unaccusativity in LFG. In K. Dziwirek, P. Farell, E. Mejias-Bikandi (eds.): *Grammatical Relations: A Cross-Theoretical Perspective*. Stanford: CSLI. pp45-57.

Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. California: University of California Press.

Chomsky, N. 1986. *Barriers*. Cambridge: MIT Press.

Haegeman, L. 1994. *Introduction to Government-Binding Theory* (2nd edition, revised). Oxford: Blackwells.

Huang, J. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* **15**, 531-574.

Huang, J. 1989. Pro-drop in Chinese: a generalized control theory. In Osvaldo Jaeggli and Ken Safir(eds.), *The Null Subject Parameter*. Dordrecht: Kluwer Academic Publishers.

Huang, J. 1991. Remarks on the status of the null object. In Freidin, R. (ed.), *Principles and Parameters in Comparative Grammar*, 56-76. Cambridge: MIT Press.

Hu, Jianhua and Pan Haihua. 2000. Deriving the Subject-Object Asymmetry in Topicalization. Presented at the International Symposium on Topic and Focus, Hong Kong, 21-23 June 2000.

Luke, K. K. 2001. Subjecthood in Cantonese. Departmental seminar. Department of Linguistics, University of Hong Kong.

Manzini, M. R. 1983. On Control and Control Theory. *Linguistic Inquiry* **14**:421-446.

Stowell, T. 1981. The origins of phrase structure. Cambridge, MA: MIT Department of Linguistics and Philosophy. Ph.D. dissertation.

An Optimality-Theoretic Account of Mandarin Complex Reflexive 'ta-ziji' (s/he-self)

Haihua Pan Jianhua Hu

City University of Hong Kong

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

An Optimality-Theoretic Account of Mandarin Complex Reflexive 'ta-ziji' (s/he-self)[1]

Haihua Pan Jianhua Hu

City University of Hong Kong

1. Introduction

As noted in Pan (1995, 1998), non-contrastive complex reflexive *ta-ziji* (s/he-self) in Chinese (i) can have a long-distance (LD) bound antecedent; (ii) allows non-c-command/sub-command antecedents; and (iii) observes some kind of blocking effect. In this paper we will show that the binding properties exhibited by *ta-ziji* can be best explained if we adopt an Optimality-theoretic (OT) account of reflexivization. We think that the blocking effect of *ta-ziji* can be derived from the prominence constraint which stipulates that the binding of a reflexive to α will be blocked by an intervening NP β iff β is not less prominent than α. After listing different constraints that regulate the interpretation of reflexives, we find that it is necessary to make a distinction between hard constraints and soft constraints, and in this aspect, we deviate from the standard OT hypothesis in the sense that all constraints are violable. Under our analysis, hard constraints are inviolable through all languages, and hence, they do not vary from language to language, whereas soft constraints are violable, and their different rankings play an active role in accounting for the different binding properties of reflexives in different languages.

2. The Binding Properties of *Ta-ziji*

Although *ta-ziji* cannot be bound across the local animate subject in (1), it can skip the local inanimate subject in (2), as noted in Pan (1998).

(1) *John$_i$ zhidao Bill xihuan ta-ziji$_i$
know like he-self
John knows that Bill likes himself.

(2) a. John$_i$ shuo naben shu hai-le ta-ziji$_i$
say that book hurt-Perf he-self
John said that that book hurt himself.

b. John$_i$ shuo naben shu fang zai ta-ziji$_i$ de jiali
say that book put at he-self DE home

[1] This research was funded in part by the Strategic Research Grant of City University of Hong Kong (#7001122), and the FHS (Faculty of Humanities and Social Sciences) Research Expense Grant of City University of Hong Kong (#8760051). We acknowledge the generous support of the University. Besides, we thank Joan Bresnan for encouraging us to attempt an OT account of reflexivization in Chinese.

John said that that book was put at his home.

Sentences like (2) show that the Chinese reflexive *ta-ziji* is constrained not by an absolute locality condition like Chomsky's (1981) Binding Condition A, but by a relative one, and thus exhibits properties fundamentally different from English reflexives like *himself*, which must be constrained by an absolute locality condition, as demonstrated below (Pan 1998: 774):

(3) a. *John $_i$ said that book hurt himself $_i$

b. *John $_i$ knew that debt brought himself into bankruptcy $_i$

The following sentences show that *ta-ziji* can also skip an animate subject, though it cannot skip a human subject.

(4) John $_i$ shuo yitiao gou zai ta-ziji $_i$ de fangjian shuijiao
say one dog at he-self DE room sleep
John said that a dog was sleeping in his room.

Although *ta-ziji* must find a compatible NP as its antecedent and can thus skip an incompatible inanimate or animate but nonhuman subject NP, it cannot skip an incompatible human subject NP, as shown below:

(5) *John$_i$ juede wo xihuan ta-ziji$_i$.
think I like he-self
John thinks I like him.

Besides, as shown in (6), *ta-ziji* can also have a sub-commanding antecedent, like the bare reflexive *ziji*, which, according to Tang (1989), can be bound to a sub-commanding antecedent, as shown in (7), with the definition of *sub-command* given in (8).

(6) John$_i$ xie de shu gei ta-ziji$_i$ dailai-le xuduo mafan.
write DE book to he-self bring-Perf many trouble
The book that John wrote brought a lot of troubles to him.

(7) a. John$_i$ de jiao'ao hai-le ziji$_i$.
DE pride hurt-Perf self
John's pride hurt him.

b. John$_i$ xie de shu gei ziji$_i$ dailai-le xuduo mafan.
write DE book to self bring-Perf many trouble
The book that John wrote brought a lot of troubles to him.

(8) β sub-commands α iff β is contained in an NP that c-commands α or that sub-commands α, and any argument containing β is in subject position.

Although *ta-ziji* may be bound by a sub-commanding antecedent, neither sub-command nor c-command is a necessary requirement on its antecedent, as exemplified below:

(9) a. Wo wei John zhaodao-le ta-ziji de zhaopian.
I for find-Perf he-self DE photo
I found John's photo for him.
b. Wo cong John nar zhaodao-le ta-ziji de zhaopian.
I from there find-Perf he-self DE photo
I found John's photos from him.

In (9) the antecedents are all contained in a PP adjunct, which does not c-command the reflexive. Since they are neither in a subject position nor contained in a c-commanding or sub-commanding NP, they do not sub-command the reflexive, either.

3. Prominence and Locality

From the above discussion, we can see that there are two important factors monitoring the interpretation of *ta-ziji*. One is Prominence, and the other Locality. Instead of saying that the locality constraint can be relativized in Chinese, as argued in Pan (1998), we want to say that the prominence constraint can be ranked higher than the locality constraint in Chinese. We think that different languages may have different rankings of these two constraints. For example, English ranks the locality constraint higher than the prominence constraint, and thus its reflexives cannot be bound to an antecedent across the local subject even though the subject is not so prominent as the potential antecedent, as exemplified in (3). However, the reflexive in Chinese can be bound to a more prominent antecedent across the local subject since in Chinese the prominence constraint is ranked higher than the locality constraint. The relevant constraints are defined as follows:

(10) Locality Constraint (LC)
A reflexive should select the closest NP as its antecedent.

(11) The Closeness Condition (Pan 1998)
α is closer to X, the reflexive, than β is iff the path from X to the minimal maximal projection dominating α is a proper subset of the path from X to the minimal maximal projection dominating β.

(12) Prominence Constraint (PC)
An anaphor cannot be bound to α across an intervening NP β which is not less prominent than α.

We think that the prominence of a NP is determined by two factors: (i) Grammatical Function; (ii) Animacy Hierarchy. The former can be represented by the feature [± Subject], and the latter [±Animate]. The value of the feature [±Animate] is determined by the following Animacy Hierarchy:

(13) Animacy Hierarchy (Chou 1992)
[+Human] > [+Animate, -Human] > [-Animate]

The interaction of the feature [± Subject] and the feature [±Animate] will yield the following results:

(14) a. [+ SUBJ, +ANIM]
b. [-SUBJ, +ANIM]
c. [+SUBJ, -ANIM]
d. [-SUBJ, -ANIM]

The feature specifications given in (14) can be ranked into the following prominence hierarchy:

(15) [+ SUBJ, +ANIM] > [-SUBJ, +ANIM] > [+SUBJ, -ANIM] > [-SUBJ, -ANIM]

Our definition of prominence is in spirit similar to Bresnan's (2001) definition of prominence in her Prominence Principle which says that "a binder excludes from its domain any personal pronouns more prominent than it". Although the basic idea underlying Bresnan's Prominence Principle is that a bindee cannot be more prominent than a binder, and thus does not concern the cases we consider here, her definition of prominence is, however, in every aspect relevant to the cases that concern us. Bresnan's (2001) prominence relation considers (i) linear order native to c-structures which model the ordering of overt perceptible expressions, (ii) syntactic rank native to f-structures which model the grammatical relations, whether expressed or unexpressed, (iii) and the thematic hierarchy. Hence, under her formulation, prominence relation can be defined along several dimensions that vary crosslinguistically. At this point, we follow Bresnan (2001) in hypothesizing that the parameterized prominence relation for Chinese is (15).

Besides the locality and prominence constraints, there are two other constraints that also play a role in reflexive binding. One is the feature compatibility constraint, and the other the I-within-I constraint.

(16) Feature Compatibility Constraint (FCC)
Two coindexed elements must have compatible features.

(17) I-within-I Constraint (IC)
An anaphor cannot be coindexed with an NP that dominates it.

The following example demonstrates that a reflexive must be bound to an antecedent with compatible features.

(18) *John$_i$ xihuan ni-ziji$_i$.
like you-self
*John likes yourself.

The I-within-I constraint can explain why the coindexation is illicit in the following sentence:

(19) *John yiwei ta-ziji$_i$ de pengyou$_i$ bu hui lai.
think he-self DE friends not will come
John thought that his own friends would not come.

4. Constraint Ranking in Chinese

Among the constraints discussed in the above section, FCC is assumed to be a hard constraint under our analysis since it is hard to imagine that it can be violated in any language. In Chinese, FCC should be ranked as high as PC, but cannot be ranked higher than PC. The reason is obvious. If FCC is ranked higher than PC, undesirable results will be produced, as exemplified by (5), repeated as (20) below:

(20) *John$_i$ juede wo$_j$ xihuan ta-ziji$_{i/j}$.
think I like he-self
John thinks I like him.

The coindexation of *John* with *ta-ziji* violates PC, according to (12), as the local subject *wo* 'I' is as prominent as *John*, but it satisfies FCC. If FCC were ranked higher than PC, we would wrongly predict that *John* can bind *ta-ziji*.

Although FCC cannot be ranked higher than PC, one might think that there is another option: it is ranked lower than PC. However, this option is not available, as shown below:

(21) Wo wei John zhaodao-le ta-ziji de zhaopian.
I for find-Perf he-self DE photo
I found John's photo for him.

In (21) *wo* is the most prominent NP, but the reflexive cannot be bound to it since they have incompatible features. The above examples in (20) and (21) show that (i) FCC cannot be violated, thus excluding the possibility that it is ranked lower than PC or any other constraints, and (ii) FCC cannot be satisfied by violating PC, thus excluding the possibility that it is ranked higher than PC. If FCC should assume the same ranking as PC, we can rank the relevant constraints as follows:

(22) Constraint Ranking
FCC/PC >>LC

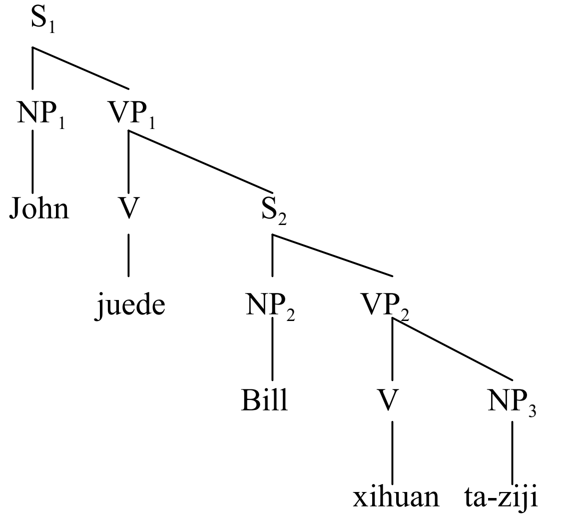
The constraint ranking given in (22) can correctly predict the binding possibilities of *ta-ziji* in the following sentences:

(23) a. John$_i$ xihuan ta-ziji$_i$.
like he-self
John likes himself.
b. John$_i$ juede Bill$_j$ xihuan ta-ziji$_{*i/j}$.
think like he-self
John$_i$ think that Bill$_j$ like himself$_{*i/j}$.
c. *John$_i$ juede wo$_j$ xihuan ta-ziji$_{i/j}$.
Think I like he-self
John think that I like him.

The coindexation between the reflexive and the antecedent in (23a) does not violate any constraint, and is thus ruled in. In (23b) there are two candidates for the reflexive. One is the local subject, and the other the matrix subject. The coindexation between the reflexive and the local subject does not violate any constraint, whereas the coindexation between the reflexive and the matrix subject violates two constraints, as demonstrated below:

(24)

candidates	FCC	PC	LC
a. ☞ Bill$_i$, ta-ziji$_i$			
b. John$_i$, ta-ziji$_i$		*!	*

(25)


The coindexation between the reflexive and the matrix subject in (23b) violates LC because the local subject *Bill* is closer to the reflexive than the matrix subject *John*, as shown in (25). The path W from the reflexive *ta-ziji* to the minimal maximal projection S_2 dominating NP_2 is {NP_3, VP_2, S_2}, and the path X from the reflexive *ta-ziji* to the minimal maximal projection S_1 dominating NP_1 is {NP_3, VP_2, S_2, VP_1, S_1}. Since W is a proper subset of X, *Bill* is closer to the reflexive than *John*, according to (11). The coindexation between the reflexive and the matrix subject also violates PC because the local subject is not less prominent than the matrix subject, according to (15).

In (23c) there are two binding possibilities for the reflexive *ta-ziji*, but neither of them is licit according to our account.

(26)

candidates	FCC	PC	LC
a. Wo_i, $ta\text{-}ziji_i$	*!		
b. $John_i$, $ta\text{-}ziji_i$		*!	*

The tableau above shows that in (23c) the coindexation between the reflexive and the matrix subject *John* violates PC, whereas the coindexation between the reflexive and the local subject violates FCC. Although (26a) violates less constraints, as compared with (26b), it does not survive because it violates a hard constraint. As a result, both candidates will be correctly ruled out by our grammar.

The constraint ranking given in (22) can also correctly predict the binding possibilities of the following sentences:

(27) a. $John_i$ shuo naben shu_j hai-le $ta\text{-}ziji_{i/*j}$
say that book hurt-Perf he-self
John said that that book hurt himself.
b. $John_i$ shuo [$Bill_j$ de xiaocongming]$_k$ hai-le $ta\text{-}ziji_{i/j/*k}$.
say DE little-trick hurt-Perf he-self
John said that Bill's little trick hurt him.
c. Wo_i wei $John_j$ zhaodao-le $ta\text{-}ziji_{*i/j}$ de zhaopian.
I for find-Perf he-self DE photo
I found John's photo for him.

The tableau for (27a) is given in (28):

(28)

candidates	FCC	PC	LC
a. $nabenshu_i$, $ta\text{-}ziji_i$	*!		
b. ☞ $John_i$, $ta\text{-}ziji_i$			*

The tableau above shows that, although (28b) violates LC, it does not violate any other higher-ranked constraints, and is thus better than (28a), which has a fatal violation of FCC. Note that (28b) does not violate PC, though it is bound to the matrix subject across the local one, since the latter is less prominent than the former.

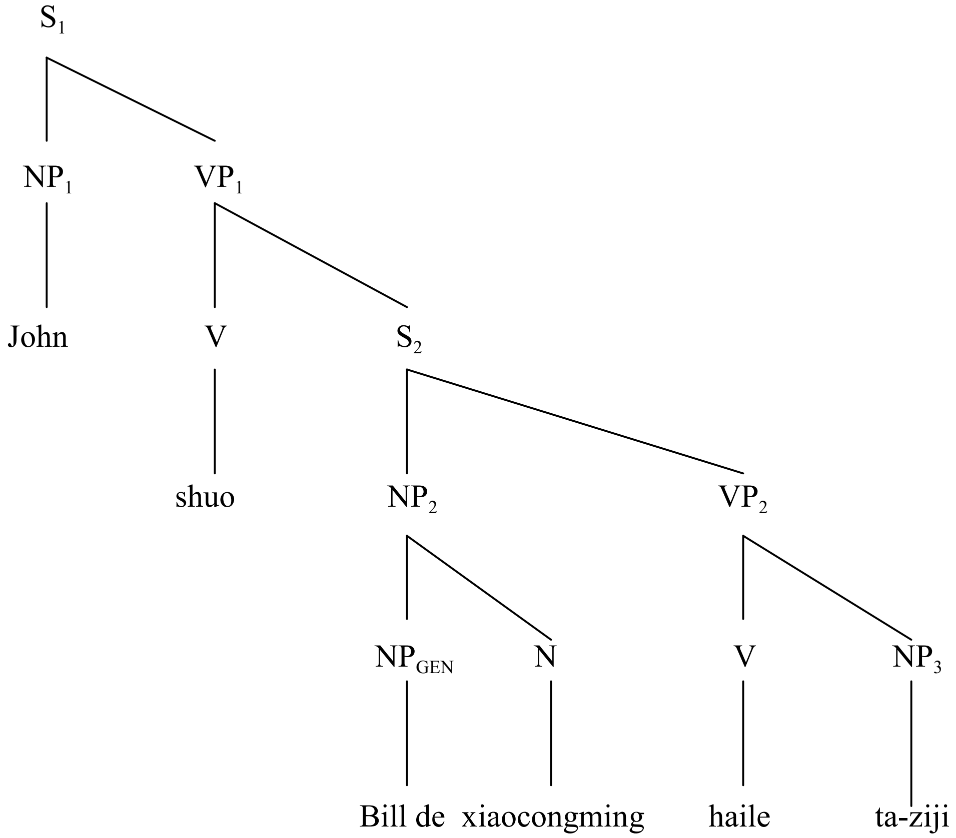
The tableau for (27b) is given in (29):

(29)

candidates	FCC	PC	LC
a. xiaocongming$_k$, ta-ziji$_k$	*!		
b. ☞ Bill$_i$, ta-ziji$_i$			*
b. ☞ John$_i$, ta-ziji$_i$			*

In the above tableau, (29a) is ruled out because it violates the highest-ranked constraint FCC, whereas both (29b) and (29c) are ruled in because the constraint they violate is ranked as the lowest in the constraint hierarchy. Note that *Bill* in (27b) is not closer to the reflexive than *John* since there is no subset relation between the two candidates, as shown below:

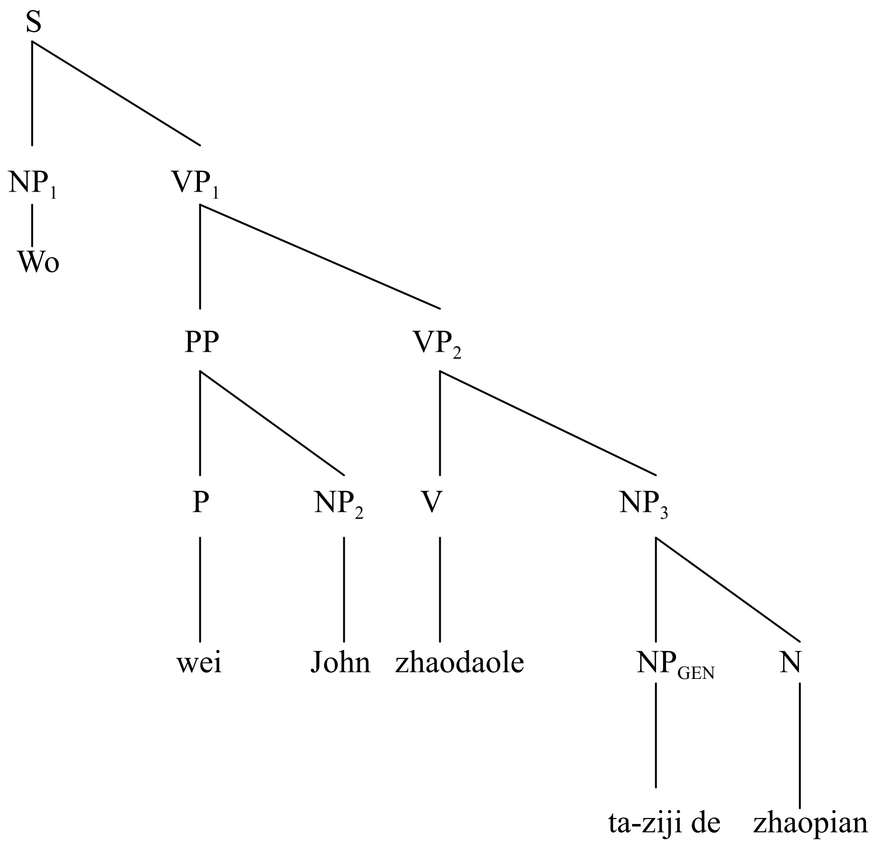
(30)



The path W from the reflexive to the minimal maximal projection S_2 dominating NP_2 is $\{NP_3, VP_2, S_2\}$, and the path X from the reflexive to the minimal maximal projection NP_2 dominating NP_{GEN} is $\{NP_3, VP_2, S_2, NP_2\}$, and the path Z from the reflexive to the minimal maximal projection dominating NP_1 is $\{NP_3, VP_2, S_2, VP_1, S_1\}$. Although W is a proper subset of X and Z respectively, there is no subset relation between X and Z, and hence, *Bill* is not closer to the reflexive than *John*, though both *John* and *Bill* are less closer to the reflexive than *Bill de xiaocongming* 'Bill's little tricks'. Note that the binding of the reflexive to the matrix subject does not violate PC since both *Bill* and *Bill de xiaocongming* are less prominent than the matrix subject.

Now, let's consider (27c), the structure of which is given in (31):

(31)



As shown in the structure above, there is no subset relation between the path W from the reflexive to the PP dominating NP_2, {NP_{GEN}, NP_3, VP_2, VP_1, PP}, and the path X from the reflexive to the S dominating NP_1, {NP_{GEN}, NP_3, VP_2, VP_1, S}. Hence, the reflexive can be bound to both *John* and *wo* without violating LC. Although LC is satisfied by both candidates, only the candidate in (32b) can survive because the other candidate incurs a fatal violation of FCC, as shown below:

(32)

candidates	FCC	PC	LC
a. Wo_i, ta-$ziji_i$	*!		
b. ☞ $John_j$, ta-$ziji_i$			

5. Blockers

Pan (1998) notes that, although subjects are blockers for *ta-ziji*, non-subjects can also become blockers if they dominate the potential antecedents, as shown below:

(33) John$_i$ quan [Bill$_j$ de baba]$_k$ kan ta-ziji$_{i/*j/k}$ de zhaopian.

Persuade DE father see he-self DE picture

John persuaded Bill's father to have a look at his picture.

In (33), the non-subject *Bill de baba* 'Bill's father' blocks the NP *Bill* it dominates from being the antecedent of *ta-ziji*, though it cannot block the matrix subject *John* from binding *ta-ziji*. We think that we need not stipulate what NP can be blockers in the grammar since the blocking effect can be derived from our PC. The following tableau shows why *Bill* is blocked in (33):

(34)

candidates	FCC	PC	LC
a. ☞ baba$_k$, ta-ziji$_k$			
b. Bill$_i$, ta-ziji$_i$		*!	*
c. John$_i$, ta-ziji$_i$			*

The above tableau shows that the binding of the reflexive to *Bill* violates both LC and PC, and thus should be filtered out. It violates PC because the closer NP *baba* 'father' is not less prominent than *Bill*. Hence, the reflexive cannot be bound to *Bill* across *baba*. Notice that the binding of the reflexive to *John* does not violate PC since *John* is more prominent than both *Bill* and *baba*. Although our account correctly predicts that (34a) is the optimal candidate, it wrongly predicts that (34c) should be ruled out since it is sub-optimal according to the above tableau. However, this prediction is false since (34c) is grammatical. We think that this problem can be solved if we reformulate our definition of LC as follows:

(35) Locality Constraint (LC)

A reflexive should select either the closest subject or the closest NP as its antecedent.

With this new definition of LC, we can give a new tableau for (33), as shown below:

(36)

candidates	FCC	PC	LC
a. ☞ baba$_k$, ta-ziji$_k$			
b. Bill$_i$, ta-ziji$_i$		*!	*
c. ☞ John$_i$, ta-ziji$_i$			

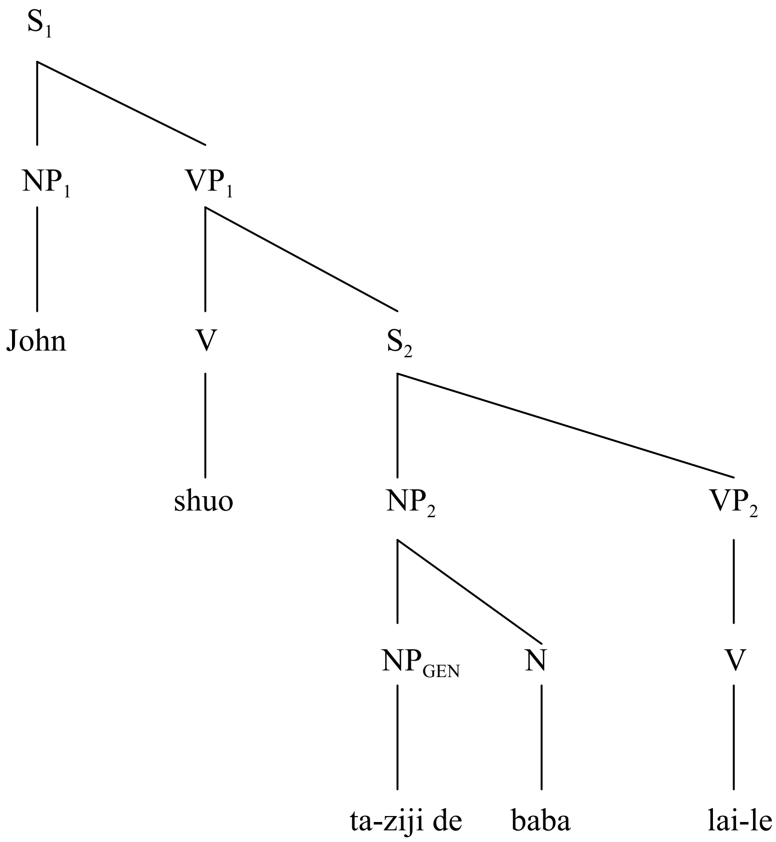
In the above tableau, (36b) violates LC since *Bill* is neither the closest subject nor the closest NP. Note that (36c) does not violate LC any more since *John* is the closest subject.

Although in the examples discussed above, the violation of PC will result in the blocking effect, it does not mean that PC is inviolable in Chinese. In fact, PC can be overrun by IC, as exemplified below:

(37) John$_i$ shuo [ta-ziji$_i$ de baba]$_k$ lai-le.
say he-self DE father come-Perf
John said that his father had come.

The structure of (37) can be represented as (38), in which the path W from the reflexive *ta-ziji* to the minimal maximal projection S_2 dominating NP_2 is $\{NP_{GEN}, NP_2, S_2\}$, and the path Z from the reflexive to the minimal maximal projection S_1 dominating NP_1 is $\{NP_{GEN}, NP_2, S_2, VP_1, S_1\}$. Since W is a subset of Z, *baba* is closer to the reflexive than *John*. Note that *baba* is also an NP that is not less prominent than *John* since the former has the same prominence value as the latter, according to (15). Hence, the coindexation of *John* with *ta-ziji* violates PC, and that of *baba* with *ta-ziji* violates IC. If IC were not ranked higher than PC, *baba* would block *John* from binding *ta-ziji*, which is at odds with the fact.

(38)



The above fact indicates that IC should dominate PC in constraint ranking, and the relevant ranking among IC, PC and LC is given in (39).

(39) Constraint Ranking
IC >> PC >> LC

The tableau for (37) is given in (40).

(40)

candidates	IC	PC	LC
a. ☞ $John_i$, $ta\text{-}ziji_i$		*	*
b. $baba_i$, $ta\text{-}ziji_i$	*!		

Note that (39) is different from the Constraint Ranking given in (22) since it does not include FCC. We find that we cannot merge (22) and (39) into one constraint ranking since in (22) FCC is ranked as high as PC, and if we include IC in (22), we do not know where to put IC in the relevant constraint ranking. If we rank IC higher than PC, it will also be ranked higher than FCC since FCC assumes the same status as PC. This result is obviously undesirable since FCC is a hard constraint that cannot be violated by satisfying IC, which is also a hard constraint. All thes suggest that there should be two constraint rankings that work for the interpretation of reflexives in Chinese. Along this line, we can reformulate (22) and (39) as (41).

(41) Constraint Ranking for Complex Reflexives in Chinese

a. FCC/PC >> LC

b. IC >> PC >> LC

The above constraint rankings show that IC can only override PC, but not FCC since there is no ranking between them. Note that, although hard constraints like FCC and IC participate in constraint ranking, they are inviolable in the sense that they cannot be ranked lower than any other constraints. We think that the variation among languages can only be reflected by different rankings of soft constraints, but not hard constraints.

6. Conclusion

In this paper we have shown that the binding of reflexives is regulated by the ranking of different constraints: IC, FCC, PC, and LC. Among these constraints, IC and FCC are considered to be hard constraints, whereas PC and LC are considered to be soft constraints. We think that it is necessary to make a distinction between hard constraints and soft constraints: the former are inviolable, whereas the latter are violable. Hence, we deviate from the standard OT hypothesis that all constraints are violable. We think that it is both conceptually and empirically plausible for us to posit the existence of hard constraints in languages. For instance, it is hard for us to imagine that there are languages in which FCC can be violated. Unlike hard constraints, soft constraints are more active in predicting and accounting for different binding properties and possibilities of reflexives in different languages. We think that different rankings of PC and LC play a major role in the interpretation of reflexives in different languages. In Chinese, PC is always ranked higher than LC, whereas in English, PC can be overrun by

LC.

References:

Bresan, Joan (2001) Lexical-Functional Syntax. Oxford: Blackwell.

Burzio, Luigi (1998) Anaphora and soft constraints. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pestetsky, eds., *Is the Best Good Enough*? Cambridge, MA: MIT Press.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Foris Publications, Dordrech, Holland.

Cole, Peter and C.-Z. Wang. (1996) Antecedents and blockers of long distance reflexives: The case of Chinese *ziji*. *Linguistic Inquiry* 27, 357-90

Hu, Jianhua and Haihua Pan. 2001. OT model and anaphoric binding. *Journal of Foreign Languages* 1, 2001.

Hu, Jianhua and Haihua Pan. To appear. NP Prominence Computation and the Reference Resolution of *Ziji*. *Contemporary Linguistics*.

Huang, C.-T. J. (1983) A note on the binding theory. *Linguistic Inquiry* 14: 554-561.

Huang, C.-C. James and C.-C. Jane Tang. (1991) The local nature of long distance reflexive in Chinese. In Jan Koster and Eric Reuland, editors, *Long distance Anaphora*. Cambridge University Press, Cambridge.

Huang, Yan. (1994) *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. Cambridge University Press, New York.

Pan, Haihua (1995) *Locality, Self-Ascription, Discourse Prominence, and Mandarin Reflexives*. Ph.D. thesis, Dept. of Linguistics, the University of Texas at Austin, Austin, TX.

Pan, Haihua (1997) *Constraints on Reflexivization in Mandarin Chinese*. Garland Publishing, Inc., New York.

Pan, Haihua (1998) Closeness, prominence and binding theory. *Natural Language & Linguistic Theory* 16, 771-815.

Pan, Haihua (2001) Why the blocking effect? In Peter Cole, Gabriella Hermon and James Huang, eds., *Long Distance Reflexives*. Syntax and Semantics Series, Academic Press, 279 – 316.

Pica, P. (1987) On the nature of reflexivization cycle. *Proceedings of the North Eastern Linguistics Society 17*: 483-499.

Prince, Alan & Paul Smolensky (1993) *Optimality theory: Constraints in generative grammar*. RuCCs Technical Report #2. Rutgers University, Center for Cognitive Science. [To appear, Cambridge, MA: MIT Press.]

Reinhart, Tanya and Eric Reuland. (1993) Reflexivity. *Linguistic Inquiry* 24, 657-720.

Sells, Peter. (1987) Aspects of logophoricity. *Linguistic Inquiry* 18, 445-479.

Tang, C. -C. J. (1989) Chinese reflexives. Natural Language and Linguistic Theory 7.1, 93-122.

Tang, C. -C. J. (1994) A note on relativized SUBJECT for reflexives in Chinese. In B. Lust, M. Sûner, and J. Whitman eds., *Syntactic theory and first language acquisition: Cross-linguistic perspectives.* Vol. 2, *Heads, projections, and learnability*. Hillsdale, N.J: Lawrence Erlbaum.

LFG for Chinese: Issues of Representation and Computation

Sun Maosong
The State Key Lab. of Intelligent Technology & Systems, Tsinghua University
Beijing 100084, China
<sms@s1000e.cs.tsinghua.edu.cn>

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Mirriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu/

ABSTRACT

LFG has been widely used to analyze English language as well as other languages from linguistic point of view [Joan Bresnan 2001; Louisa Sadler 1996], including Chinese [Lian-Cheng Chief 1996; One-Soon Her. 1997]. A new direction in LFG research field is applying it to language computation, ranging from parsing to machine translation [Louisa Sadler, Josef van Genabith, and Andy Way 2000; Mark Johnson 2000; Miriam Butt, Stefanie Dipper, Anette Frank, and Tracy Holloway King 1999]. However, the LFG-based work in Chinese computing is rather rare [Lian-Cheng Chief, Chu-Ren Huang, Keh-Jiann Chen *et al* 1998].

The current framework of LFG shows two folds when being employed in Chinese computing tasks: it is quite powerful for linguistic representation, but seems not to be strong enough for Chinese computation – there exists some room for improving the formalism of LFG. This paper will focus on these two issues, suggesting some possible augmentations on LFG paradigm, though the idea is still preliminary. The author believes linguistic resources, such as annotated corpora, mainly semantics-oriented, are also required to make manipulations on the augmented paradigm possible. The total solution is based on not only academic research but also engineering realization – it will not work without either.

1. Introduction*

LFG has been widely used to analyze English as well as other languages from a linguistic point of view [Joan Bresnan 2001; Louisa Sadler 1996], including Chinese [Lian-Cheng Chief 1996; One-Soon Her. 1997]. A new direction in LFG research field is applying it to language computation, ranging from parsing to machine translation [Louisa Sadler, Josef van Genabith, and Andy Way 2000; Mark Johnson 2000; Miriam Butt, Stefanie Dipper, Anette Frank, and Tracy Holloway King 1999]. However, the LFG-based work in Chinese computing is rather rare [Lian-Cheng Chief, Chu-Ren Huang, Keh-Jiann Chen *et al* 1998].

The current framework of LFG shows two folds when being employed in Chinese computing tasks: it is quite powerful for linguistic representation, but seems not to be strong enough for Chinese computation – there exists some room for improving the formalism of LFG. This paper will center on these two issues.

2. LFG for Representing Chinese Linguistic Phenomena

LFG is powerful in describing linguistic phenomena of Chinese. Even a very sophisticated sentential construction could be successfully explained by LFG. Take sentence 1 as an example:

(1)	Zhang-san	fang4	gou3	yao3	si3	le	Li-si.
	person1	send	dog	bite	die	AUX	person2
	N1	V1	N2	V2	V3	AUX	N3

Zhang-san sent the dog to bite Li-si, and Li-si died.

The following observations hold for this quite complex sentence: (i) N1 and N2 are SUBJECT and OBJECT of V1 respectively; (ii) V2 and V3 form a verbal phrase VP in c-structure (V3 serves as the complement of V2 syntactically); (iii) from f-structure point of view, N2 and N3 should logically be SUBJECT and OBJECT of V2, meanwhile N3 logically be SUBJECT of V3; and (iv) “yao3 si3 le Li-si” is XCOMP of V1 (see Fig 1).

* This research is supported by 973 Project of China under grant no G1998030507 and Foundation for the University Key Teacher by the Ministry of Education of China.

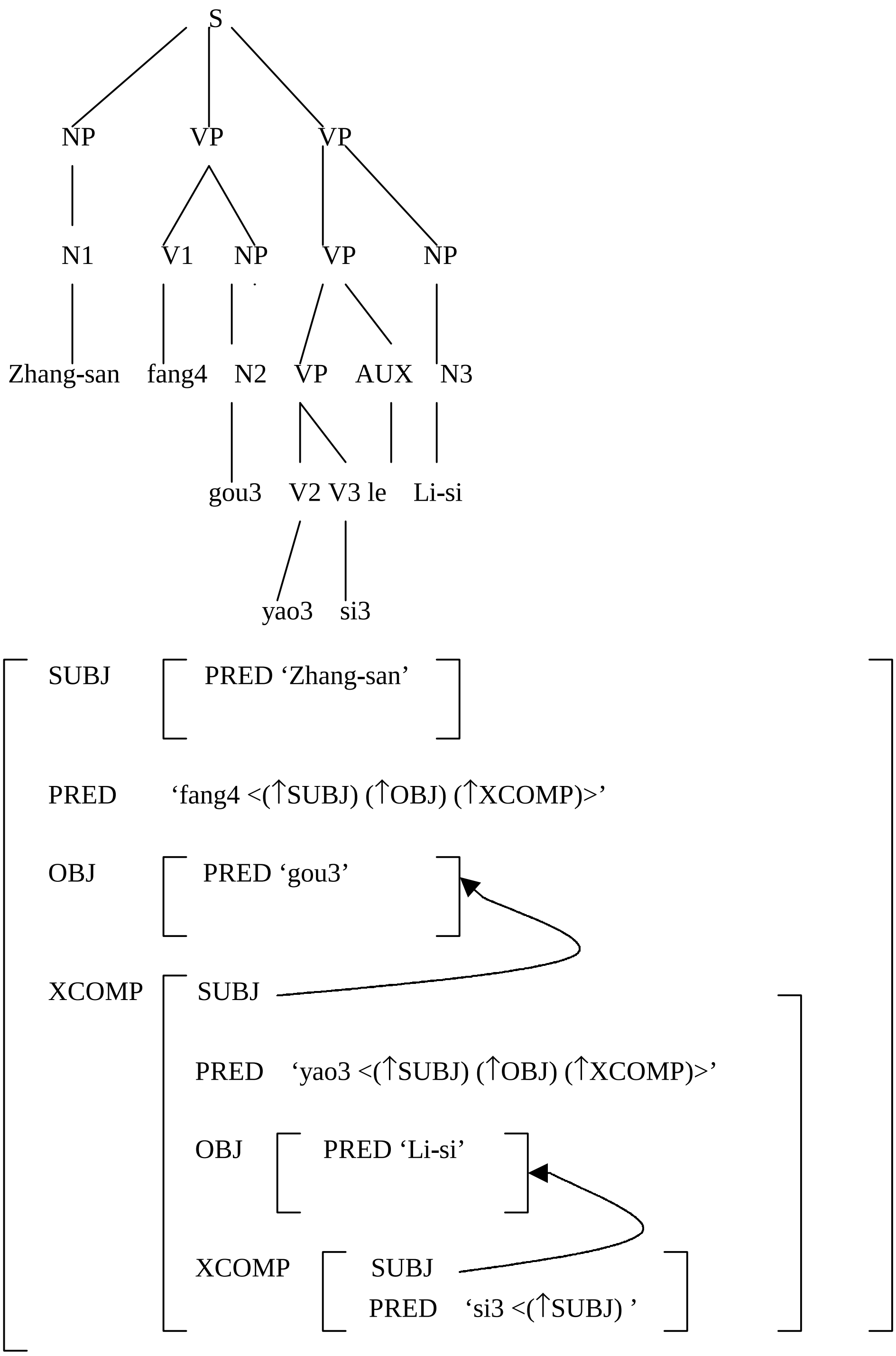


Fig.1. The c-structure and f-structure of sentence (1)

Though the mapping from c-structure to f-structure for (1) is not straightforward (note that V2 and V3 should be combined syntactically while split out semantically), it can still be built up quite easily supposing that one has already comprehended the sentence in advance. We would get similar conclusions when dealing with other types of typical linguistic constructions in Chinese.

3. LFG for Chinese Computing

3.1. The Role of Semantics in Computation

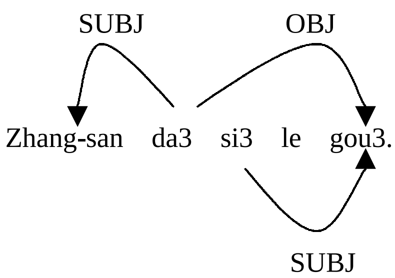
Here, I would like to show that LFG, as a computational formalism, still has some limitations for Chinese computing. The fact that Chinese is an inflection-free language (for instance, neither change in form nor explicit marker is used when a verb functions as main verb, clause, infinitive, modifier of nouns, or head of noun phrases) may result in large number of ambiguities at every linguistic level for machines. The mapping between c-structure and f-structure, as well as the mapping between f-structure and a-structure are extremely difficult to figure out, if semantic information is not provided sufficiently. Consider a group of sentences:

(2a) Zhang-san da3 si3 le gou3.
person hit die AUX dog
N1 V1 V2 AUX N2
Zhang-San hit the dog, and the dog died.

(2b) Zhang-san he1 zui4 le jiu3
person drink drunk AUX wine
N1 V1 V2 AUX N2
Zhang-san drank (the wine), and (Zhang-san) got drunk.

(2c) Zhang-san ku1 zhong3 le yan3jing1.
person cry ‘get turgid’ AUX eye
N1 V1 V2 AUX N2
Zhang-San cried, and (his) eyes got turgid.

The c-structures of these three sentences are patterned in the same way, but their f-structures are quite different, as illustrated in Fig.2: in (2a), N2 is both OBJECT of V1 and SUBJECT of XCOMP; in (2b), N2 is still OBJECT of V1, but SUBJECT of XCOMP becomes N1; in (2c), N2 serves only as SUBJECT of XCOMP, no longer OBJECT of V1.



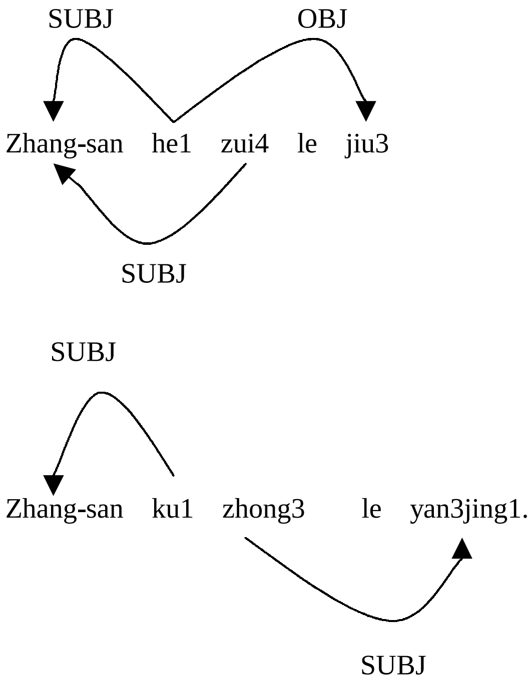


Fig.2. The f-structures (sketch) of sentences (2a) (2b) and (2c)

What makes the distinction here? The answer is obvious: it is nothing but semantic constraints among V1, V2, N1 and N2 control the one-to-many mapping processes from c-structure to f-structures.

Similar cases can be frequently encountered in Chinese. Consider another group of sentences which concerns "V1+N1+de+N2", a popular syntactically ambiguous construction in Chinese:

(3a) Yao3 lie4ren2 de gou3
bite hunter AUX dog
V1 N1 AUX N2
The dog that bites the hunter (NP)
To bite the hunter's dog (VP)

(3b) Yao3 lie4ren2 de ji1
bite hunter AUX chicken
V1 N1 AUX N2
To bite the hunter's chicken (VP)

(3c) Yao3 tu4zi3 de gou3
bite rabbit AUX dog
V1 N1 AUX N2
The dog that bites the rabbit (NP)

Both the c-structures and f-structures of these sentences differ this time. Again, semantic constraints among V1, N1 and N2 play critical role in the relevant analyses, determining which sentence out of (3a) (3b) and (3c) is realized as ‘true’ syntactic ambiguity and which one does not (Fig. 3):

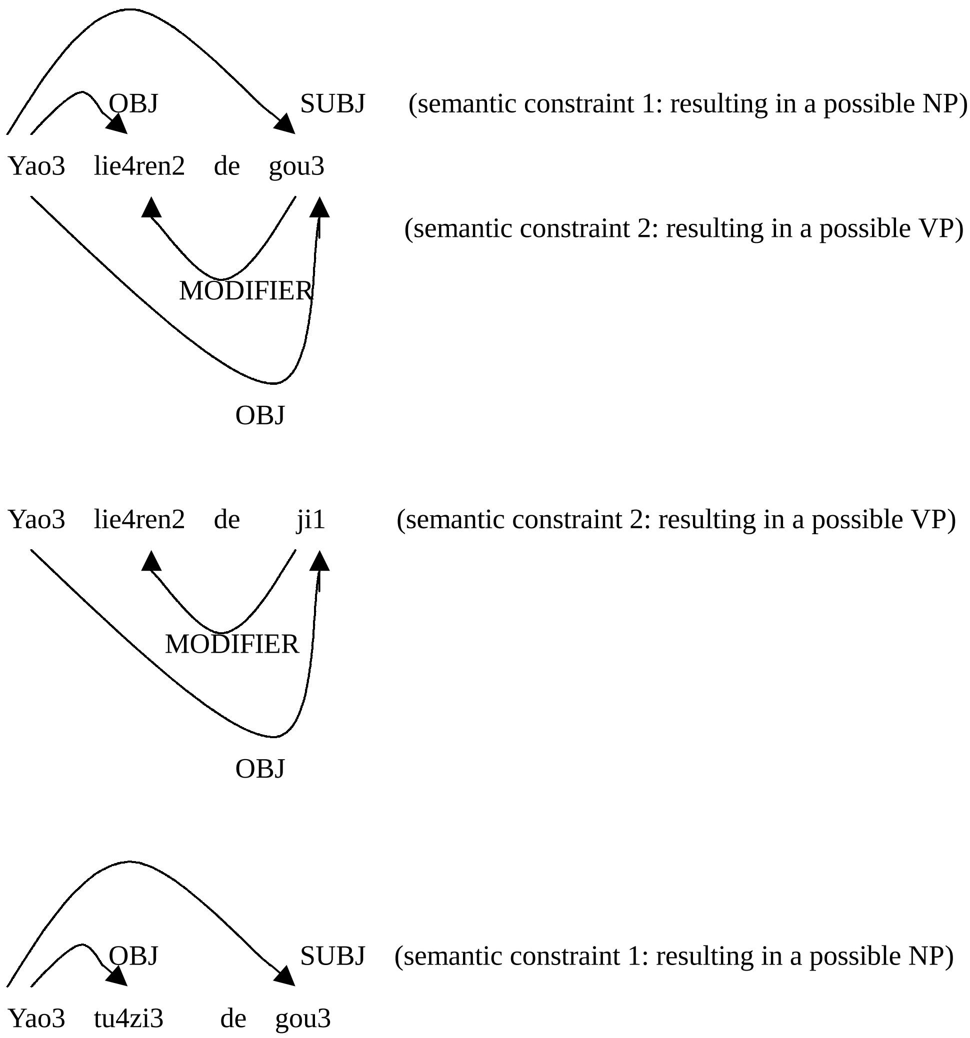


Fig.3 Semantic constraints in potentially ambiguous construction “V1+N1+de+N2”

The situation can be even more complicated if word segmentation ambiguities are to be included. Incorrect segmentation may still lead to a syntactically well-formed but semantically ill-formed ‘sentence’. For example, given the input sentence “Zhang san fang4 huo3 shao1 si3 le mei4 mei4” (note that each token here is a Chinese character, rather than word, and there is no spacing between either adjacent characters or adjacent words in original writings), the correct segmentation for it should be:

(4a) Zhang-san fang4 huo3 shao1 si3 le mei4mei4.
person make fire burn die AUX sister
N1 V1 N2 V2 V3 AUX N3
Zhang-San set up a fire, to burn (my) sister, and sister died.

But another possible segmentation exists, -- it is well-formed syntactically but almost ill-formed semantically (and, the sense of ‘fang4’ is changed from ‘make’ in (4a) to ‘put on’ in (4b)):

(4b) Zhang-san fang4 huo3shao1 si3 le mei4mei4.
person put on cake die AUX sister
N1 V1 N2 V2 AUX N3
Zhang-San put on cake, and sister died.

In order to filter out (4b), a computational mechanism at semantic level is absolutely necessary.

3.2. Possible Augmentation on LFG Framework

The point addressed here is that semantic analysis is likely to be in a dominant position in computing Chinese sentences. Manipulations on a-structure, f-structure and c-structure should be carried out jointly and in parallel. To render LFG truly computable for Chinese, I believe that some augmentation is needed accordingly:

(i) Experience tells us it is easy for human to reveal those semantic constraints, but how about machines? Recall sentence (1). Suppose a machine is asked to derive f-structure from this input sentence. The pattern of the c-structure of the sentence tail, “gou3 yao3 si3 le Li-si”, is totally the same as that of the sentences (2a) (2b) and (2c). Which f-structure in Fig.2 should be assigned to this fragment (The correct one is (2a))? Of course, we need to feed all the relevant knowledge to the machine. To enable the machine to treat unrestricted texts, the knowledge ought to be given in detail and systematically, -- in more computational terms, it must take every combinatorial possibility of constraints among V1, V2, N1 and N2 into account. An accurate way of providing such knowledge is to take ‘word’ as basic factoring unit in lexicon, that is, attempting to enumerate every collocation-like ‘semantic’ correspondence between every possible word pair (In fact, the lexicon organized under current paradigm of LFG involves this sort of information implicitly).

da3: V, (↑PRED) = ‘ da3 < (↑SUBJ)(↑OBJ)(↑XCOMP)>’
(↑SUBJ PRED) = ‘Zhang-san’
(↑OBJ PRED) = ‘gou3’
(↑XCOMP SUBJ PRED) = ‘gou3’
(↑XCOMP PRED) = ‘si3< (↑SUBJ)>’

he1: V, (↑PRED) = ‘ he1 < (↑SUBJ)(↑OBJ)(↑XCOMP)>’
(↑SUBJ PRED) = ‘Zhang-san’
(↑OBJ PRED) = ‘jiu3’
(↑XCOMP SUBJ PRED) = ‘Zhang-san’
(↑XCOMP PRED) = ‘zui4< (↑SUBJ)>’

ku1: V, (↑PRED) = ‘ ku1 <(↑SUBJ)(↑ XCOMP)>’
(↑SUBJ PRED) = ‘Zhang-san’
(↑XCOMP SUBJ PRED) = ‘yan3jing1’
(↑XCOMP PRED) = ‘zhong3< (↑SUBJ)>’

In the case, the number of combination can be potentially $|V|^2 * |N|^2$, where $|V|$ and $|N|$ is number of verbs and nouns in the lexicon respectively. Imagine what a complex picture it would be! It is impossible to establish such a lexicon when facing the real world of Chinese language.

An alternative solution is to make some degree of approximations, shifting from ‘word’ to ‘semantic class of word’, and describing semantic constraints over semantic classes rather than word. The advantage of doing so is that the complexity of the task can be reduced dramatically – the number of word classes is at least one or two orders of magnitude smaller than that of words – so as to make the computation feasible:

da3: V, (↑PRED) = '#HIT <(↑SUBJ)(↑OBJ)(↑XCOMP)>'

(↑SUBJ PRED) = ?#HUMAN

(↑OBJ PRED) = ?#ANIMATE

(↑XCOMP SUBJ PRED)

(↑XCOMP PRED) = '?#DIE< (↑SUBJ)>'

he1: V, (↑PRED) = ' #DRINK <(↑SUBJ)(↑OBJ)(↑XCOMP)>'

(↑SUBJ PRED) = ?#HUMAN

(↑OBJ PRED) = ?#WINE

(↑XCOMP SUBJ PRED)

(↑XCOMP PRED) = '?#GET-DRUNK< (↑SUBJ)>'

ku1: V, (↑PRED) = '#CRY < (↑SUBJ)(↑XCOMP)>'

(↑SUBJ PRED) = ?#HUAMN

(↑XCOMP SUBJ PRED) = ?#EYE

(↑XCOMP PRED) = '?#GET-TURGID < (↑SUBJ)>'

Where the symbol '#' denotes the succeeding entity is a semantic class, and '?' means that the machine need to find a word with the semantic class specified by the succeeding '#' in the input sentence.

(ii) In addition to the lexicon, a WordNet-like semantic system (including a conceptual hierarchy and a relation system among concepts, in particular among action concepts) is also indispensable, as required by the inference mechanism in computation.

(iii) To cope with a variety of ambiguities in conducting c-structure, f-structure as well as a-structure of any Chinese sentence effectively and efficiently, certain statistical mechanism should be incorporated into the LFG paradigm. For instance, we say that (4b) is ill-formed in meaning, -- this statement is relative: both "Zhang-san fang4 huo3shao1" and "Zhang-san si3 le mei4mei4" are well-formed semantically, so their combination "Zhang-san fang4 huo3shao1 si3 le mei4mei4" may still appear some degree of rationality. Though the sentence (4a) can be approved by 'pure' logical-form based calculations in terms of computational resources provided in (i) and (ii), the disapproval of the sentence (4b) will largely depend on the logical relation between two actions, #PUT-ON-ITEMS and #DIE. The decision could be made in terms of the probability of (4b):

PROB('Zhang-san fang4 huo3shao1', 'Zhang-san si3 le mei4mei')
≈PROB('fang4'(put on))*PROB('die')*
PROB(#PUT-ON-ITEMS | 'fang4')*PROB(#DIE | 'die')*
PROB(#PUT-ON-ITEMS, #DIE)
≈PROB(#PUT-ON-ITEMS, #DIE)

All the above statistical parameters are to be derived from large scale annotated corpora.

(iv) In line with (iii), the operation 'unification' in LFG ought to be augmented to fit the statistical calculation. A new attribute 'PROB' should be added into both static lexical entries in lexicon and f-structures dynamically generated during parsing procedure. The value of 'PROB' of two unified feature sets is calculated from the value of 'PROB' of each, in principle.

4. Conclusion

Inspired by work on Chinese computing, this paper has suggested some augmentations on LFG paradigm, though the idea is very preliminary. Other resources, such as annotated corpora, mainly semantics-oriented, are also required to make manipulations on the augmented paradigm possible. It is obvious that the total solution is based on not only academic research but also engineering realization – it will not work without either.

References

Anette Frank. 2000. Automatic F-structure Annotation of Treebank Trees. *Proceedings of LFG'00*.

Joan Bresnan. 2001. *Lexical-functional syntax*. Oxford: Blackwell Publishers.

Lian-Cheng Chief. 1996. An LFG Account of Mandarin Reflexive Verbs. *Proceedings of LFG'96.*

Lian-Cheng Chief, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai, and Lili Chang. 1998. What Can Near Synonyms Tell Us. *Proceedings of LFG'98*.

Louisa Sadler. 1996. New Developments in LFG. In Keith Brown and Jim Miller, editors, *Concise Encyclopedia of Syntactic Theories*. Elsevier Science, Oxford.

Louisa Sadler, Josef van Genabith, and Andy Way, 2000. Automatic F-structure Annotation of Treebank Trees and CFGs Extracted from Treebanks. *Proceedings of LFG' 00.*

Mark Johnson, 2000. Stochastic Lexical-Functional Grammar. *Proceedings of LFG' 00.*

Miriam Butt, Stefanie Dipper, Anette Frank, and Tracy Holloway King. 1999. ParGram Project. *Proceedings of LFG' 99.*

One-Soon Her. 1997. The Lexical Mapping Theory and Mandarin Resultative Compounds. *Proceedings of LFG'97*.

On the function COMP in Cantonese

Adams Bodomo and Sophia Lee
Department of Linguistics
University of Hong Kong

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Mirriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publication

http://csli-publications.stanford.edu/

Abstract

This paper discusses the function COMP with data from Cantonese. We first point out some possible problems for previous proposals for the elimination of COMP. We next take up the concept of 'mixed language' using Cantonese data. Initial results show that Cantonese is indeed a mixed language. We do, however, propose that the concept of mixed languages should be extended to include degrees or extents to which a language can be regarded as a mixed language. Our analysis in this paper supports a finer-grained categorization of grammatical functions in linguistic theory.

1. **Introduction**[*]

Grammatical function categorization and specification constitute an important issue for most formal theories of syntax. However, not all theoretical frameworks have the same taxonomy of functions. While familiar ones such as SUBJ and OBJ are recognized and differentiated in almost all theories of grammar, some others such as the OBL_{q} , COMP and XCOMP are not that familiar across frameworks. Some frameworks, such as LFG, are thus more finer-grained than others in terms of grammatical function categorization. In this paper we look more closely at the existence and relevance of one function, COMP, in the structure of Cantonese, a Yue dialect of Chinese as spoken in Hong Kong. COMP (or closed complement) was first introduced as a grammatical function in Lexical-Functional Grammar in Bresnan (1982a & b). However, its existence is not without controversy. Alsina, Mohanan and Mohanan (1996a) propose that clausal complements, which are commonly considered as bearing COMP function, can be taken as bearing the OBJ function since they possess the same kinds of syntactic properties as OBJ do. Other works consider that there is evidence for COMP as a distinct grammatical function (Lødrup 1996; Curly 1996; Dalrymple and Lødrup 2000). In this paper, we propose to contribute to this discussion with analysis of data on Cantonese COMPs. In section 2, we introduce and illustrate the existence of COMP and other functions such as OBJ in the structure of Cantonese. In section 3, we question whether it is possible to successfully argue against the existence of COMP as a grammatical function (Alsina, Mohanan and Mohanan 1996). In section 4, we take up the concept of *mixed language*, as introduced by Dalrymple and Lødrup (2000), as a motivation for retaining COMP as a grammatical function in Cantonese. We show that Cantonese is an OBJ/COMP mixed language based on the simple fact that there is

[*] We are grateful for comments on this paper from the following people: Helge Lødrup, Mary Dalrymple, and participants at LFG 2001.

alternation between COMPs and OBJs. Following this position, we propose that the concept of mixed language should be extended to include degrees or extents to which a language can be regarded as a mixed language. This will be briefly demonstrated in section 5. Finally, we show that a finer-grained categorization of grammatical functions in a linguistic theory like LFG is desirable.

2. Complement Functions and the Structure of Cantonese

Cantonese like other dialects of Chinese, has some very unique characteristics including pro-drop without a rich inflectional morphology (Luke, Bodomo and Nancarrow 2001), long distance anaphora (Pan and Hu 2001) and a quite complex verbal complementation. It is this latter aspect of Chinese that is addressed in this paper. We show below in examples (1) – (5) the full range of grammatical functions, such as OBJ, OBJ_{θ}, OBL_{θ}, COMP and XCOMP that captures verbal complementation in Cantonese, respectively.

(1)	*Nei5*	*gaau3*	[*ngo5*]		
	2.SG	teach	me		
	'You teach me.'				

(2)	*Nei5*	*bei2-zo2*	*bun2*	*syu1*	[*ngo5*]
	2.SG	give-PERF	CL	book	1.SG
	'You have given me a book. '				

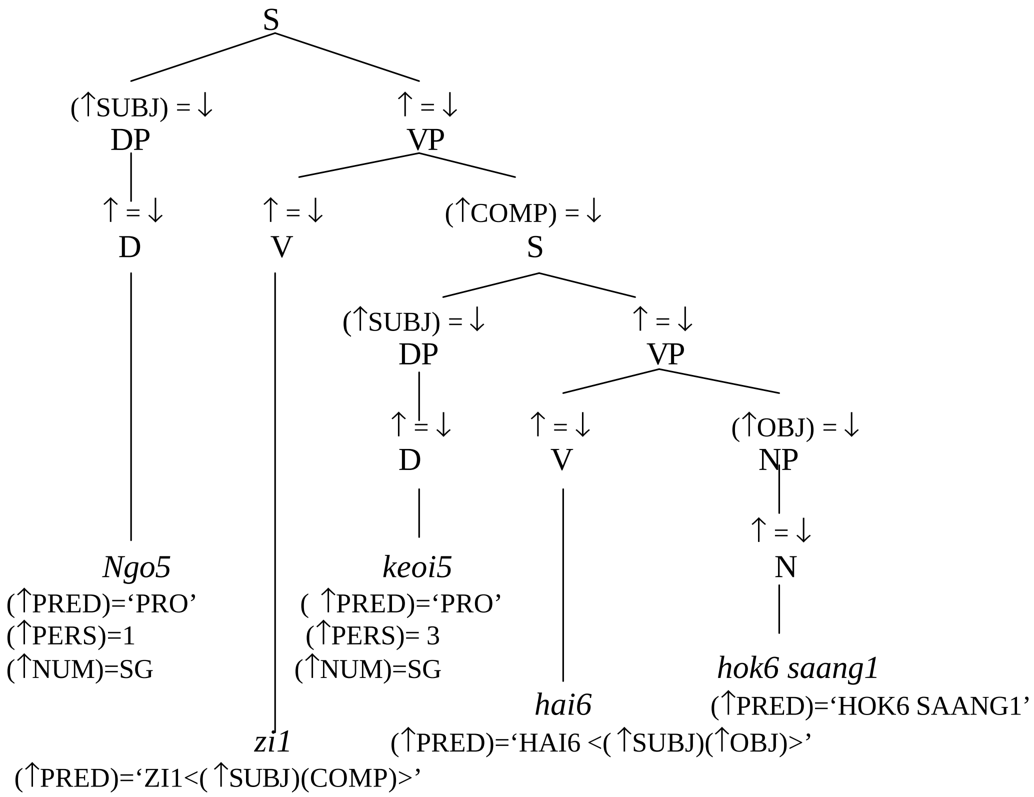
(3)	*Ngo5*	*zyu6*	[*hai2*	*zung1waan4*]
	1.SG	live	at	Central
	'I live in Central.'			

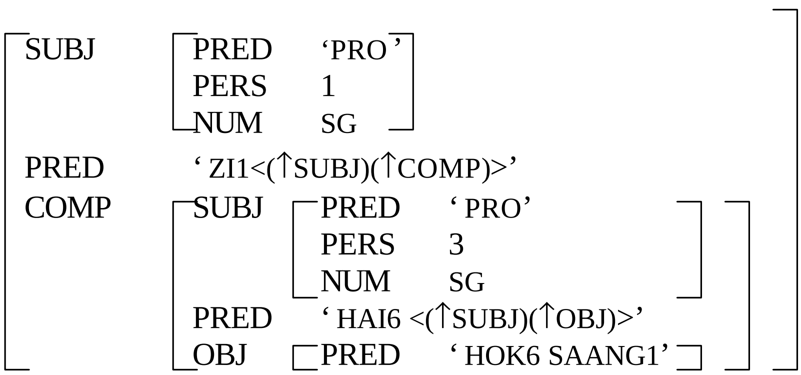
(4)	***Ngo5***	***zi1***	**[*keoi5***	***hai6***	***hok6***	***saang1*]**
	1.SG	knows	3.SG	be	student	
	'I knows that s/he is a student.'					

(5)	*Keoi5*	*jiu3*	*ngo5*	[*zau2*]
	3.SG	want	1.SG	leave
	'He/she asked me to leave'			

COMP is the embedded sentence which contains its own subcategorized arguments as illustrated in (4). In most languages it usually appears together with a complementizer which serves as a sentence introducer. In Cantonese, however, there is no obvious complementizer as *that* in English. The annotated c- and f-structure diagrams in (6) illustrate the occurrence of COMP in the structure of Cantonese.

(6) c- and f-structures of example (4): *Ngo5 zi1 keoi5 hai6 hok6 saang1*





COMP is realized as S in terms of phrase structure category as shown in the c-structure. The f-structure shows us that it consists of a PRED *hai6* which subcategorizes for two arguments, i.e. *keoi5* as the SUBJ and *hok6 saang1* as the OBJ. Neither of them needs to be controlled by the matrix functions. Having illustrated the occurrence of COMP as a grammatical function in Cantonese, we will try to examine a debate in the literature about the existence of COMP.

3. Should COMP be eliminated as a Grammatical Function?

As mentioned earlier, not all works in LFG agree that COMP should exist in the taxonomy of grammatical functions. In this section we question the relevance of the arguments against COMP with respect to Cantonese.

Alsina, Mohanan and Mohanan (1996) takes the position that all the sentential complements, in fact, bear the grammatical function OBJ or OBJ_q and hence COMP is unnecessary as a grammatical function and should thus be eliminated from the LFG taxonomy of grammatical functions. There are various reasons for this position. First, it is argued that the distinction between OBJ and COMP at f-structure is duplicated as NP and S respectively at c-structure. This duplication of information at the two levels of representation results in a redundancy and therefore the distinction of the two functions should be made in terms of phrase structure category at c-structure, but not of grammatical functions at f-structure.

However, this would depend on the kind of language one is dealing with. It is known that phrase structure category alone is not able to differentiate all the various types of syntactic behaviour. This is especially true for those non-configurational languages like Warlpiri and Malayalam in which their grammatical functions are not uniformly represented in terms of continuous phrases at c-structure. Even Cantonese, though not one such non-configurational language, provides an interesting piece of evidence here as shown in the case of a ‘pro -drop’ sentence in (7):

(7) $_{S}$[*Keoi5* $_{VP}$[*waa6* $_{VP}$[*m4* *leng3*]]]
3.SG say NEG pretty
‘S/he said that (it is) not pretty.’

Though COMP is usually realized in Cantonese as S in terms of phrase structure category, it is sometimes realized as VP as shown in (7) due to the absence of an overt subject in the embedded sentence. This piece of evidence in Cantonese indicates that phrase structure might not provide as clear an evidence for distinguishing OBJ and COMP in terms of NP and S as is suggested by Alsina, Mohanan and Mohanan (1996).

Another argument in favor of eliminating COMP is the claim that verbs which take a sentential complement can be passivized just as OBJ can be. Again, this might work for some languages but not for all. We believe that it is not a general trend for many languages. In Cantonese, COMP cannot normally be passivized as in (9):

(8) *Ngo5* *zi1 dou3-zo2* [*keoi5* *saat3-gwo3* *jan4*]
1.SG know-PERF 3.SG kill-PERF people
‘I know that s/he has killed a person.’

(9) ?*[*Keoi5* *saat3-gwo3* *jan4*] *bei2*[1] *ngo5* *zi1 dou3-zo2*
3.SG kill-PERF people BEI 1.SG know-PERF
‘That s/he has killed a person is known by me.’

The only way to get a structure that looks like a passivization of COMP would be to nominalize it by adding “li1+CL+N” or “ge3 + N” right after it. Consider the example below in (10):

(10) [*Keoi5* *saat3-gwo3* *jan4*] *li1* *gin6* *si6* *bei2* *ngo5* *zi1 dou3-zo2*
3.SG kill-PERF people this CL matter BEI 1.SG know-PERF
‘That s/he has killed a person is known by me.’

In (10), the S *Keoi5 saat3-gwo3 jan4* is moved to the initial position of the sentence,

[1] ‘ *bei2*’ serves as the passive marker in Cantonese.

however, it becomes the modifier of the SUBJ *li1 gin6 si6*.

Lastly, a major argument for the elimination of COMP as a grammatical function advanced by Alsina, Mohanan and Mohanan (1996) comes from the structure of the Lexical Mapping Theory (LMT)'s four-way classification of grammatical functions (Bresnan and Kanerva 1989, Bresnan 2001). The current LMT model decomposes grammatical functions into two binary features, $\pm r$ and $\pm o$. It is obvious that COMP has not been taken into consideration during the development of the theory. We don't believe, however, that it is a good enough reason to not include COMP in the taxonomy of LFG functions just because it has not been accounted for by the LMT. In any case, it is not only COMP that is not accounted for by the LMT. Other functions like XCOMP are not included. A better approach would probably be to consider expanding the generative power of LMT to account for more functions and not to eliminate COMP and others because of the apparent inability of LMT to handle them.

Indeed various alternative proposals have been made to augment the current LMT model. For instance, Zaenen and Engdahl (1994) propose that COMP and XCOMP bear the thematic role PROPOSITION.

Given the above, among other reasons, we take the position in this paper that COMP exists as a grammatical function in Cantonese.

4. Cantonese as an OBJ/COMP mixed language

Following from our arguments that COMP exists as a grammatical function in Cantonese, we would like to show in this section that some clausal complements function as COMPs and some as OBJs. The syntactic behaviour of clausal complements is language-specific. It has been shown that in some languages like Icelandic and Spanish, all clausal complements bear a uniform grammatical function, OBJ. Some clausal complements bear either OBJ or COMP function in other languages.

This manifestation of variability in the structure of languages in this respect has led Dalrymple and Lødrup (2000) to propose the term *mixed language* to describe a language in which clausal complements exists both as OBJs and COMPs. The existence of OBJ/COMP mixed languages implies the need for the distinction between COMP and OBJ function because even a hierarchically-defined distinction among grammatical functions cannot predict the different behavior of clausal complements in mixed languages (Dalrymple and Lødrup 2000). Following the introduction of the idea of mixed language, we will now try to examine if Cantonese is an OBJ/COMP mixed language.

The first and most important step in deciding if a language is an OBJ/COMP mixed language, which we define as one in which some clausal complements can function as OBJs while others function as COMPs is to see if there is an alternation of NP OBJs and COMPs in the complementation properties of some predicates in the language.

4.1. Alternation with NP object

In Cantonese, verbal predicates like *zi1* (*dou3*) 'know' and *seon3* 'believe' that take OBJ clausal complements also allow NP objects, i.e. they take either NP/DP/CLP[2] or S OBJs:

(11) a. *Ngo5* *zi1 dou3* $_{S}$[*keoi5* *hai6 hou2* *jan4*]
1.SG know 3.SG be good people
'I know that s/he is a good person.'

b. *Ngo5* *zi1 dou3* $_{DP}$[*li1* *gin6 si6*]
1.SG know this CL matter
'I know (about) this.'

[2] CLP = Classifier phrase

(12) a. *Ngo5* *seon3* S[*keoi5* *hai6 hou2* *jan4*]
1.SG believe 3.SG be good people
'I believe that s/he is a good person.'

b. *Ngo5* *seon3* DP[*keoi5*]
1.SG believe 3.SG
'I believe (in) him/her.'

On the other hand, verbal predicates like *gok3 dak1* 'think, feel' and *hei1 mong6* 'hope' that take COMP clausal complements do not allow nominal OBJs:

(13) a. *Ngo5* *gok3 dak1* S[*keoi5* *hai6 hou2* *jan4*]
1.SG think 3.SG be good people
'I think that s/he is a good person.'

*b. *Ngo5* *gok3 dak1* DP[*keoi5*]
1.SG think 3.SG

(14) a. *Ngo5* *hei1 mong6* S[*keoi5* *hai6 hou2* *jan4*]
1.SG hope 3.SG be good people
'I hope that s/he is a good person.'

*b. *Ngo5* *hei1 mong6* DP[*keoi5*]
1.SG hope 3.SG

These verbs allow only COMP clausal complements. This shows that clausal complements do function differently depending on the matrix predicate in which not all of them can be taken as bearing an OBJ function as proposed by Alsina et al. (1996).

The above analysis of Cantonese as a mixed language is based on the criteria that were used in Dalrymple and Lødrup (2000). However if the distinction between OBJ and COMP languages is attributable to the pronominalization by an NP proform, this may not be a good criterion for Cantonese since Cantonese pronouns refer primarily to animate entities, and are rarely used to replace phrases or sentences that encode facts, propositions and ideas. On this score, one might argue that Cantonese is solely a COMP language. We should note, however, that we cannot pronominalize even clear cases of inanimate OBJ NPs in Cantonese as shown in (15).

(15) a. *Ngo5* *zi1* *ni1 gin6 si6*
1.SG know this matter
'I know this matter.'

??b. *Ngo5* *zi1* *keoi5*
1.SG know 3.SG
'I know it.'

The above two types of predicates show that some clausal complements function as OBJs and others function as COMPs in Cantonese. On this score alone, Cantonese is clearly an OBJ/COMP mixed language.

4.2. Other criteria

In addition to the COMP alternation with NP OBJs that is given as a first parse in establishing a language as an OBJ/COMP mixed language, Dalrymple and Lødrup (2000) have also mentioned others including coordination, passivization, unbounded dependency, and complementation facts. We show in this subsection that even though we have established Cantonese as an OBJ/COMP mixed language on account of the facts of NP OBJ alternation with COMP, these other extra facts do not seem to apply well in Cantonese grammar. This thus raises the issue as to what kind of OBJ/COMP mixed language Cantonese is, if it is at all.[3]

4.2.1 Coordination

Dalrymple and Lødrup (2000) bring up the issue of NP object coordination with a clause, as shown in (16), as a feature of all the three languages, English, German and Swedish that they consider as mixed languages.

(16) Pat remembered the appointment and that it was important to be on time. (Dalrymple and Lødrup 2000)

In this construction, an NP object can be coordinated with a clause which also bears the

[3] We do not rule out the case that the facts of Cantonese might lead one to argue that it is, indeed, a non-mixed language. Lødrup (email communication, June 2001) has hinted at this alternative analysis.

OBJ function. We note here that this kind of construction is very unnatural in Cantonese; in fact this is not a usual way to coordinate two clauses in this language, as shown by the marginally or hedgingly acceptable sentences in (17) and (18).

(17) ? *Ngo5* *ming4 baak6* *nei5 ge3* *gam2 sau6* *tung4 maai4* [*nei5*
1.SG understand 2.SG POSS feeling and 2.SG

ji2 ging1 *zeon6-zo2* *lik6*]
already try-PERF strength
'I understand your feeling and that you have tried your best.'

(18) ? *Ngo5* *do1 ze6* *ngo5* *baa4 baa1* *maa4 maa1* *zung6 jau2* [*di1*
1.SG thank 1.SG father mother and some

bang4 jau2 *jat1 zik6* *gam2* *zi1 ci4* *ngo5*]
friend all the way so support 1.SG
'I thank my parents and that my friends have supported me all the way.'

This is also the case in Mandarin Chinese. Chao (1968: 271) indicates that 'since the expressions must be comparable in structure, it is not good to have a nominal phrase and a clause in coordination'.

In any case, this kind of coordination is not always available even in English, German and Swedish, as observed by Dalrymple and Lødrup's (2000) quotation of the following sentence:

(19) *He proposed [a 20 % reduction for the elderly] and [that the office be moved to the suburbs]. (Emonds 1970:85)

NP object coordination with an OBJ clausal complement is thus not a strong enough evidence in determining the status of a language as an OBJ/COMP mixed language.

4.2.2 Passivization

Another piece of evidence that could distinguish between OBJ and COMP and thus help in deciding the status of a language comes from the facts of LMT, as observed by Dalrymple and Lødrup (2000). Following LMT, SUBJ and OBJ functions are classified as –r, while COMP is classified as +r (proposed by Zaenen and Engdahl 1994). It is then predicted that the –r argument can be realized as SUBJ as in the passive, while the +r argument cannot be realized as SUBJ.

Cantonese, however, does not seem to be in line with this prediction in which both OBJ clausal complements as in (20) and COMP clausal complements in (21) can by no means become the SUBJ of the passivized verbs:

(20) *[*Keoi5 hai6 hou2 jan4*] *bei2 ngo5 seon3*
3.SG be good people BEI 1.SG believe

(21) *[*Keoi5 hai6 hou2 jan4*] *bei2 ngo5 gok3 dak1*
3.SG be good people BEI 1.SG think

In fact, not all -r arguments can become the subject of the passivized verbs. This depends very much on the verb type. Verbs like *gwaa3 zyu6* 'miss', *zung1 ji3* 'like', *zang1* 'hate', and etc. do not have their passive counterparts regardless of the types of complements they take (neither nominal NP OBJ (22) nor OBJ clausal complement (23)):

(22) a. *Ngo5 zung1 ji3 li1 fu3 ngaan5 geng2*
1.SG like this CL glasses
'I like this pair of glasses.'

*b. *Li1 fu3 ngaan5 geng2 bei2 ngo5 zung1 ji3*
this CL glasses BEI 1.SG like

(23) a. *Ngo5 zung1 ji3* [*nei5 daai3 li1 fu3 ngaan5 geng2*]
1.SG like 2.SG wear this CL glasses
'I like that you wear this pair of glasses.'

*b. [*Nei5 daai3 li1 fu3 ngaan5 geng2*] *bei2 ngo5 zung1 ji3*
2.SG wear this CL glasses BEI 1.SG like

As we can see then, Cantonese does not seem to distinguish OBJ clausal complement from COMP clausal complement in this respect.

4.2.3 Unbounded dependency

The facts of unbounded dependency provide yet another parameter for comparing OBJ and COMPs. Dalrymple and Lødrup (2000) observe that most languages (e.g. English, German, and Swedish) allow OBJ arguments to enter into an unbounded dependency, but not COMP. However, it is not uncommon that Cantonese allows topicalized COMP, i.e. a COMP can be identified with a TOPIC. In this case, both OBJ and COMP clausal complements are allowed to enter into an unbounded dependency as in (24) and (25) respectively:

(24)	[*Keoi5*	*hai6*	*hou2*	*jan4*],	*ngo5*	*zi1 dou3*
	3.SG	be	good	people	1.SG	know
	'S/he is a good person, I know.'					

(25)	[*Keoi5*	*hai6*	*hou2*	*jan4*],	*ngo5*	*gok3 dak1*
	3.SG	be	good	people	1.SG	think
	'S/he is a good person, I think.'					

It would therefore seem that, on this score, Cantonese again does not come out as an OBJ/COMP mixed language, as it is hard to distinguish between OBJ and COMP in Cantonese using this piece of evidence.

4.2.4 Complementation of nouns, adjectives and prepositions

A further piece of evidence as observed by Dalrymple and Lødrup (2000) comes from the facts of differences in category complementation. In English, nouns and adjectives (intransitive categories) are expected not to take OBJ clauses but COMP clauses. In contrast, prepositions (transitive category) are predicted to take OBJ clauses. Unlike English, nouns in Cantonese do not take COMP clauses. On the other hand, adjectives are expected to take COMP clauses as in (26):

(26)	*Ngo5*	*hou2*	*hoi1 sam1*	[*gaan1*	*gung1 si1*	*ceng2-zo2*	*ngo5*]
	1.SG	very	happy	CL	company	employ-PREP	1.SG
	'I am so happy that the company has employed me.'						

There is a productive construction which occurs along with an obligatory adjectival modifier *dou3* 'to the extent that' as in (27):

(27)	*Keoi5*	*gui6*	*dou3*	[*deoi3*	*goek3*	*jyun5*	*saai3*]
	3.SG	tired	to the extent	CL	legs	feeble	all

'S/he is so tired that his/her legs are feeble.'

However, the sentential complement does not seem to function as a COMP, but as an ADJUNCT.

There is no common agreement that prepositions form a distinct class of category in Cantonese. If we take the position that coverbs are not prepositions (Bodomo 1999, 2000), this test would no longer be valid. On the other hand, if coverbs are in fact prepositions in Cantonese, then it is expected that prepositions do not take OBJ clauses.

(28)	*Keoi5*	*can3*	[*ngo5 dei6*	*m4*	*wai4 ji3*]	*tau1-zo2*	*bun2*	*syu1*
	3.SG	during	1.PL	NEG	aware	steal-PERF	CL	book

'S/he stole the book when we did not pay attention.'

(29)	*Keoi5*	*soeng2*	*jau4*	[*ngo5*	*zyu2 ci4*	*li1*	*go3*	*wui2*]
	3.SG	want	from	1.SG	chair	DET	CL	meeting

'S/he wants me to chair this meeting.'

(30)	*Keoi5*	*zi1*	*gwan1 jyu1*	[*ngo5 dei6*	*hap6 zok3*	*ge3*	*si6*]
	3.SG	know	about	1.PL	cooperate	POSS	matter

'S/he knows about the matter on which we cooperated.'

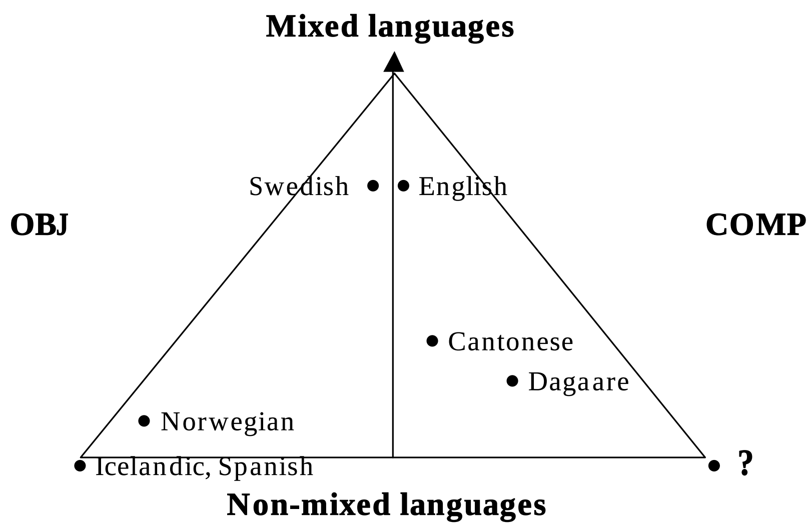
The coverb *can3* in (28) seems to have taken a sentential complement clause *ngo5 dei6 m4 wai4 ji3*, however, it functions as an adjunct in this case. In (29), the coverb *jau6* takes an OBJ *ngo5* instead of the sentential complement *ngo5 zyu2 ci4 li1 go3 wui2*. In (30), the coverb *gwan1 jyu1* takes a nominalized sentence, but not a COMP clause.

We have argued in this section that though Cantonese does not seem to have satisfied most of the criteria to be considered as a mixed language (Dalrymple and Lødrup 2000), it still is a mixed language. It is simply because a distinction between two kinds of clausal complements, i.e. OBJ clausal complement and COMP clausal complement, is crucial in this issue, which we have been able to do for Cantonese with the facts NP OBJ alternation with COMP. Although they are both realized as S, they behave in different ways. One behaves more like an OBJ, while the other is more of a COMP. Therefore, it is not plausible to differentiate the two in terms of phrase structure category in c-structure only as suggested by Alsina, Mohanan and Mohanan (1996), but also in terms of grammatical functions at the level of f-structure. This therefore leaves us with the issue that OBJ/COMP mixed languages may indeed differ in the frequency, productivity and extent to which OBJ and COMP co-occur in a language.

5. Degree of mixedness of a language

In this section, based on the discussion in sections 3 and 4, we speculate on the nature of OBJ/COMP mixed languages. We propose that the extent to which OBJ and COMPs co-occur and alternate may differ from language to language. As we can see from section 3, there is no distinct boundary to differentiate mixed languages from non-mixed languages. We propose, as illustrated below in (31) that a degree of mixedness should be introduced.

(31) Finer gradation of the concept of mixed languages



→ The degree of mixedness from non-mixed languages to elaborately mixed languages

The figure in (31) is a triangle, with an arrow from the base indicating increase is the degree of mixedness, such that the further away from the base a language is the more mixed it gets in terms of OBJ/COMP occurrences. English and Swedish are thus the most mixed. Cantonese seems to be more mixed than Dagaare, a Gur language of West Africa, and Norwegian, with Icelandic and Spanish on the baseline, showing that they are non-mixed languages. These positions are not meant to be precise ones on the diagram but only indicative of language types based on trends in the behaviour of sentential complements in the language.

The triangle has two halves, indicating a division into language types, based on whether the sentential clauses behave more like OBJ or COMPs in the language. English and Swedish are each on a separate half, indicating that though both are OBJ/COMP mixed languages, English is more of a COMP language while Swedish is more of an OBJ language.

To further illustrate degrees of mixedness, let us look at some Dagaare data. It seems to be a mixed language mainly on the score of OBJ NP alternation with clausal complement COMP.

(32) a. *ǹ bɔ̀ngé lá à dɔ́ɔ́*
1SG know TOP DEF man
'I know the man.'

b. *ǹ bɔ̀ngɛ̀ɛ̀ lá ká à dɔ́ɔ́ é lá néng-vèlàá*
1.SG know.PERF TOP COMP DEF man be TOP person-good
'I know that the man is a good person.'

(33) a. *ǹ téɛ́ré ká ò yɛ́lɛ́ vèɛ̀lɛ́ lá*
1.SG think COMP 3.SG matter good TOP
'I think he is good.'

?*b. *ǹ téɛ́ré lá à dɔ́ɔ́*
1.SG think TOP DEF man
*'I think the man.'

Most of the criteria like the facts of coordination, passivization, topicalization of OBJ/COMP etc do not apply well for Dagaare, though we shall not go into the details in this paper. Yet on the strength of NP object alternation with COMP, it may still be regarded as an OBJ/COMP mixed language.

Since we do not expect all languages to meet most or all of these tests, it would be useful to extend the notion of OBJ/COMP mixed language and accept the fact that some languages have a richer manifestation of OBJ/COMP occurrence than others. On the one hand, languages like Cantonese and Dagaare are *marginal OBJ/COMP mixed languages* while others like English and Swedish are *elaborate OBJ/COMP mixed languages*. The claim we make here is that the empirical evidence for NP OBJ alternation with clausal COMP is so strong that even in the case of marginally mixed languages like Cantonese and Dagaare which can be regarded as the weaker case for distinguishing COMP from OBJ, one still has to do it because one cannot ignore the facts of NP OBJ and COMP alternation.

6. Conclusion

This paper has discussed the nature of the grammatical function COMP in the structure of Cantonese. This grammatical function has often been ignored and not given much focus in many works on Cantonese, but we show in this paper that it is one of the salient issues in a discussion of the relatively complex complementation properties of Chinese predicates.

We have argued that despite suggestions from works such as Alsina, Mohanan and Mohanan (1996) for the elimination of COMP as a grammatical function from the LFG taxonomy of functions as it creates duplication and thus redundancy across levels of representation, this grammatical function does actually exist in the structure of Cantonese.

This is because some verbal predicates have NP objects which alternate with clausal complements that are objects, while other verbal predicates that cannot take objects NPs do have clausal complements which are COMPs. These COMPs behave differently from clausal complements that are objects, alternating with Object NPs in many ways. Following Dalrymple and Lødrup (2000) which introduces the concept of mixed languages to describe such languages, we show that Cantonese is a mixed language on this score.

But we have also drawn attention to the fact that though Cantonese is a mixed language, it does not meet many of the tests for mixed languages as suggested by Dalrymple and Lødrup (2000). We have suggested the introduction of a degree of mixedness, such that languages that behave like Cantonese should be treated as marginal OBJ/COMP mixed languages while languages like English that fulfil many of the criteria outlined should be treated as elaborate OBJ/COMP mixed languages.

One issue that has attracted grammarians is the extent to which we should abstract grammatical functions. Like the debate about a taxonomy of argument roles, there is also the debate about whether we should have fewer grammatical functions (coarse-grained approach) or a much wider category of grammatical functions (finer-grained approach). We believe that the recognition of another function, COMP, in the structure of Cantonese allows one to have a better understanding of the nature of sentential complementation in Chinese. Moreover a new member in the taxonomy of LFG functions allows us to model and express typological differences across languages.

On the strengths of these two points, we believe that a finer-grained distinction of grammatical functions is desirable in theories of grammar.

References:

Alsina, A., T. Mohanan, and K. P. Mohanan. 1996a. Responses to the discussion on COMP in LFG. Submission to the LFG list, 6 September 1996.

Alsina, A., T. Mohanan, and K. P. Mohanan. 1996b. Untitled submission to the LFG list. 3 September 1996.

Alsina, A., T. Mohanan, and K. P. Mohanan. 1996c. On COMP. Submission to the LFG list. 1 October 1996.

Bodomo, A. B. 1999. Double Objects and Serial Verbs in Competition: Expressing the Benefactive in Cantonese. Paper read at the Annual Research Forum, Linguistic Society of Hong Kong, December 1999, Chinese University of Hong Kong, Hong Kong.

Bodomo, A. B. 2000. Expressing the Benefactive in Dagaare and Cantonese. Paper read at the West African Language Society Conference, August 2000, Accra, Ghana.

Bresnan, J. 1982a. The passive in Lexical Theory. In Bresnan (ed.), 1982. *Mental representation of Grammatical Relations*. London: The MIT Press. pp.3-86.

Bresnan, J. 1982b. Control and Complementation. In Bresnan (ed.), 1982. *Mental representation of Grammatical Relations*. London: The MIT Press. pp. 282-390.

Bresnan, J. and J. Kanerva. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20.1:1-50.

Bresnan, J. 2001. *Lexical-Functional Syntax*. UK: Blackwells.

Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. California: University of California Press.

Curly, C. 1996. Even More on COMP. Submission to the LFG list. 6 September 1996.

Dalrymple, M. and H. Lødrup. 2000. The Grammatical Functions of Complement Clauses. In Butt and King (ed.), Proceedings of the LFG00 Conference, University of California, Berkeley: CSLI Publications. http://www-csli.stanford.edu/publications/

Li, C. and S. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. California: University of California Press.

Luke K. K., A. B. Bodomo and O. Nancarrow. 2001. The Subject Condition in Cantonese. Workshop on LFG approaches to Chinese, LFG2001.

Lødrup, H. 1996. COMP. Submission to the LFG list. 3 September 1996.

Pan, H and J. Hu. 2001. An Optimality-Theoretic Account of Mandarin Complex Reflexive 'ta -ziji' (s/he -self). Workshop on LFG approaches to Chinese, LFG2001.

Zaenan, A. and E. Engdahl. 1994. Descriptive and theoretical syntax in the lexicon. In Atkins B. T. S. and A. Zampolli (eds.), *Computational Approaches to the Lexicon*, pp. 181-212. Oxford: Oxford University Press.

The Treatment of Tense

Workshop for LFG 2001
University of Hong Kong, Hong Kong

Organized by

Miriam Butt
Universität Konstanz

Table of Contents

THE TREATMENT OF TENSE

Miriam Butt
University of Konstanz

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications
http://csli-publications.stanford.edu/

1 Introduction

This workshop sought to discuss issues with respect to the treatment of tense which have arisen in recent years. By way of an introduction to the workshop, I first layed out the classic approach to the representation of temporal and aspectual information within LFG and then moved on to briefly present some of the more recent developments. In particular, recent work on nominal tense with clausal scope and nonconfigurational tense taken together with the well known problems surrounding the determination of compositional and relational tense, points to the need of a careful reconsideration of the issues involved. My introductory remarks are reproduced in the following sections.

Before delving into the linguistics of tense, however, I would like to thank all of the participants of the workshop for being extremely cooperative and tailoring their contributions exactly as specified by the theme of the workshop. I would also like to thank Henriette de Swart for providing advice and support when it was much needed.

2 The Classic Approach

The classic approach to tense within LFG as practiced in Bresnan 1982, for example, includes the postulation of a TENSE feature at f-structure. The range of its values have never been precisely defined, but they are assumed to include at least PRES(ent) and PAST. As far as I can tell, there is no real consideration of the possible representation of aspectual information such as progressive or perfect. For English, Falk 1984 does encode the morphological form of the participle in terms of a PART feature whose values are PROG or PERF.

The encoding of the morphological form of the participle reflects the existence of periphrastic auxiliary constructions in English. In order to ensure that only the right combination of auxiliary and participle is analyzed or generated, information about the form of the participle must be available at some level of representation. As is well known, auxiliaries in English (and German) select the form of their dependent in a cascading relationship. This kind of binary dependence would seem to be at the root of the crosstheoretic analysis of auxiliaries as raising verbs. As illustrated in (1), auxiliaries subcategorize for a non-thematic subject and an XCOMP in the classic LFG approaches.

(1) a. The driver is turning the lever.

b.

$$\left[\begin{array}{ll}
\text{PRED} & \text{‘be} < \text{XCOMP} > \text{SUBJ’} \\
\text{TENSE} & \text{PRES} \\
\text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘driver’} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \text{DEF} \end{array}\right] \\
\text{XCOMP} & \left[\begin{array}{ll} \text{PRED} & \text{‘turn} < \text{SUBJ, OBJ} > \text{’} \\ \text{SUBJ} & [\ \ ] \\ \text{OBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘lever’} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \text{DEF} \end{array}\right] \end{array}\right]
\end{array}\right]$$

3 Revision of the Classic Approach

Two undesirable side effects of the classic approach are that although auxiliaries are generally grouped with functional elements, the raising analysis treats them as semantically active predicational elements. That is, auxiliaries are treated on a par with modals and other raising verbs without necessarily sharing all of the properties associated with these other classes of verbs. More recent approaches, such as Bresnan 2001 have reconsidered this assumption and taken a different stance. In keeping with their functional status, auxiliaries are situated in functional categories such as I. Furthermore, auxiliaries are not assumed to have a subcategorization frame, but simply contribute features such as TENSE at the level of f-structure. A sample f-structure is shown in (2).

(2)

a. The driver will turn the lever.

b.

$$\left[\begin{array}{ll} \text{PRED} & \text{`turn} < \text{SUBJ}, \text{OBJ} > \text{'} \\ \text{TENSE} & \text{FUT} \\ \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`driver'} \\ \text{CASE} & \text{NOM} \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \text{DEF} \end{array}\right] \\ \text{OBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`lever'} \\ \text{CASE} & \text{ACC} \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \text{DEF} \end{array}\right] \end{array}\right]$$

With respect to crosslinguistic analysis, the analysis represented by (2) has considerable advantages, as it abstracts away from the surface form in a manner that is consistent with the general analytical endeavour of LFG (see Butt et al. 1996 on the advantages for machine translation applications). For example, it is now possible to abstract away from the difference between periphrastic and morphologically encoded tense.

This revision of the classical approach thus clearly represents a step in the right direction. However, there is still no provision made for the representation of aspectual information on a par with temporal information.

Another problem which remains to be solved is the issue of local dependencies between auxiliaries and participles in periphrastic constructions. Based on a suggestion by Ron Kaplan, Butt et al. 1996 propose to use another level of representation at which such local well-formedness conditions can be stated. This level was dubbed m-structure because it was taken to encode those features which represented local morphological well-formedness conditions with no useful functional or semantic role.[1] The m-structure which ensures the correct local dependencies for the example in (2) is shown in (3).

(3)

$$\text{m-structure} \left[\begin{array}{ll} \text{AUX} & + \\ \text{FIN} & + \\ \text{DEP} & \left[\begin{array}{ll} \text{FIN} & - \\ \text{VFORM} & \text{BASE} \end{array}\right] \end{array}\right]$$

[1]For example, strong/weak adjective agreement in German has no conceivable functional or semantic use. The agreement system between determiners, adjectives and nouns in German is so arcane as to be incomprehensible (and indeed represents a major problem for German language learners). However, the checking of agreement is needed in order to ensure well-formedness. As such, these agreement features are a good candidate for m-structure.

With the problem of local dependencies solved and the contribution of the auxiliaries whittled down to the contribution of a temporal feature, another possibility for the treatment of periphrastic constructions emerges. For examples like *The driver will have turned the lever.*, the *will* taken together with the *have* could be encoded as the composite tense FUTPERF.

The implementation of this approach is described in Butt et al. 1996. It has the advantage that a higher level of analysis with respect to tense is done at the level of f-structure since the value of tense is calculated based on all of the clausal information available. However, the calculation of composite tenses is hard work in that the lexical entries for items like *will* must encode a large number of disjunctions, as shown in (4) for the German auxiliary *wird*. Note that the disjunctions governing the passive have been left out, (4) shows only part of the entry.

(4)	wird	AUX	$(\uparrow$SUBJ CASE$) =$ NOM	`subject agreement`
			$(\uparrow$SUBJ NUM$) =$ SG	
			$(\mu$ M* AUX$) = +$	`this is an auxiliary`
			$\{$ $(\mu$ M* DEP VFORM$) =_c$ BASE	`Option 1:  if bare infinitive`
			$(\mu$ M* DEP DEP VFORM$) \neq$ PERFP	`no perf.  particple`
			$(\uparrow$PASSIVE$) \neq +$	`and no passive`
			"simple future: wird drehen	
			$(\uparrow$TENSE$) =$ FUT	`then simple future`
			$\mid$	`Option 2`
			$(\mu$ M* DEP VFORM$) =_c$ BASE	`if bare infinitive`
			$(\mu$ M* DEP DEP VFORM$) =_c$ PERFP	`and a perf.  participle`
			$(\uparrow$PASSIVE$) \neq +$	`and no passive`
			"future perfect: wird gedreht haben"	
			$(\uparrow$TENSE$) =$ FUTPERF $\}$	`then future perfect`

The computation of these disjunctions turns out to be fairly costly. In addition, it is also not clear whether the calculation of composite tenses always arrives at the most desirable solution. Within the ParGram project (Butt, King, Niño, Segond 1999), this approach was adopted at first, but then abandoned as a consideration of a large amount of data within the large-scale grammar development project forced a realization that the analyses were often either inconsistent or impossible to compute correctly based on the clausal information available at f-structure. (see section 4).

The current solution within the ParGram project is to simply register the morphosyntactic information available in terms of non-composite features such as PERF or FUT, as shown in (5).

(5) a. The driver will have turned the lever.

b.

$$\left[\begin{array}{ll}
\text{PRED} & \text{`turn} < \text{SUBJ, OBJ} > \text{'} \\
\text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`driver'} \\ \text{CASE} & \text{NOM} \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \text{DEF} \end{array}\right] \\
\text{OBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`lever'} \\ \text{CASE} & \text{ACC} \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \text{DEF} \end{array}\right] \\
\text{TNS-ASP} & \left[\begin{array}{ll} \text{TENSE} & \text{FUT} \\ \text{ASP} & \text{PERF} \end{array}\right]
\end{array}\right]$$

4 Further Problematic Data

Some of the problems encountered as part of the ParGram effort were well-known problems such as the interpretation of relative tense. In (6), for example, morphological past tense must actually be interpreted semantically as future due to the fact that the embedded (or relative clause) must be interpreted *relative* to the matrix clause (the examples are taken from Crouch 1993).

(6) a. If I smile when I get out, then the interview **went** well.

b. By 1998, everybody will have at least one friend who **died** from AIDS.

The calculation of this type of relative tense is in principle possible at the level of f-structure, however, given the fact that the computational effort is significant, the question remains whether the strategy of trying to get the tense value exactly right at the level of f-structure is a feasible one.

Furthermore, the well-known possibility of polysemy with respect to phenomena such as the English perfect would seem to provide an insurmountable challenge for a complete analysis of tense and aspect at the level of f-structure. Some examples (based on Kiparsky 2001) to illustrate this problem are shown in (7). The possible semantic readings of the morphosyntactic perfect are given in bold face at the beginning of the sentence.

(7) a. **Existential:** Fred has visited Paris several times.

b. **Universal:** I have known him since 1960.

c. **Resultative:** The police have probably caught the subject by now.

d. **Recent Past:** Musharraf has been sworn in as president. (hot news reading)

Again, the ParGram solution of simply registering the morphosyntactic information as ASP PERF would seem to be the most pragmatic idea. Under this approach, the relevant morphosyntactic information is simply registered at f-structure and then passed on to semantics for further interpretation.

However, for the following questions arise: If the level of f-structure can only effectively be used for a simple registration of morphosyntactic information, then why implicate f-structure at all? Why not register the local morphosyntactic information at m-structure and have semantics be able to access that information?

A further argument for this view would seem to come from Nordlinger and Bresnan's 1996 study of Wambaya in which future tense must be seen as a composite out of the features FUTURE and UNCERTAIN. The morphology determining the composite tense information is distributed all over the clause, as shown in (8). Nordlinger and Bresnan propose to use three features for a decompositional analysis of tense: [±PAST], [±FUTURE], [±UNCERTAIN]. A Wambaya example together with the proposed analysis is shown in (8).

(8) a. Ngaj-ba ny-ng-u
see-Uncertain 2Sg.A-1.O-Fut
'You will see me.' (Wambaya)

b.

$$\left[\begin{array}{ll} \text{FUTURE} & \left[\begin{array}{ll} \text{PAST} & - \\ \text{FUTURE} & + \\ \text{UNCERTAIN} & + \end{array}\right] \end{array}\right]$$

Again, the calculation of the tense feature FUTURE involves the careful piecing together of information from various parts of the clause: a process whose complexity should not be underestimated. The question of whether f-structure should simply be left out of the equation thus rears its head.

5 Workshop Contributions

The problems and considerations summarized above were taken as a provokative starting point for the discussion engendered by the workshop. None of the workshop participants agreed with the conclusion that f-structure has no real role to play with respect to the determination of tense. However, the discussion engendered by the workshop produced a more refined idea of the role of f-structure and the type of information that usefully could be represented at that level.

The workshop began with a contribution by Louisa Sadler and Rachel Nordlinger, who continued their investigations into the expression of nominal tense by presenting data from languages in which nominal tense is used to locate the nominal temporally independent of the clausal specification. While they have worked out an account of nominal tense with nominal scope within LFG (Nordlinger and Sadler 2000), the phenomena of nominal tense with nominal scope remains to be tackled within a formal approach. Their contribution to this workshop can thus be viewed as presenting new and potentially problematic data that any formal theory of grammar must take on eventually (cf. Nordlinger and Sadler's 2000 discussion with respect to Bender and Sag's 2000 HPSG analysis).

Christoph Schwarze's contribution considered the nature of features at f-structure and concluded on the basis of Romance data that the traditional TENSE feature must indeed still be given a place at f-structure.

Sheila Glasbey took on the same issue from a semanticist's point of view. Under the assumption that f-structure is the primary input for semantic structure, Glasbey also concluded that the representation of temporal and aspectual information makes sense at the level of f-structure. However, Glasbey provides a much more detailed theory in terms of the crosslinguistically relevant features that should be encoded at f-structure. Glasbey also formulated proposals for the interaction of lexically encoded underspecified information with clausal information. This lexically underspecified information is specified as part of the clausal analysis. The features encoded at f-structure are based on insights from the semantic literature and thus provide just the right information for a temporal and aspectual analysis of continuous discourse.

References

Bender, Emily M., and Ivan A. Sag. 2000. Incorporating Contracted Auxiliaries in English. In *Grammatical Interfaces in HPSG*, ed. Ronnie Cann, Claire Grover, and Philip Miller. 17–32. Stanford, California: CSLI Publications.

Bresnan, Joan (ed.). 1982. *The Mental Representation of Grammatical Relations.* Cambridge, Massachusetts: MIT Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax.* Oxford: Blackwell.

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño, and Frédérique Segond. 1999. *A Grammar Writer's Cookbook.* Stanford, California: CSLI Publications.

Butt, Miriam, María-Eugenia Niño, and Frédérique Segond. 1996. Multilingual Processing of Auxiliaries in LFG. In *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference*, ed. D. Gibbon. 111–122. Berlin: Mouton De Gruyter. Universität Bielefeld, 7 - 9 October 1996.

Crouch, Richard. 1993. Tense in Simple Conditionals. In *Proceedings of the 9th Amsterdam Colloquium.*

Falk, Yehuda N. 1984. The English auxiliary system: A Lexical-Functional analysis. *Language* 60(3):483–509.

Kiparsky, Paul. 2001. Event Structure and the Perfect. Unpublished Ms., Stanford University.

Nordlinger, Rachel, and Joan Bresnan. 1996. Non-configurational tense in Wambaya. In *On-line Proceedings of the First LFG Conference, Rank Xerox, Grenoble, August 26–28, 1996*, ed. Miriam Butt and Tracy Holloway King.

Nordlinger, Rachel, and Louisa Sadler. 2000. Tense as a Nominal Category. In *On-Line Proceedings of the LFG00 Conference*, ed. Miriam Butt and Tracy Holloway King. 197–214. Stanford, California: CSLI Publications. University of California, Berkeley.

NOMINAL TENSE WITH NOMINAL SCOPE: A PRELIMINARY SKETCH

Louisa Sadler Rachel Nordlinger

louisa@essex.ac.uk racheln@unimelb.edu.au

Proceedings of the LFG01 Conference

University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)

2001

CSLI Publications

http://csli-publications.stanford.edu

Abstract

The morphosyntactic categories of tense, aspect and mood are traditionally considered to be properties of verbs. However, the morphological expression of these categories within the nominal system is attested across a range of languages. Drawing on data from a number of languages, we provide a preliminary sketch of the phenomenon of *nominal tense with nominal scope* - that is, cases in which nominal tense serves to temporally locate the nominal independent of the temporal specification of a governing predicate.

1 Types of Nominal Tense

Although the categories of tense, aspect and (to a lesser extent) mood are traditionally considered to be properties of verbs, the morphological expression of tense, aspect and/or mood (henceforth TAM) on nominals is attested across a range of languages. We distinguish two major subcases of nominal tense, that is, of cases in which TAM marking occurs on a nominal or other constituent of NP/DP.[1]

In some cases, a dependent nominal or nominal phrase (of whatever grammatical function) bears some TAM marking which serves to temporally, aspectually or modally specify the clausal predicate which is itself distinct from the nominal argument. This is illustrated by examples (1)-(2)) in which the case marking of the dependent NPs changes to reflect the tense (future vs. nonfuture) of the clausal predicate:

(1)	*Ngamari-lu*	*ngunytyi-ka*	*ngali-nha*	*mangarni-marru-nga-nha*	*kathi-nha.*
	mother-ERG	give-PAST	we.DU-ACC	bone-having-GEN-ACC	meat-ACC.
	Mother gave us the doctor's meat. (Pitta Pitta (Australia), Blake 1987:60, ex. 4.12)				
(2)	*Ngamari-ngu*	*ngunytyi*	*ngali-ku*	*mangarni-marru-nga-ku*	*kathi-ku.*
	mother-NOM.FUT	give	we.DU-ACC.FUT	bone-having-GEN-ACC.FUT	meat-ACC.FUT.
	Mother will give us the doctor's meat. (*op. cit.*, ex. 4.14)				

We refer to this as nominal tense with clausal scope (NT-CS for short). In Nordlinger and Sadler (2000) we discuss several cases of NT-CS, demonstrating that this phenomenon is well attested and showing how it can be simply analysed in LFG using inside out statements. Further work is needed to draw up a typology of exponence for NT-CS, and on the description of the interaction at the clausal level of nominally and verbally expressed clausal TAM features. We will have nothing more to say in the present paper about this phenomenon (but see Nordlinger and Sadler (in preparation)).

In other cases, the nominal tense has scope only over the nominal constituent itself, that is, is interpreted with respect to the nominal phrase (rather than the verb phrase). This paper is concerned with the phenomenon of nominal tense with nominal scope (NT-NS for short). Here two subcases are conceptually distinct. In what are sometimes called nominal sentences, the nominal itself is not an argument of a verbal predicate but serves as the main predicate of the proposition, without showing any signs of undergoing morphological derivation to form a verb (see, for example, 4.3 below). When such a nominal (or noun phrase) is tensed, we refer to this as nominal tense on nominal predicates (NT-NP for short). In many languages, subordinate 'clauses' often involve the use of predicates which are transparently nominal in category but which bear TAM marking, and this is another situation in which NT-NP is attested (Koptjevskaya-Tamm 1993). Although we exemplify both these sorts of construction briefly below, for reasons of space we will concentrate mainly in this paper

[1]For ease of exposition, we use the term **nominal tense** throughout to refer to TAM marking on nominal elements. Throughout this paper we interpret nominal tense to mean nominal tense with nominal scope, in the sense to be defined below.

on the second subcase. Here tense is marked on a *dependent* [2] nominal and serves to temporally locate that nominal *independent of the temporal specification of the governing predicate.* We can refer to this as nominal tense on a nominal dependent (NT-ND for short). Most of the data in this paper exemplifies this phenomenon. The distinctions between cases of nominal tense outlined above can be summarised in the the following diagram:

(3) nominal tense

clausal scope[NT-CS] nominal scope[NT-NS]

nominal predicate[NT-NP] nominal dependent[NT-ND]

The phenomenon of NT-NS raises a number of interesting questions. **Morphosyntactically**, it seems that we need to accommodate TAM features associated with *nominal* morphological and syntactic structures. This runs counter to the view that features are strictly typed into "nominal" and "verbal" features. The data shows that morphological tense on nominals is syntactically active. **Semantically**, quantificational and referential approaches to the semantics make different predictions as to the temporal location of nominal arguments: in quantificational approaches the temporal location associated with the event has scope also over the nominal argument while referential approaches (e.g. (Enc 1986)) allow in principle for events and arguments to be independently temporally located (for an insightful discussion of the temporal location of nominal arguments, see the dissertation by (Tonhauser 2000)). The data supports this latter view and also illustrates a full rage of referential, deictic, anaphoric and discourse related interpretations.

As noted above, the existence of TAM marking on nominals has received very little attention to date. The present paper is basically descriptive in its aim, constituting a preliminary attempt to provide some sort of overview of the phenomenon of NT-NS by bringing together an illustrative (but not exhaustive) set of data.[3] We hope, in presenting this brief overview, to show that this phenomenon is much less marginal than the paucity of theoretical discussion in the literature might lead one to expect, and to provide the impetus for future theoretical work on the issue.

2 Nominal Tense Affixes in Tariana

Tariana is an Arawak language from north-west Amazonia, Brazil.[4] It is a polysynthetic language which combines head marking and dependent marking morphology. Nouns and verbs are heavily inflected, with both prefixing and suffixing, and upwards of 15 possible structural 'positions' may be identified in each. In Tariana, nouns can be inflected for either past or future tense, and the tense marking is interpreted with respect to the nominal predicate itself (the examples that we discuss are all cases of NT-ND). The occurrence of tense morphology on nominals is very widespread, indeed Aikhenvald reports that around 40% of nouns in texts are tense inflected. The nominal tense system is much simpler than that on verbs (which also involves evidentiality distinctions), and most of the forms are quite distinct from their verbal counterparts.

There is a single form for nominal future tense, -*pena*, which denotes that a predicate holds in the future, for

[2] As far as we can judge from the data we have seen so far, it is clear that nominal tense may occur on nominal arguments but it remains an open question whether it occurs also on nominal adjuncts, though we expect that it does.

[3] We thank members of the LINGTYP list for producing many useful leads and in particular Alexandra (Sasha) Aikhenvald, Brent Galloway, Bill Lewis and Tom Payne for access to unpublished data.

[4] The data provided is courtesy of Sasha Aikhenvald, and is taken from her forthcoming grammar *The Tariana language of Northwest Amazonia*, CUP (Aikhenvald's numbering given in brackets throughout this section)

example: *wa-tʃimaɾi-pena* (1PL-son.in.law-NOM.FUT) 'our future son-in-law', *pi-ya-dapana-pena* (2SG-POSS-house-NOM.FUT) 'your future house' and (4):[5]

(4) *kayu-maka hĩ waɾipeɾe unyane-pena di-kakwa-pidana*
so-AFF DEM:ANIM Walipere flood-NOM.FUT 3sg.NF-plan-REM.P.REP
Thus Walipere was planning the future flood. (ex. 9.4)

Nominal past tense has three forms: *miki-ɾi* for masculine singular nouns, -*miki-ɾu* for feminine singular nouns, and -*miki* for plural nouns. -*miki-* is used to refer to a previous state of the head noun (e.g. English 'ex-'). It is used more with animates than inanimates, but possible with both, e.g. *correio-miki-ɾi* (post office-PST-NF) 'old/former post office'; *du-sa-do-miki-ɾu* (3SG.NF-spouse-FEM-NOM.PST-FEM) 'his late spouse' and (5):

(5) *thepi di-maɾe-pidena eta-miki-ɾi-nuku*
to.water 3SG.NF-throw.CAUS-REM.P.REP eagle-NOM.PST-NF-TOP.NON.A/S
He threw the remains of the eagle (lit. what used to be the eagle) into water. (ex. 9.15)

Clausal (that is, verbal) tense is marked in Tariana by floating tense/evidentiality clitics which attach to any focussed constituent in a clause. Naturally, it is possible for these clitics to attach to nominals, along with 'real' nominal tense as illustrated above. The clausal clitics in the following examples are glossed PRES.NONVIS 'present nonvisible' and PRES.VIS 'present visible'.

(6) *kayu-maka diha nawiki-nha ñamu na-nite*
so-AFF he person-PAUS evil.spirit 3PL.say-TOP.ADV+CL:ANIM
nawiki-miki-ɾi-mha
person-NOM.PST-NF-PRES.NONVIS
So this man called evil spirit ñamu, he is the one who used to be a person (lit. he is an ex-person) (ex. 9.10)

(7) *pi-ya-dapana-pena-naka*
2SG-POSS-house-NOM.FUT-PRES.VIS
This is your future house (I can see it)

(8) *pi-ya-dapana-miki-ɾi-naka*
2SG-POSS-house-NOM.PST-NF-PRES.VIS
This is what used to be your house (I can see it)

In the following example, a single nominalised form is marked for tense three times: a past verbform is nominalised, takes nominal past marking and then a clausal tense clitic:

(9) *yatu ka-pusuku-kaɾi-miki-ɾi-mha diha*
snuff REL-mix-PST.REL.NF-NOM.PST-NF-PRES.NONVIS he
He is the one who used to mix up snuff a long time ago (and he's not doing it anymore). (ex. 9.20)

-*kaɾi* refers to the fact that he used to mix the snuff before the moment of speech, -*miki-ɾi* refers to the fact that he has stopped doing so, and -*mha* is the clausal (present-nonvisual) tense/evidentiality marker.

[5] Non-obvious abbreviations in these examples include: AFF 'affix', CL.ANIM 'animate classifier', DEM.ANIM 'demonstrative animate', NF 'non-feminine', PAUS 'pausal', PRES.VIS 'present visible', REL 'relative', REM.P.REP 'remote past reported', TOP.ADV 'topic advancement', TOP.NON.A/S 'topical non-subject clitic'.

3 Category Neutral Tense Affixes in Halkomelem

Our second example of nominal tense comes from Halkomelem, a Salish language spoken on the Northwest coast of North America (Burton 1997, Gerdts 1988, Galloway 1993). This language is predicate-initial and head-marking. As in Tariana, nominals can be inflected with one of two tense markers: future tense, expressing 'will be', and past tense encoding meanings such as 'former, -ex, late (dead)'. In contrast to Tariana, however, the same set of affixes mark tense on nominals and on verbal predicates, a point to which we return below. The following past tense examples are from Burton (1997:67), who discusses the various interpretations available for past tense nominals:[6]

(10) *tel mál*
my father

(11) *tel má:l-elh*
my father-PST
my late father

(12) *te sqwemá:y*
the dog

(13) *te sqwemá:y-elh*
the dog-PST
the dead dog

The 'deceased' reading exemplified above occurs only with animate nouns. When the noun refers to a non-cancellable (or 'lifetime') property such as 'father' or 'dog' (i.e. one cannot cease to be a father or a dog without ceasing to exist) then past tense animate nouns always have the deceased reading. When the noun refers to a cancellable property, however, Burton notes that an alternative reading of 'former, -ex' is possible (Burton 1997:74):

(14) *stó:les-elh*
wife-PST
dead wife, ex-wife

(15) *siyó:ye-lh*
friend-PST
dead friend, former friend

When the past tense marker is used with a possessed inanimate noun, it marks the fact that the possession relation was in the past, or that the possessed item has been destroyed (Burton 1997:67-68):

(16) *tel xeltel-elh*
my pencil-PST
my former pencil, used to be my pencil,
my destroyed pencil

(17) *tel pukw-elh*
my book-PST
my former book, used to be my book,
my destroyed book

In (18) the suffix *-elh* is attached to a pre-verbal auxiliary and marks clausal past tense, thus illustrating the fact that the same tense affix can attach indiscriminately to nouns and verbs. Clausal tense marking is not obligatory in Halkomelem; it is possible for the verb to remain uninflected for tense, in which case the tense of the clause is determined by contextual and pragmatic considerations (Burton 1997:68), as exemplified in (28) below.

(18) *i-lh imex tel sí :le*
AUX-PST walk my grandfather
My grandfather walked.

[6] For simplicity, these Halkomelem sentences are given in the established practical orthography. For information about the pronunciation of Halkomelem, and the IPA equivalent of each grapheme see Galloway (1993). Some of the interlinear glosses have also been simplified for ease of exposition. Again, the reader is referred to Galloway (1993) for further details.

A number of influential papers have argued that the Salish languages lack a distinction between the lexical categories, including the basic distinction between noun and verb (see e.g. (Jelinek and Demers 1994, Jelinek 1996)). In essence, on this view, the differentiation between "propositional" and "referring" syntactic categories comes about solely through the functional "shell": basically, every lexical item projects an IP, which can itself be selected by a D, to form a DP. The most persuasive evidence for this analysis comes from inflectional morphology. For example, in St'át'imcets any open class element can inflect with person/number subject markers to form clauses, as shown below:

(19) *qwatsáts-kacw*
leave-2SG.SUBJ
You left/leave

(20) *smúlats-kacw*
woman-2SG.SUBJ
You are a woman

(21) *xzúm-lhkacw*
big-2SG.SUBJ
You are big (Demirdache and Matthewson 1995:81)

and the defi nitepast and future tense particles encliticize to the main predicate (the bracketting shown is that of (Demirdache and Matthewson 1995)).

(22) *[qwatsáts-ø tu7] [kw-s Gertie]*
leave-3ABS DEF.PAST DET-NOM Gertie
Gertie left

(23) *[qwatsáts-ø kelh] [kw-s Gertie]*
leave-3ABS FUT DET-NOM Gertie
Gertie will leave

(24) *[plísmen-ø tu7] [kw-s Bill]*
policeman-3ABS DEF.PAST DET-NOM Bill
Bill was a policeman

(25) *[plísmen-ø kelh] [kw-s Bill]*
policeman-3ABS FUT DET-NOM Bill
Bill will be a policeman

(26) *[xzum-ø tu7] [ti s-géw'p-a]*
big-3ABS DEF.PAST DET NOM-meet-DET
The meeting was big

(27) *[xzum-ø kelh] [ti s-géw'p-a]*
big-3ABS FUT DET NOM-meet-DET
The meeting will be big (Demirdache and Matthewson 1995:81-82)

On the category neutral view of Salish languages, there are no bare lexical projections, that is, no NPs and VPs, only IPs and DPs. This is naturally of importance to the phenomenon under discussion here since on this view (13) might be analysed as $[_{DP}$ the $[_{IP}$ PRO was dog $]]$ "the one which was a dog", with tense occurring not nominally but on the main predicate of the IP.

However, Demirdache and Matthewson (1995) provide evidence that the syntax of St'át'imcets requires reference to "bare" lexical projections (the categories AP and NP), and therefore evidence against the category neutral view. First they show that there is evidence for a head-final relative clause in which the head must be an uninflected noun (that is, having neither a determiner nor subject person/number inflections). They then show that the class of complex predicates in the language are predicate nominals (NPs), in which an AP modifies an NP in the structure $[_{NP}$ AP NP $]$. This position is supported for Halkomelem by Galloway (1993) who argues on morphological and syntactic grounds for a distinction bewteen nouns and verbs in Halkomelem: for example, affixes encoding possession and diminutives are possible only with nouns (verbs must first be nominalised), while only verbs may be inflected with subject and object pronominal affixes and valency-changing morphs (pp 238, 371-2). Accepting the force of these arguments against the category neutral view of Salish languages, we suggest that the Halkomelem examples discussed in this section are indeed cases of tense affixes occurring on nominals with nominal scope.

Despite the fact that they are encoded with the same morphological marker, clausal tense and nominal tense are completely independent categories in Halkomelem, and can vary independently of each other. In (28) and (29) a past-tense inflected nominal (with the deceased reading) co-occurs with clausal past time reference (non-overt in (28) and overt in (29)). But in (30) a past-tense nominal occurs in a future tense clause, indicating that nominal and clausal tense marking are distinct. Such examples establish that tense marking on nominals is semantically distinct from clausal tense, otherwise we would expect such cases of conflict in tense values to result in ungrammaticality.

(28) *kw'étlexwes tel má:l-elh te sqwemá:y*
see my father-PST the dog
My (late) father saw the dog. (Burton 1997:68)

(29) *éwe-lh kw'étslexw the-l sí:lá:-lh*
NEG.be-PST see the(f)-my grandparent-PST
He didn't see his late grandmother. (Brent Galloway, p.c.)

(30) *El-'éliyemet-tsel-cha the-l sí:lá:-lh*
RDP-dream.about-1SG.SUBJ-FUT the(f)-my grandparent-PST
I'll be dreaming about my late grandmother. (Brent Galloway, p.c.)

To summarize, there is a single past tense morpheme for both verbs and nouns. With verbs and auxiliaries, it places the action in the past. With nouns, it identifies as past the time as which the referent of the NP had the property denoted by the noun, or, in the case of a possessive construction, the time at which the possessive relation held.

The role of future tense marking on nominals is rather different. In combination with the future suffix *-cha*, also used to mark future tense with verbs (see (30) above), the noun functions as a (future) stative predicate. In these cases, the tensed nominal constitutes the clausal predicate, and thus what we have here is a case of NT-NP.[7] The suffix *-s* in (32) functions to nominalize the following phrase so that it can be relativized.

(31) *Swíyeqe-cha*
man-FUT
It will be a man.

(32) *Swíyeqe-cha kw'-a'-s hákw'eles*
man-FUT the-your-NMZR remember
It will be a man that you remember.

[7]These examples are from Brent Galloway, p.c.

4 Nominal Tense and Possession in Hixkaryana

We turn now to our third example language, the Carib language Hixkaryana, spoken in Northern Brazil. This is a mixed ergative, head marking and polysynthetic language, again with very rich verbal morphology. The basic word order is OVS.

Hixkaryana illustrates a rather different situation in which tense marking in nominals is inextricably linked with possession. The tense suffixes are distinct from those used with verbal stems. The language has a set of nominal suffixes which express present, past and remote past possession, and depossession.[8] The choice of basic possessive suffix (expressing present possession) is lexically conditioned, with most nouns taking *-ɾɨ*, *-nɨ*, and other forms being *-tɨ*, *-tʃe*, *-ø*. The depossession suffix is *-nano*. The past possessive suffixes are *-thɨɾɨ(-tʃhɨɾɨ)*, and *-tho (-tʃho)* — the forms in parenthesis being phonologically conditioned allomorphs. Remote past possession is indicated by the suffixes *-nhɨɾɨ* and *-nho*. Of these possessive suffixes, *-tho* and *-nho* occur with first person, first person exclusive and third person (with preceding NP) prefixes, and *-thɨɾɨ* and *-nhɨɾɨ* with the remaining person-marking prefixes.[9] The following examples are from Derbyshire (1979:98-99).

(33) *ɾo-kanawa-ɾɨ*
1-canoe-POSSD
my canoe

(34) *ɾo-kanawa-tho*
1-canoe-POSSD.PST
my former canoe

(35) *ow-ot-tɨ* *(owotɨ)*
2-meat-POSSD
your meat

(36) *ow-wo-tɨ-thɨɾɨ*
2-meat-POSSD-PST
that which was your meat

(37) *ɨ-he-tʃe*
3-wife-POSSD
his wife

(38) *ɨ-he-tʃe-nhɨɾɨ*
3-wife-POSSD-REM
his former wife

(39) *ɾo-katxho-ø*
1-things-POSSD
my things

(40) *ɨ-katxho-ø-thɨɾɨ*
3-things-POSSD-PST
his old things

There is some evidence of the use of the past possessed marker on non-possessed words. The forms *-nhɨɾɨ* and *-nho* also inflect non-possessed words as in *waha-nho* "one who had been a killer", and *toto-tho-nhɨɾɨ* (human-DEV-REM.PST) "one who had been a human being" (Derbyshire 1979:99).

Past and remote possession marking is also found on deverbal nominals, where the possessor morphology codes one of the core arguments (either subject or object), and is very often obligatory. The following action and result nominalizations illustrate:

(41) *ro-to-thɨɾɨ-nhɨɾɨ*
1-go-PST-REM
my going long ago (Derbyshire 1979:99)

[8] According to Derbyshire (1999) the depossession suffix appears (although rarely) on inalienably possessed nouns to indicate more general reference.

[9] The PST forms *replace* the POSSD form *ɾi* but follow the other POSSD forms. In the examples in this section, some changes have been made in the abbreviations used in the examples taken from (Derbyshire 1979) in order to increase consistency in glossing across the languages discussed. Non-obvious abbreviations in these examples include DEV 'devalued', REM 'remote'.

(42) *e-karyma-tʃhɨrɨ-nhɨrɨ*
3-tell-PST-REM
its telling long ago (Derbyshire 1979:99)

(43) *o-n-menho– thɨrɨ-nhɨrɨ*
2-OBJ NOMZ-write-PST-REM
the thing you wrote long ago' (Derbyshire 1979:99)

Derbyshire notes that the remote past may be used in these cases to distinguish between earlier and later events, even if the early one is not actually remote from the utterance time. Although the details provided are quite sparse, Derbyshire (1979) suggests that the interpretation of *-nhɨrɨ* appears to be dependent on the choice of nominalising suffi x, with *-nye* he interprets it as meaning 'past', with *-saho* he interprets it as meaning 'remote past' and with *-hɨnɨ* he interprets it as meaning 'past':

(44) *oy-okaryma-nye-nhɨrɨ*
2-tell-DOER NOMZ-PST
the one who told about you (Derbyshire 1979:99)

(45) *e-karyma-xaho-nhɨrɨ*
3-tell-OBJ PST ACT-REMP
thing that was told long ago (Derbyshire 1979:99)

(46) *r-okaryma-nye-hɨnɨ-nhɨrɨ*
1-tell-DOER NOMZ-NEG-PST
not the one who told about me (Derbyshire 1979:99)

In Hixkaryana, nominalized forms very frequently serve as clausal predicates; in fact, Derbyshire (1979) observes that nominalization is the dominant form of (clausal) subordination. Derivational suffi xes attached to nominalized forms are used to derive a variety of adverbial (pseudo-)clauses (e.g 'during' clauses, purposes clauses, 'until' clauses). For example, temporal adverbial (subordinate) clauses denoting simultaneous action involve suffi xation of *-toko* to a nominalized verbstem. The fact that the resultant wordform is nominal, rather than verbal, is clearly indicated by the morphology — the form is possessed, with the prefi x referring to the subject of the action with intransitive stems and the object of the action with transitive stems:

(47) *o-horohɨ-nɨ-toko*
2-stop-ACT NOMZ-SIMUL
when you stop (Derbyshire 1979: 177)

(48) *r-aryma-nɨ-toko*
1-throw-ACT NOMZ-SIMUL
when my being thrown (Derbyshire 1979: 178)

(49) and (50) illustrate the difference between (verbal) main clauses and (nominal) subordinate clauses, with the possessed/tense forms occuring in the nominalizations.

(49) *karyhe toye (ø-to-ye) kamara*
quickly went (3S-go-REM) jaguar
The jaguar went quickly

(50) *karyhe kamara tothɨrɨ (=ø-to-thɨrɨ)*
quickly jaguar (3-go-POSS.PST)
the quickly going of the jaguar (Derbyshire 1979:23)

Finally, before ending this section, we should note that the marking of a past/present distinction alongside possession is not limited to Hixkaryana among the Carib languages. In his overview survey of Carib, (Derbyshire 1999) notes that the following other Carib languages also mark form/past possession morphologically: Apalai, Wai Wai, Carib, Dekwana, Trio, Wayana, Makushi.

5 Nominal Tense and Definiteness in Somali

In a series of recent papers, Lecarme provides extensive discussion of the phenomenon of nominal tense in Somali (see (Lecarme 1996, Lecarme 1999)), and we draw extensively on this work in discussing Somali in this section. Somali is a Cushitic language, and is spoken in Somalia. In this language, definite determiners (which are nominal affixes) encode a past/nonpast distinction on the heads of noun phrases. Affixal determiners involve an initial consonant (*-k* and its allomorphs with masculine stems and *-t* and its allomorphs with feminine stems — these forms are subject to sandhi rules), followed by *-ii*, *-u* or *-a*: *-ii* forms are +past (case neutral) and *-u*, *-a* forms are -past, nominative and non-nominative respectively.[10]

(51)

initial C		[-past]	[+past]
-k/t-	[+nom]	-u	-ii
-k/t-	[-nom]	-a	-ii

(Lecarme 1999: 335)

The tensed determiners shown in the table above are in paradigmatic opposition with a separate deictic system involving near/far demonstratives, which do not have a temporal interpretation:

(52)

	proximal	remote
-k/t-	-án 'this'	-áas 'that'
-k/t-	éer 'that (far away)'	-óo 'that (very far away)'

(based on Lecarme 1999: 335)

The following is an example of nominal tense on the head noun within a complex noun phrase (the affairs could be completed, or referred to earlier in the discourse).

(53) *arrimí -hii Gúddi-ga Sare ee Tawrád-du*
affairs-DET.M.PST Committee-DET.M upper and Revolution-DET.F.NOM
the affairs of the Supreme Council of the Revolution (Lecarme 1999:235)

Somali also has a set of possessive determiners which are suffixed to nominal heads, and undergo a set of sandhi rules similar to those affecting definite determiners. These also show past/nonpast distinctions: *gúri* 'your house' *gúrigàagii* 'your house.PST' (Saeed 1999:115), with the meaning 'your former house'.

[10]In Somali, case is marked on the rightmost constituent of the nominal phrase, while definiteness is marked on the head noun.

5.1 Interpretation of Nominal Tense

Nominal tense is independent of clausal tense, in that it may coincide or differ from the tense of the verbal predicate. Thus in (54) and (55) below, the termporal location of the clause and that of the nominal happen to coincide, while in (56) and (57) they do not.

(54) *dhibaatá-da Khalí ij-ku welí way taagán tahay*
problem-DET.F Gulf-DET.M.NOM still FOC.3S permanent is
The crisis of the Gulf still persists

(55) *dhibaatá-dii Khalíij-ku wáy dhammaatay*
problem-DET.F.PST Gulf-DET.M.NOM FOC.3S end.PST
The crisis of the Gulf ended (Lecarme 1999:335)

(56) *ardáy-da baan kasin su'áash-aadii*
students-DET.F FOC.NEG understand.PST question-DET.F.POSS2S.PST
The students (who are present/I am telling you about) did not understand your question.

(57) *ardáy-dii wáy joogaan*
students-DET.F.PST FOC.3P are present.-PST
The students (I told you about) are present. (Lecarme 1999:335)

Lecarme also provides some examples of the choice of determiner leading to specifi cpresuppositions: (58) is only appropriate if the speaker believes the exhibition is closed at Utterance time, and (59) only appropriate if the speaker believes it is open. Likewise, the choice of determiner in (60) is determined by whether the journey is still in progress or not.

(58) *bandhíg-gii máad daawatay?*
exhibition-DET.M.PST Q.2S see.PST
Have you seen the exhibition(still running/closed at UT) (Lecarme 1999:338)

(59) *bandhíg-ga máad daawatay?*
exhibition-DET.M Q.2S see.PST
Have you seen the exhibition(still running/closed at UT) (Lecarme 1999:338)

(60) *búugganu sáfarkayga/-ii buu tilmáanayaa*
book.DET.M.NOM journey.DET.M.POSS1S./PST FOC.3S relates
This book relates to my journey (Lecarme 1996:7)

A possible view of the nonpast/past distintion in affi xaldeterminers might be that it encodes the presence (or absence) of 'present relevance'. There are several indications, however, that the distinction is temporal in character. In discussing the choice of determiners in examples (56) and (57) Lecarme notes that the second of these sentences, (57), "is only possible if the discourse has already mentioned some past time which is taken as the reference point, that is, a time already given in the context" (Lecarme 1996:6). Further evidence that this phenomenon is not just marking of present relevance is provided by the fact that overt (nominal) temporal modifi ersmust occur with a matching tense marking. In the examples below, the temporal modifi er 'next year' selects a non-past determiner, while 'last year' selects a past determiner: clearly, depending on the nature of the event, some event taking place last year may still have 'present relevance'.

(61) *sánnad-ka/*-kii dambe*
year-DET.M next
next year

(62) *sánnad-kii/*-ka hore*
year-DET.M.PST before
last year (Lecarme 1999:342)

Note that in (57), *ardáy-dii* (students-DET.F.PST) does not have the interpretation 'ex-students': that is, it is not the event or predication time (that is, the time of being a student) which is fixed in the past, but rather the reference time. The addition of a temporal modifier such as *hore* 'before' may unambiguously fix the event or the reference time. In (63), *ardáyday-dii* excludes the possibility of the individuals still being students at the time of utterance (that is, the nominal predication or 'event' itself is temporally restricted).

(63) *ardáyday-dii hore dhammáan-t-ood (waa ilá soo xariiran).*
students-DET.F.POSS1S.PST before entirety-DET.F-POSS3P (are in contact with me)
All my ex-students (are in contact with me). (Lecarme 1999:342)

Definite temporal modifiers like *shálay* 'yesterday' provide a (contextually determined) reference time:

(64) *qabqabashá-dii shálay*
arrests-DET.F.PST yesterday
Yesterday's arrests (Lecarme 1999:342)

Although nominal tense may occur as the sole overt expression of tense in so-called *nominal clauses*, note that the nominal tense does not determine the clausal interpretation in these cases. The determiner in (65) is present while that in (66) is past, but both clauses are given 'present' interpretations. The lexically empty word *waa*, glossed here by Lecarme as a focus marker, occurs in positive, declarative verbless sentences. This particle, analysed by Saeed (1999) as a sentence type marker, often serves to focus the complement NP in sentences such as these.

(65) *búug-gii wáa kan*
book-DET.M.PST FOC DET.M.DEM
Here is the book (distant but in sight, I have in mind, I told you about)

(66) *nimá-kii waa macallimíin*
men-DET.M.PST FOC teachers
The men (over there, I have in mind, I told you about) are teachers (Lecarme 1999:335)

To summarise, then, in Somali definite determiners (which are nominal affixes) show a past/non-past distinction. Nominal tense and verbal tense have different morphological forms. Nominal tense is independent of verbal tense in Somali and its domain is restricted to the DP. It may be interpreted referentially, deictically or anaphorically (and thus, the interpretation of nominal tense is not always determined by the discourse context).

5.2 Morphosyntactic Aspects

Nominal tense is implicated in the syntax of Somali in several intriguing ways. One of these concerns tense agreement within the noun phrase. Adjectives used attributively also inflect for tense, sharing the tense endings of the highly irregular verb 'be' (*ø* -PST, *-aa* PST.M, *-ayd* PST.FS). Adjectives agree in gender and tense, with adjectival tense only appearing when the noun is definite (and therefore tensed) (Lecarme 1996:4, Lecarme 1999:343).[11] The following examples show gender and tense agreement with masculine singular, feminine singular and feminine plural nouns respectively.

(67) *árday-ga wanaagsan*
student-DET.M good

(68) *árday-gii wanaagsan-aa*
student-DET.M.PST good-PST
the good student

(69) *ardayád-da wanaagsan*
student(f)-DET.F. good

(70) *ardayád-dii wanaagdan-ayd*
student(f)-DET.F.PST good-PST.F
the good student

(71) *ardáy-da wan-wanaagsan*
students-DET.F PL-good
the good students

(72) *ardáy-dii wan-wanaagsan-aa*
students-DET.F.PST PL-good-PST
the good students

Predicative nouns in modifierposition also display concord, agreeeing in tense and definiteness. These nominal modifiers occur with a copula element: the copula shows tense agreement and the nominal definiteness agreement. This concord phenomenon appears with proper names and with idioms (see (75)-(77)):

(73) *ardayád-da soomaalí -da ah*
student(f)-DET.F Somali-DET.F be
the Somali student(f)

(74) *ardayád-dii soomaalí -da ahayd*
student(f)-DET.F.PST Somali-DET.F be.PST
the Somali student(f) (who phoned you) (Lecarme 1999:344)

(75) *dád fará badan*
people fingers many
many people

(76) *dád-ka fará-ha badan*
people-DET.M fingers-DET.M many
the numerous people

(77) *dád-kii fará-ha badn-aa*
people-DET.M.PST fingers-DET.M many-PST
the numerous people (past) (Lecarme 1999:344)

Lecarme (1999) further argues that nominal tense is implicated in licensing certain possessor constructions in Somali.[12] The syntactic expression of possession can take the form of either a construct state construction or a prenominal genitive construction. In the construct state construction the head and the possessor are both obligatorily definite and tense is marked only on the head:

(78) *dhibaatá-dii Khalí ij-ku (wáy dhammaatay)*
problem-DET.F.PST Gulf-DET.M.NOM (FOC.3FS end.PST)
the crisis of the Gulf (ended) (Lecarme 1999:345)

[11] Number may be marked through optional reduplication in Somali adjectives.
[12] Her insight is essentially that the availability of the +/-T feature within DP makes the possessor relation visible.

In the prenominal genitive construction, tense appears on the head, after the genitive infix coding the prenominal possessor:

(79) *Khalíij-ka dhibaatá-d-iis-ii (wáy dhammaatay)*
Gulf-DET.M problem-DET.F-POSS3M-PST (FOC.3FS end.PST)
the Gulf crisis (ended) (Lecarme 1999:345)

As noted above, the construct state construction requires the head noun to be definite. Relational nouns (which lack autonomous reference) and other non-referential nouns (e.g. *láf-(ta)* 'soul, self') can take a definite determiner, but are unable to take tense marking. Such nouns permit only the prenominal genitive construction (as they lack tense, although definite).

(80) *inán-tii yar-ayd iyo waláalkeed/*-ii)*
girl-DET.F.PST small-PST.FS and brother.DET.M.POSS2FS/*PST
the small girl and her brother (Lecarme 1999:348)

(81) *af-soomáali-ga láf-t-iisa/(*-ii)*
language-Somali-DET.M bone-DET.F-POSS3MS/*PST
the Somali language in itself (Lecarme 1999: 348)

As we have seen, nominal tense in Somali is independent of clausal tense, is inextricably associated with definiteness in affixal determiners, and can be used referentially or deictically or be linked to the domain of discourse. In the final section we look briefly at an example of languages in which nominal tense is essentially bound up with deictic markers.

6 Nominal Tense, Deixis and Visibility in Iraqw and Mao Naga

In some languages nominal tense markers also express notions of visibility/nonvisibility, spatial or discourse proximity, and deixis. Such languages include the Cushitic language Iraqw, spoken in Tanzania (Mous 1993). Iraqw has four affixal determiners (glossed DEM for 'demonstrative') which encode meanings of spatial proximity and visibility as follows: *-í* or *-ká* 'near the speaker'; *-síng* 'near the addressee'; *-qá'* 'near neither of them but still visible, or mentioned earlier'; and *-dá'* 'far away, or mentioned earlier' (Mous 1993: 90). Some of these determiners also encode temporal information. When used to refer back to a previously mentioned noun, *-qá'* is used in the present tense and *-dá'* in the past tense (Mous 1993:91):[13]

(82) *xa'i i-na túu' xa'i-dá' ka kwáahh*
trees O.N-PAST uproot:3SG.M:PAST trees-DEM4 O.3:IMPS:OBJ.N:PERF throw:PST
He uprooted trees ... Those trees were thrown away.

(83) *gwara-r-qá' hhiya-'ée' i-r gwâa'-i*
death-F-DEM3 brother-1.SG.POSS S.3-INST die:INT-INF:S.3
Is that a death for my brother to die? (after a sentence about the way he died)

[13] Note that *i-na*, *ka* and *i-r* in the following examples are verbal auxiliaries containing subject/object markers and TAM information. Abbreviations include: DEM 'demonstrative', F 'feminine', IMPS 'impersonal subject', INF 'infinitive', INT 'interrogative', O.3 'third person object', O.N 'neuter object', PERF 'perfect', S.3 'third person subject'

Thus, with *-qá'* we see the association of visibility with present tense, and with *-dá'* the association of spatial distance with past tense.

A somewhat similar association is found in Mao Naga, a Tibeto-Burman language from India (Giridhar 1994). In this language, nouns can be inflected with one of three 'individuation' markers, which encode meanings of (in)visibility, location with respect to speaker, and discourse prominence: *-hi* marks visible or spatially proximate entities; *-ṭi* marks nonvisible entities which are known or familiar; and *-sü* marks nonvisible entities which are unfamiliar, or only vaguely remembered (Giridhar 1994:118-9).[14]

Interestingly for our purposes, these markers can also encode temporal distance with *-hi* referring to the present time (the time of utterance), *-ṭi* the past, and *-sü* the future (pp. 135-6). Thus:

(84)
hata-li-hi 'in the current week'
hata-li-ṭi 'in the past week'
hata-li-sü 'in the week to come'

(85)
ovo koso-hi 'the current work'
ovo koso-ṭi 'the past/done work'
ovo koso-sü 'the work still to be done'

(86)
ni cümüi-hi 'your current wife'
ni cümüi-ṭ i 'your past wife'
ni cümüi-sü 'your future wife'

As in Iraqw, the Mao Naga data shows the association of visibility (and spatial proximity) with present tense (*-hi*). Nonvisibility is associated both with past tense (*-ṭi*) and future tense (*-sü*), the difference being one of familiarity and/or discourse prominence (not surprisingly, past tense is associated with familiarity, and future tense with nonfamiliarity).

7 Conclusion

In this paper we have attempted to provide a preliminary sketch of the phenomenon of (morphologically expressed) nominal tense with nominal scope, distinguishing this from nominal tense with clausal scope. We have seen that NT-NS occurs in a range of languages from distinct language families, and that it sometimes, but not always, shares a set of exponents with verbal tense. The expression of nominal tense may be tied up with the expression of possession (Hixkaryana, Somali), definiteness (Somali), deictics or demonstratives (Iraqw and Mao Naga). Syntactically, data from Somali show that nominal tense is syntactically active, controlling concord phenomena and playing a crucial role in the grammaticality of construct state constructions, suggesting the presence at f-structure of (nominal) tense features alongside nominal agreement and definiteness features. Semantically, data from several languages show a range of clear temporal interpretations for nominal tense with nominal scope, suggesting that the semantics of nominal structures in these languages involves reference and event times (where the event can equally well bey the "possession event" in possessive nominals). Nominal tense affixes may also have a deictic interpretation, or be related to the expression of (non)visibility.

[14] We are grateful to D.N.S. Bhat for drawing our attention to this Mao Naga data.

Bibliography

Blake, B. 1987. *Australian Aboriginal Grammar*. London: Croom Helm.

Burton, S. C. 1997. Past Tense on Nouns as Death, Destruction, and Loss. In K. Kusomoto (Ed.), *NELS 27*, 65–77.

Demirdache, H., and L. Matthewson. 1995. On the Universality of Syntactic Categories. In *NELS 25*, 79–93.

Derbyshire, D. 1979. *Hixkaryana.* Amsterdam: North Holland.

Derbyshire, D. 1999. Carib. In R. M. W. Dixon and A. Aikhenvald (Eds.), *The Amazonian Languages*, 23–64. Cambridge, UK: Cambridge University Press.

Enc, M. 1986. Towards a Referential Analysis of Temporal Expressions. *Linguistics and Philosophy* 9:405–426.

Galloway, B. 1993. *A Grammar of Upriver Halkomelem.* Berkeley: University of California Press.

Gerdts, D. B. 1988. *Object and Absolutive in Halkomelem Salish.* New York: Garland.

Giridhar, P. P. 1994. *Mao Naga Grammar*. Mysore: Central Institute of Indian Languages.

Jelinek, E. 1996. Quantification in Straits Salish. In Emmon Bach, Eloise Jelinek, Angelika Kratzer and Barbara Partee (Ed.), *Quantification in Natural Languages*. Dordrecht: Kluwer.

Jelinek, E., and R. Demers. 1994. Predicates and Pronominal Arguments in Straits. *Language* 70:697–736.

Koptjevskaya-Tamm, M. 1993. *Nominalizations.* London: Routledge.

Lecarme, J. 1996. Tense in the nominal system. In J. Lecarme, J. Lowenstamm and U. Shlonsky (Ed.), *Studies in Afroasiatic Grammar*. The Hague: Holland Academic Graphics.

Lecarme, J. 1999. Nominal Tense and Tense Theory. In Francis Corblin and Carmen Dobrovie-Sorin and Jean-Marie Marandin (Ed.), *Empirical Issues in Formal Syntax and Semantics 2*, 333–354. The Hague: Thesus.

Mous, M. 1993. *A Grammar of Iraqw*. Hamburg: Helmut Buske Verlag.

Nordlinger, R., and L. Sadler. 2000. Tense as a Nominal Category. In M. Butt and T. H. King (Eds.), *Proceedings of LFG 2000*

Nordlinger, R., and L. Sadler. in preparation. Nominal Tense: A Typological Survey.

Saeed, J. 1999. *Somali*. Amsterdam: John Benjamins.

Tonhauser, J. 2000. Dynamic Semantics and the Temporal Interpretation of Noun Phrases. Diploma thesis, IMS, Stuttgart.

Do Sentences Have Tense?

Christoph Schwarze
University of Konstanz

Proceedings of the LFG01 Conference
University of Hong Kong, Hong Kong

Miriam Butt and Tracy Holloway King (Editors)\

2001

CSLI Publications
http://csli-publications.stanford.edu

Christoph Schwarze (University of Konstanz)

Do sentences have tense?

6th International LFG Conference, Hong Kong, June 25-27, 2001

Workshop on Tense

0. Introduction

Under the classical treatment in LFG, tense is a feature encoded in morphology or in the lexicon and projected into f-structures. It is formulated, as all other features, as an attribute-value equation; the values of the tense attribute generally are the labels by which traditional grammar refers to the tenses of the specific language.

A fact often not made explicit by syntacticians is that these labels need to be interpreted in semantics. The semantics of tenses has been formulated as a theory of temporal reference since the seminal work of Reichenbach (1947). Subsequently this theory has been integrated into a theory of context change, which widens the domain of tense semantics from the sentence to discourse (Kamp. 1979, Kamp & Rohrer 1983). Functional linguistics since Weinrich (1964, 1971) investigates the role tenses play in textual organization and discourse pragmatics.

The aim of this workshop is to discuss a conception in which tense and temporal reference is distributed in a new way. It is assumed that tense features belong to morphological structure and are not projected to f-structure. Instead, they are mapped onto f-structure features based on the insights of tense semantics. Since there are no tense features in f-structure, according to this approach, sentences do not have tense.

In this contribution to the workshop, I will discuss this idea from a slightly skeptical point of view.

1. Some general reflections on features

Features are a format of representation. There is a wide range of different facts, which can be expressed by features. Therefore, features have no ontology. Reflecting on features can only mean to be aware of the use, which a certain type of linguistic thinking makes of features and how this use can be improved.

Since features do not have an ontology, it is not obvious why different kinds of features should be defined for each layer of representation. Features flow from layer to layer. But not all features flow in the same way.

2. How to model the morphology-syntax interface

Some features (μ-features) are needed only in morphology. Tense auxiliaries are assigned a μ-tense value in morphology, which is not passed on as such to syntax, but is needed for the elaboration of the tense value at the level of the sentence (see below). So-called past participles of passivizable verbs have two mutually exclusive features at the sentence level (past vs. passive), but morphology does not make such a difference. Therefore there is only one morphological feature for both, call it "μ-PARTICIPLE = II", which must be replaced with "PARTICIPLE = PAST" or "PARTICIPLE = PASSIVE" before lexical insertion. A third case regards inflectional classes. They can be treated as morphological agreement operating on inflectional class features, such as "μ-CLASS = n". Schwarze (1999a). But there is no need to postulate a full set of μ-features, most of which are trivially mapped to syntax[1]. On the contrary, it is linguistically interesting to see how some features flow across borders, from morphology to syntax and semantics, and others, the morphological features proper, are confined to morphology (cf. Mayo 2000:170).

3. Kinds of features

A feature system may be described according to two points of view: their origin, i.e. the place where they are encoded (3.1), and their flow within the grammar and their effect (3.2).

---

[1] An alternative to such a full set of μ-features is to introduce, in order to account for morphological structure, a specific set of morphological c-structure categories. Schwarze (1999b) uses lower case letters for lexical words (n, v, a, …) and other morphological segments: n/v, v/n, n/n, … for affixes of word formation, n_infl, v_infl, … for inflectional affixes.

3.1 Where are features encoded?

3.1.1 Lexically encoded features

Most features are or may be lexically encoded. For reasons of computational realism, the lexicon must be modeled as being organized into various regions (Mayo (2000). Hence lexically encoded features can be located more specifically.

The largest region is ***Lexical Conceptual Structure*** (LCS, cf. Mayo 2000:195). All lexical words are encoded in that region, mainly with their predicate features (1). Agreement features (2) are encoded in LCS if they are properties of lexical words. This is the case, e.g. for gender in languages like French or German, where all nouns have a lexically encoded gender value. Inflectional class features are encoded in LCS under the same condition. (3) refers to the main classes of verbs inflection of Italian; the values are the respective infinitive endings.

(1) PREDICATE = 'SEE <(SUBJ), (OBJ)>', ...

(2) GENDER = {MAS, FEM}
NUMBER = {SG, PL, ...}

(3) μ-CLASS = {ARE, IRE, ERE}

Other features, concerning, e.g., agreement and tense, may be encoded in LCS if the are properties of irregularly inflected word forms. These, in fact, are stored as syntactic words in LCS.

Last, but not least, grammatical functions may be encoded in LCS, namely in the lexical entries of control verbs.

The next region is the region of ***syntactic operators*** (Opr, cf. Mayo 2000: 197). It contains "grammatical" prepositions, determiners, negation particles and similar forms. Examples of the features encoded in Opr are given under (4).

(4) P_CASE = {A, DE, ...}
SPEC = {DEFINITE, DEMONSTRATIVE, POSSESSIVE, PARTITIVE, ...}
NEG = +

Grammatical functions may also be thought to be encoded in Opr, e.g. if there are pronouns which are specialized for the subject function.

Finally, there is the ***morphological*** region of the lexicon (Mrph, cf. Mayo 2000: 1999). It contains all kinds of affixes. Among the features encoded in Mrph we find again the agreement and the inflectional class features, which, as other features, thus have competing different origins. Moreover, Mrph contains the μ-features for tensed auxiliaries and, in the classical treatment of tense, the tense and mode values that are expressed by suffixes.

Those derivational affixes, which operate on the lexical meaning of their bases, encode a derivational predicate (DPRED, cf. Mayo et al. 1995:932), which triggers a semantic change. In the lexical model proposed in Mayo et al. (1995), Schwarze (1999b) and Mayo (2000) DPRED must be resolved, i.e. a new predicate must be derived, before lexical insertion.

3.1.2 Features encoded as rule annotations

One of the most basic properties of LFG is that grammatical functions are normally introduced as c-structure annotations. The same holds for sentence types (declarative, interrogative etc.) But according to the conception of morphology advocated here, features may also be introduced by annotations to morphological rule. (5) is a rule which, for languages like English, French or German derives singular nouns without adding a suffix. According to the conventions mentioned in fn. 1 above, *N* is a syntactic noun, accepted by the syntax, and *n* a lexical noun, which has no number.

(5) $$\mathrm{N} \rightarrow \underset{(\uparrow \text{ NUM}) = \text{SG}}{\underset{\uparrow = \downarrow}{\mathrm{n}}}$$

3.1.3 Derived features

A small number of features are ***lexically derived***. The functional analysis of participles mentioned above is such a case. Two lexical rules may by formulated, (6) for the participles of all verbs, and (7) for the participles of passivizable verbs:

(6) μ-PARTICIPLE = II ⇒ PARTICIPLE = PAST

(7) μ-PARTICIPLE = II ⇒ PARTICIPLE = PASSIVE
If OBJ ∈ the grammatical functions governed by the verb

Obviously, the formulation of the condition in (7) is only a sketch.

Other features are ***constructionally derived***. This is the case for the representation of compound tenses (8) and aspect periphrases (9), furthermore for the number and gender of coordinated noun phrases (10), and for the imperative mode, when the verb has the same form as the infinitive (11) or the subjunctive (12).

(8) Fr. Il a plu.
it has rained
'It rained'
TENSE = COMPOUND_PAST

(9) Fr. Il va pleuvoir.
it goes rain
'It is going to rain'
TENSE = PRESENT, ASPECT = IMMINENT

(10) Fr. le garçon et la fille
[the boy-MAS-SG and the GIRL-FEM-SG]
'The boy and the girl'
NUM = PL, GEN = MAS

(11) It. Non preoccuparti.
not worry-INFINITIVE you
'Don't worry'
MODE = IMPERATIVE

(12) It. Non si preoccupi.
not REFL-PERS3 worry-PERS3-SUBJUNCTIVE
'Don't worry' (politeness form)
MODE = IMPERATIVE

Since tense is the topic of this workshop, I will briefly show how compound tenses can be treated in terms of constructionally derived features. The values of the TENSE attribute are a closed list, which varies cross-linguistically. For French they are listed under (13).

(13) TENSE = {PRESENT, FUTURE, CONDITIONAL, IMPERFECT, SIMPLE_PAST, COMPOUND_PAST, PLUPERFECT, PAST_ANTERIOR, FUTURE_ANTERIOR}[2]

Five of these tenses (PRESENT, FUTURE, CONDITIONAL, IMPERFECT, SIMPLE_PAST) are realized as simple tenses, i.e.. by inflection only. The remaining four (COMPOUND_PAST, PLUPERFECT, PAST_ANTERIOR, FUTURE_ANTERIOR) are realized as compound tenses, i.e. by an auxiliary, which is inflected as a verb, and a past participle. Since the compound

[2] One might want to classify the conditional as a mode, rather than as tense. I will not go into this problem in this paper.

tenses form a closed list, they can be derived by a list of inferences, which operate on f-descriptions, namely:

(14) ***Derived tense features***

(↓ AUX μ-TENSE) = PRES & (↓ PART) = PAST ⇒
(↑ TENSE)= COMPOUND–PAST

(↓ AUX μ-TENSE) = FUT & (↓ PART) = PAST ⇒
(↑ TENSE) = FUT_ANT

(↓ AUX μ-TENSE) = IMPERF & (↓ PART) = PAST ⇒
(↑ TENSE) = PLUPERFECT

(↓ AUX μ-TENSE) = SIMPLE_PERF & (↓ PART) = PAST ⇒
(↑ TENSE) = PAST_ANT

3.2 What features do and how they flow

Inflectional class features support morphological wellformedness. They stay in morphology.

Predicates express the distinctiveness of lexical and pronominal meanings. They are pointers to the semantics. They are projected from the lexicon to f-structure and to semantic structure (σ-structure).

Gender features support grammatical and anaphoric agreement. They are projected from morphology and the lexicon to f-structure, and, to the extent they are used for anaphoric agreement, to σ-structure, where they are needed for discourse semantics.

Person and number features are similar. They support grammatical and anaphoric agreement and may be interpreted in sentence and discourse semantics (anaphora). If this is the case, they are passed on from morphology and the lexicon to f-structure and σ-structure.

Specification features support sentence as well as discourse semantics. They flow from the lexicon to f-structure and σ-structure.

PCase features distinguish subclasses of grammatical relations (objects, obliques, infinitival complements, adjuncts). With the exception of those which distinguish subclasses of infinitival complements, they are projected from the lexicon to f-structure and σ-structure.

Tense and aspect[3] features, in the classical view, are interpreted in sentence and discourse semantics and pragmatics. If a language has tense agreement (*consecutio temporum*) they support mere wellformedness conditions. They are projected from morphology to f-structure and, in most cases, to σ-structure.

Mode features may support mere syntactic wellformedness conditions, but may also be interpreted in sentence semantics and discourse pragmatics. If this is the case, they are projected to f-structure and to σ-structure.

Sentence type features support syntactic wellformedness and express prototypical kinds of speech acts per default. They are shared by f- and σ-structure.

Grammatical functions support syntactic wellformedness and are a criterion for distinguishing semantic argument positions. They stay at the f-structure level, but are mapped onto semantic arguments in σ-structure.

4. F-structures, σ-structures and semantics

At this point, it may be useful to mention that f-structures are not semantic representations. As was implied in the characterizations given in section 3.2, they contain two kinds of information, information, which is needed for constraints on syntactic and semantic wellformedness, and information that must be available to semantics. This part of the information contained in f-structures is mapped onto σ-structure. And σ-structure, which is f-structure stripped of that information, which is not needed for semantics, is often semantically underspecified. Thus the values of the SPEC attribute, e.g., are related to, but do not determine quantification. SPEC = DEF, e.g., may be interpreted as the iota-operator, as the generic quantifier etc. It is the task of semantics proper to resolve such cases of underspecification.

[3] It is useful to make a distinction between tense and aspect feature on formal grounds, because aspect periphrases interact with (morphological) tense. But semantically, tense and aspect features are not clearly separated.

4.1 To what degree should f-structure prepare semantic interpretation?

F-structures may vary with respect to whether they have a closer resemblance with c-structures or with semantic representations.

They are naturally remote from c-structures if they are mapped onto a set of very different c-structures. NUMBER is an example. On one hand it originates in morphology, on the other hand it is derived at the sentence level in coordinated NPs.

In actual grammar writing, there often are alternatives regarding the f-structure representation one wants to have: should a given partial f-structure be closer to c-structure or closer to semantics? Lexical prepositions are an example. The partial f-structure is closer to c-structure, if the preposition is represented as a one-place predicate (15). It is closer to semantics, if the preposition is a two-place predicate (16):

(15) PRED = 'DANS <(ARG)>'
(16) PRED = 'DANS <(ARG1), (ARG2)>'

In (15), the right sister of the preposition is ARG; in (4), it is ARG2. When local PPs are represented as in (15), f-structure only identifies the localizing object (the GROUND); semantics gets no cue how to find the localized object (the FIGURE). On the contrary, when local PPs are represented as in (16), f-structure helps to also identify the localized object (the FIGURE), via lexical semantics. ARG1 is identical with the SUBJECT if the verb is an intransitive verb of motion, with the OBJECT if the verb is a transitive verb of motion, and it is the event variable if the PP is an ADJUNCT.

A treatment that is close to semantics gives semantics a more precise input. It is not clear, however, whether it reduces computational complexity.

4.2 How can temporal information be made more explicit at the f-structure level?

Tense features, as used in the classical LFG treatment, are remote from semantic representations of the Reichenbach tradition. When tense is morphologically expressed by a suffix, there is a formal similarity between (morphological) c-structure and the representation of tense at f-structure, even though f-structure displaces tense from the verb to the sentence. But there is no resemblance at all between c-structure and f-structure when the value of the TENSE attribute is constructionally derived. Why, then, should one not treat tense in such a way that its functional representation is equally remote from morphological and constructional encoding? To simply eliminate tense features from f-structure seems to be a plausible reaction to this situation.

As a concrete research program, that solution would mean that one would have to reformulate some version of temporal semantics in the format of the LFG feature calculus. Under (17) - (25) I give a sketchy overview of the features one would need to introduce. Notice that some of the values are sorted variables. *t* for times, *s* for states of affairs (events, states, etc.). Moreover, the values of attributes representing relations between time intervals (23) have a kind of complexity, which is not used in classical LFG.

(17) STATE_OF_AFFAIRS = {EVENT, STATE, … }

(18) EVENT = {s, …}

(19) STATE = {s, …}

(20) TIME = {UTTERANCE_TIME, EVENT_TIME, REFERENCE_TIME}

(21) UTTERANCE_TIME = {t, …}

(22) TIME–INTERVAL = {t, …}

(23) TEMPORAL–RELATION = {BEFORE (ARG1), (ARG2), AFTER (ARG1), (ARG2), COINCIDING (ARG1), (ARG2)}

(24) ARG1 = {t,…}

(25) ARG2 = {t,…}

Reformulating standard logic into a feature calculus is probably feasible. But is it reasonable? A temporal semantics will be needed anyway. Does it make sense to shift such an amount of structure from semantics to syntax? I do not see the criteria on the base of which one could decide whether this should be done or not, and there is a risk that the efforts which such a research program requires may not bring about an advance in linguistics. Of course this is just a doubt.

But there are more tangible objections to the research program under discussion. One is that a given tense may be ambiguous and that the various interpretations must be retrieved from the context. A second objection is that the meaning of tense is not restricted to temporal reference.

4.3 The semantic variation of tenses

The semantics of temporal reference has to deal with the polysemy of tenses. The French Imparfait is an example:

(26) Il y a une heure, il **pleuvait** encore.
there is an hour it rained-IMPERF still
'An hour ago, it was still raining.'

The sentence refers to a single event, where the event time is before the utterance time, and it is of unbounded duration.

(27) Quand j'étais jeune, **j'allais** nager tous les jours.
when I was young I went-IMPERF swim all the days
'When I was young, I went swimming every day.'

The sentence refers to a sequence of repeated events, where the event time is before the utterance time and the repetition unbounded.

A more intricate polysemy is inherent to the French Simple Perfect, as used in literary narrative. When verb forms in the Simple Perfect follow each other in a text (28), the default is that these verbs refer to a sequence of events; with each verb the text world advances in time.

(28) Sara **alla** dans la cuisine, **avala** un bol de café froid et **sortit** sur la véranda (Klein & Kleineidam 1994:268)

'Sara went into the kitchen, swallowed a bowl of cold coffee and went out to the veranda.'

This interpretation depends on various factors. It requires the identity of the subject and scripts for coherent sequences of events. If these conditions are not met, a series of verbs in the Simple Perfect may very well refer to events, which are not in a sequence, cf. (29):

(29) Après le baccalauréat, Charles **commença** ses études de médecine, Gaston **fit** le tour du monde et Édouard se **maria**.

'After high school graduation, Charles began to study medicine, Gaston went on a world tour, and Édouard got married.'

This example shows that at the text level, there is additional systematic semantic variation in temporal reference.

4.4 Temporal reference vs. the semantics of tense

The semantics of temporal reference does not capture the whole range of polysemy that tenses may have. Therefore a distinction must be made between the semantics of temporal reference and the semantics of tense.

The ***semantics of temporal reference*** extracts relevant information from tenses and the lexical semantics of verbs, temporal adverbs and prepositions, and other lexical material. The ***semantics of tense*** must account for the whole range of information conveyed by tenses.

In addition to temporal reference tenses also convey modal and pragmatic information. Take the French Future as an example:

(30) Je sais qu'elle **partira** demain.
I know that she leaves-FUTURE tomorrow
'I know that she will leave tomorrow.'

The sentence refers to an event whose time is after utterance time.

(31) Elle **aura** une trentaine d'années.
she has-FUTURE a thirty-APPROXIMATION of years
'She may be about thirty years old.'

The sentence refers to a situation whose time coincides with the utterance time; it conveys the modality of a supposition.

(32) Pour moi ce **sera** un thé.
for me it is-FUTURE a tea
'I will have a cup of tea.'

The sentence refers to an event whose time is after the utterance time; it conveys the pragmatic value of an order; it is specific to the speech-act of ordering in a café or restaurant.

(33) Pour vous répondre, je vous **dirai** que ...
for you answer I you say-FUTURE that ...
'To give you an answer, I'd say that ...'

The sentence refers to an event whose time coincides with the utterance time; it conveys the pragmatic value of politeness.

If tense values were to be banned from f-structure, the complete modal and pragmatic information would have to be made explicit at the passage from morphological tense to f-structure. Therefore the program to replace the tense features of the sentence with more explicit features is far more complex than it seems to be.

5. Conclusion

So what should be the answer to the question I formulated in the title? The answer differs, depending on the point of view.

From the point of view of theory, the answer is negative. Sentences have a form and a meaning. Regarding the expression of temporal information, they have tensed verbs on the one hand, and time reference on the other. It cannot empirically be shown that sentences "have tense".

From the point of view of method, however, the answer is less obvious. F-structure is an intermediate layer of representation between constituent structure and semantics. Its main justification is its contribution to the expression of wellformedness conditions for sentences. A second important justification is its usefulness for cross-linguistic analyses and typology. In fact, f-structures, even though they are not universal, are a better *tertium comparationis* than c-structures. And there might be a third justification of postulating a level of f-structure, namely that it reformulates morphological and surface syntactic information in such a way that semantics can be done in more general terms.

The substitution of μ-features with tense features proper to f-structure, as far as compound tenses and aspect periphrases are concerned, goes in that direction. It frees semantics from the task of computing constructionally encoded temporal information. Tense features at the f-structure level are well motivated under this perspective. It is doubtlessly possible to without them. But I hope to have shown that such a revision of the model will meet more difficulties than it seems to at a first glance.

References

Mayo, Bruce. 2000. A Computational Model of Derivational Morphology. Ph.D. Thesis Universität Hamburg. http://allegro.sub.uni-hamburg.de/emedien?dm=f&hos=uni&fb=18

_____, Marie-Theres Schepping, Christoph Schwarze and Angela Zaffanella. 1995. "Semantics in the derivational morphology of Italian: implications for the structure of the lexicon". *Linguistics* 882-938.

Kamp, Hans. 1979. "Events, Instants and Temporal Reference". In: Rainer Bäuerle, Urs Egli and Arnim von Stechow (eds.). *Semantics from Different Points of View*. Berlin / Heidelberg / New York: Springer. pp. 376-417.

_____ And Christian Rohrer. 1983. "Tense in Texts". In: Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow (eds.). *Meaning, Use, and Interpretation of Language*. Berlin / New York: de Gruyter. pp. 250-269.

Klein, Hans-Wilhelm and Kleineidam, Hartmut. 1994. *Grammatik des heutigen Französisch für Schule und Studium*. Neubearbeitung. Stuttgart: Klett.

Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. London: Mc Millan.

Schwarze, Christoph. 1996. "The syntax of Romance auxiliaries". In: Miriam Butt & Tracy Holloway (eds.). *LFG-Workshop. Proceedings of the First LFG Conference*. Rank Xerox Research Center, Grenoble, August 26-28, 1996. pp. 418-433. http://csli-www.stanford.edu/publications
Also published in printing as 1998. "A lexical-functional analysis of Romance auxiliaries". *Theoretical Linguistics* 24. 83-105.

_____. 1999a. "Inflectional classes in Lexical-Functional Morphology". In: Miriam Butt & Tracy Holloway King (eds.). *The Proceedings of the LFG '99 Conference*. University of Manchester: CSLI Publications. http://csli-publications.stanford.edu/LFG/4/lfg99-toc.html

_____. 1999b. "Lexical-Functional Morphology and the structure of the lexicon". In: Lunella Mereu (ed.). *Boundaries of Morphology and Syntax*. (Current Issues in Linguistic Theory, 180). Amsterdam - Philadelphia: J. Benjamins. pp. 73-95.

Weinrich, Harald. 1971. *Tempus. Besprochene und erzählte Welt*. Zweite, völlig neubearbeitete Auflage. Stuttgart: Kohlhammer. First edition 1964.

TENSE, ASPECT AND THE TEMPORAL STRUCTURE OF DISCOURSE: TOWARDS AN LFG ACCOUNT

Sheila Glasbey

Department of Applied Computing
University of Dundee
Dundee DD1 4HN
Scotland, UK

srg@computing.dundee.ac.uk

Proceedings of the LFG01 Conference
University of Hong Kong,
Hong Kong

Miriam Butt and Tracey Holloway King (Editors)
2001

CSLI Publications
http://csli-publication.stanford.edu/

Abstract

This paper is based on my contribution to the workshop on Tense and Aspect at LFG01. In it, I consider the question of how information relating to tense, aspect and the temporal structure of discourse should best be encoded in LFG. I will focus in particular on what temporal information needs to be represented if we are to be able to map from the LFG formalism to some semantic representation which captures the temporal structure of continuous discourse. In particular, I will derive a preliminary set of features which are intended to be used cross-linguistically to represent tense- and aspect- related information.

Acknowledgements

I would like to thank Miriam Butt for inviting me to give a presentation at the Tense and Aspect Workshop at LFG01. I am also very grateful to all the conference participants who gave me helpful advice and comments, and contributed to a very lively and stimulating discussion.

1. Introduction

This paper addresses the question of what “temporal” information (including tense, aspect and related properties) would need to be encoded in an LFG representation, if we wished to use such a representation in order to derive the temporal structure of a piece of continuous discourse.

This very general question immediately introduces a set of more specific questions: for example:

(i) What “features” (in the LFG sense) are needed to encode the relevant information?

(ii) At what “level” should these features be represented?

Most of what follows will address question (i) – but let us first briefly consider (ii). By “level” I mean level of structure, so that one possibility might be to encode all temporal information at f-structure. Indeed, if f-structure is taken to be the sole input to semantic representation (as is stated in, for example, Butt et al 1996 p.3 and Frank and Zaenen (draft) p.1), then this appears to be a sensible way to proceed. As this stance is taken by a number of LFG researchers, I will adopt it as a working assumption here, although none of what follows is dependent upon it, as far as I can see. I will show temporal and related information represented at f-structure, but there is no reason, to my knowledge, why some of this information could not be “shared out” (e.g. some of it could appear at c-structure, provided that c-structure is visible to whatever procedure is responsible for mapping to a semantic formalism). Discussion of this matter at the workshop revealed that a number of participant researchers would wish to argue for this latter stance. I have no objection to it, but since it appears to be less straightforward than the "f-structure" assumption, I will adopt the latter in what follows. Further research is needed to establish which approach is the more useful.

A second working assumption I will make is that temporal (and related) information should be encoded in as language-independent a way as possible. This accords with the general spirit of LFG, and would make the representation suitable as input to, for example, structural transfer for machine translation, or to mapping to a semantic formalism such as Underspecified Discourse Representation Theory (UDRT). (See, for example, van Genabith and Crouch 1997 for details of such a mapping from f-structure to UDRT.)

My current aim, therefore, is to propose a method of encoding at f-structure, in as language-independent a way as possible, all the information that is pertinent to the temporal interpretation of discourse.

I will now turn to question (i), to consider how best this might be done.

2. The contribution of temporal information to the temporal interpretation of discourse

By 'the temporal interpretation of discourse' I mean such things as determining the temporal relations (succession, inclusion, overlap, etc) between events and states as presented by the text. Of course, a vast amount of work has been done on this subject (see Kamp and Reyle 1993 for a comprehensive summary), and I cannot hope to do justice to it here. Rather than try to do so, I will simply highlight a few significant findings, and draw out some implications for possible f-structure representation.

Since what I am calling 'temporal information' is often conveyed by tense and aspect features, I will from now on refer to such information as 'tense/aspect related' (or 'TA-related') information. It should be borne in mind, however, that across the languages of the world, TA-related information can be conveyed in a wide variety of ways (e.g. tense markers may be placed on NPs in some languages[1]), so I am not intending to set any limits on what can count as TA-related information.

2.1 Narrative movement

It is well known that in multi-sentence sequences, a series of "events" is generally interpreted differently from a series containing "states". For example, in the following sequence:

(1) Mary got up. She brushed her teeth.

[1] See, for example, Nordlinger and Sadler 2000.

we have two events, and the most natural interpretation is one where Mary's teeth-brushing follows her getting up. If, however, we replace the second sentence by a stative:

(2) Mary got up. She felt very ill.

then a second interpretation is available, where Mary's feeling ill begins before she gets up and continues afterwards. Another possible interpretation is that she begins to feel ill just as, or shortly after, she gets up.[2]

If the second sentence is a progressive, then only the "overlap" reading is possible:

(3) Mary got up. She was brushing her teeth.

can only mean that Mary is already brushing her teeth when she gets up.

Thus progressives are similar but not identical to lexical statives in terms of the temporal relations that they allow. Because of the similarity, the English progressive has been taken by some to be a stativiser (Vlach 1981, Moens and Steedman 1988). There are good arguments against taking this approach, however (see Smith 1991 and Glasbey 1994) and we will not adopt it in this paper (see Section 2.2 for further discussion).

Roughly speaking, we can say that two successive sentences that are non-stative and have simple (non-progressive) aspect give rise to a "forward movement" or "temporal succession" interpretation, whereby the event described second follows the one described first.[3] If the second sentence is a progressive, then only the temporal overlap interpretation is available. If the second sentence is stative (with simple aspect), then either of these interpretations is possible.

Clearly, then, in our TA representation we will need to include stative/non-stative and progressive/simple aspect information. In fact, since other work[4] has shown that the determination of temporal structure depends not only upon these distinctions but on Vendler class (the distinction between states, activities, accomplishments and achievements, as proposed in Vendler 1967), then it appears that we may need to encode Vendler class (otherwise known as aspectual class) in our representations.

The presence of such aspectual information at f-structure will be necessary in order to enable mapping to a UDRS, or to some other semantic formalism which similarly places restrictions on possible temporal relations between events and states.

[2] Of course, world knowledge or contextual knowledge often helps to select whichever interpretation is most salient in a given context.

[3] Lascarides and Asher 1993 argue that the forward movement or "narrative progression" interpretation should be seen as a default, to be inferred in the absence of other information. Alternative interpretations, such as unordered lists and elaborations, are possible if certain other conditions are met.

[4] See, Kamp and Reyle 1993 and Glasbey 1994, for example, for summaries.

A preliminary attempt at a suitable f-structure for example (1) might look as follows:

Mary got up:

$$\left[\begin{array}{ll}
\text{PRED} & \text{'get_up<SUBJ>'} \\
\text{TIME_REF} & \text{PAST} \\
\text{STATIVE} & - \\
\text{ASP_CLASS} & \text{ACHIEVEMENT} \\
\text{SUBJ} & \left[\begin{array}{l} \text{PRED 'Mary'} \\ \text{........} \end{array}\right]
\end{array}\right]$$

She brushed her teeth:

$$\left[\begin{array}{ll}
\text{PRED} & \text{'brush<SUBJ, OBJ>'} \\
\text{TIME_REF} & \text{PAST} \\
\text{STATIVE} & - \\
\text{ASP_CLASS} & \text{ACCOMPLISHMENT} \\
\text{SUBJ} & \left[\begin{array}{l} \text{PRED pronoun} \\ \text{.........} \end{array}\right] \\
\text{OBJ} & \left[\begin{array}{l} \text{PRED 'teeth'} \\ \text{..........} \end{array}\right]
\end{array}\right]$$

NB: I have ignored the possessive pronoun in this representation.

Assuming a system that maps from these f-structures to an appropriate semantic formalism such as UDRT, the respective values of the ASP_CLASS features of the two clauses (ACHIEVEMENT and ACCOMPLISHMENT) would contribute to determining the temporal relationship between the corresponding event referents in the combined UDRS for the discourse.

There is a problem with the above, however, in that the determination of sentencial or clausal aspect is compositional (see, for example, Verkuyl 1972, Krifka 1992, Verkuyl 1993). That is, the aspectual class of the sentence 'Mary brushed her teeth' is not determined solely by the aspectual class of the verb 'brush'. Other constituents, such as the subject and object arguments and any prepositional and adverbial phrases, may also have an effect on the aspectual class of the sentence as a whole.

In some accounts (e.g. Moens and Steedman 1988), the lexical entry for a verb is given a "basic" aspectual class, which may then be modified in certain ways by other constituents. So we might classify the verb 'eat' as an activity. If this combines with an "unbounded"[5] object argument such as 'bread', then the aspectual class of the VP 'ate bread' remains an activity. If the object NP is bounded (e.g. 'an apple'), then the VP 'ate an apple' will be an accomplishment. The aspectual class of the sentence as a whole is only determined when we additionally take into account the properties of the subject NP. If the subject NP is bounded (e.g. 'John') then the aspectual class of the sentence 'John ate an apple' is an accomplishment. If the subject NP is unbounded (e.g. 'people') then the sentence as a whole is an activity.

Prepositional phrases (etc) may also have an effect, so that while 'John climbed' is an activity, 'John climbed to the top' is an accomplishment.

Deriving aspectual class in a non-monotonic way like this would be a problem for deriving f-structures in the standard information-preserving manner employed in LFG. It is possible, however, to see the determination of sentential/clausal aspectual class as an information-preserving process. Viewed this way, sentence aspectual class is composed from information contributed by the various constituents. The verb, the subject NP, the object NP(s), and any prepositional and adverbial phrases each contribute aspectual information, which is combined according to compositional rules to give the aspectual class of the whole sentence. This type of approach has been used primarily by Verkuyl (1972, 1993) and is also employed by Krifka (1992) and others. It has been embodied in a number of implementations, including a feature-based Prolog grammar described in Glasbey 1994. Rather than assigning a basic aspectual class to the verb, which is then modified by other constituents, a feature called TERM(INATIVE) is used. This means roughly the same as 'bounded', and the feature make take the value +, – or "unspecified". Examples of verb classifications are:

'eat': UNSPECIFIED
'stroke': –TERM
'like': –TERM

A verb like 'stroke' which is –TERM will always combine with its arguments to give a –TERM VP, irrespective of the TERM properties of those arguments. A verb which is unspecified for TERM will combine to give a VP (and finally a sentence) the value of whose TERM feature is determined by the TERM values of its object and subject

[5] Various related terms are used in the literature. An closely-related alternative (used in Krifka 1992) is 'cumulative'.

arguments and any prepositional phrases (etc). Thus, for example, 'John ate an apple' would be +TERM because 'ate' has its TERM value unspecified, 'John' is +TERM and 'an apple' is +TERM.[6]

Clearly, across different languages the information contributed to overall sentence aspectual class by the various sentential constituents would vary (i.e. the mapping from c-structure to f-structure would differ). The aim, however, would be to produce a fully language-independent set of aspectual features. Much more work is needed, of course, to develop such a set of features.[7]

Examples of f-structures based on the TERM feature will be given at the end of Section 2. I will now move on to consider the progressive and related constructions.

2.2 Progressives, imperfectives and other unfinished business

We have already looked briefly at some of the properties of the English progressive and seen that it affects the temporal properties of discourse. It seems reasonable, therefore, to include a feature +/–PROG in the f-structure for the VP. This is problematic, however, given our ultimate aim of language-independence, because the English progressive has a number of properties which distinguish it from imperfectives in other languages.

One very obvious difference is that English progressives combine only reluctantly with stative verbs (? 'Mary was loving John') whereas in French, for example, the imparfait combines readily with stative verbs ('Marie aimait Jean').

There are implications here for aspectual composition and discourse structure, so we need to be careful not to use a single feature PROG cross-linguistically. An alternative suggestion is given below.

First, we should consider in more depth the proposal (mentioned earlier) made by some researchers to treat the English progressive as a stativiser. Moens and Steedman (1988) are among those who adopt this approach, but it has been shown to be unsatisfactory in a number of respects. Smith (1991) gives a range of cross-linguistic data which shows that the English progressive is not equivalent to the cross-linguistic property "imperfective".[8] Treating progressives as statives causes problems in English, too, as is shown by Glasbey

---

[6] It should be noted that the issue has been somewhat simplified here for the sake of exposition. Care needs to be taken, for example, with scope ambiguities in sentences like 'Everyone climbed the mountain in an hour'.

[7] It should, however, be noted that Smith (1991) bases her account of aspectual composition (which is closely related to the one just given) on data from a range of languages including English, French, Russian, Mandarin Chinese and Navajo – so at least some first steps towards language-independence have been made.

[8] There is, of course, a major assumption here that the notion "imperfective" across languages can be captured by a single feature.

(1998) in an analysis of ‘at the time’ and ‘at the same time’. The data in question is as follows:

4. Daniel climbed Ben Nevis.

5. # Jonny climbed Snowdon at the time.

6. Jonny was climbing Snowdon at the time.

7. Jonny was a young boy at the time.

8. Jonny climbed Snowdon at the same time.

9. Jonny was climbing Snowdon at the same time.

10. # Jonny was a young boy at the same time.

Sentence (4) is to be followed by one of sentences (5) – (10). It will be seen that with ‘at the same time’ the progressive in (9) does not pattern with the stative in (10). A detailed analysis is given in Glasbey 1998 – here I will simply say that in order to account for the data, it is necessary to distinguish progressives from (lexical) statives, and this runs counter to an account that treats the progressive as a stativiser.

Given that (i) it appears to be unsatisfactory to treat progressives as states, and (ii) it appears unsatisfactory to treat the English progressive in the same way as a cross-linguistic imperfective, I propose to use two features. One is PROG, which may take values of {+/–} and will be used for the English progressive (and any equivalent constructions found in other languages). The second feature, which I will call VIEWPOINT, after Smith 1991, is based on the PERFECTIVE/IMPERFECTIVE distinction, used widely in studies of the Slavic languages and others, and adopted by Smith (1991) in her cross-linguistic analysis of aspect.[9]

2.3 Perfects

Another TA feature we will need to encode at f-structure is PERF(ECT), as in:

(11) Em has climbed Ben Nevis.

The need for this feature seems uncontroversial, as the perfect (in English and a range of other languages) clearly contributes to the temporal structure of discourse. I will assume

[9] In fact Smith proposes a third value “NEUTRAL”, which she argues is needed for some of the languages in her analysis. I will retain an open mind here on whether this third value will be required. Once again, further research is needed.

here for simplicity that a single PERF feature (taking values +/–) is sufficient cross-linguistically, although this may turn out not to be the case.

2.4 Time reference: past, present and future

Up until now I have concentrated on aspectual features as opposed to tense ones. Of course, many languages have constructions which encode tense and aspect features simultaneously, and sometimes temporal reference is underspecified. It is clear, however, that we will need some way of expressing, at least, whether the temporal reference of a sentence is PAST, PRESENT or FUTURE (or some underspecified combination of these values). I propose therefore, to incorporate a feature TIME_REF which can take the above values.

2.5 Summary: Proposed TA-related features for f-structure

In sections 2.1 – 2.4, I have proposed a number of TA-related features that will be needed in our f-structures. These are STATIVE {+/–}, TERM {+/–}, PROG {+/–}, VIEWPOINT {PERFECTIVE/IMPERFECTIVE}, PERFECT {+/–} and TIME_REF {PAST/PRESENT/FUTURE}.

In order to illustrate the use of the above features, I will now give a number of lexical entries for verbs, and an f-structure for an associated sentence, to indicate how the required feature composition is achieved.

Lexical entries for verbs:[10]

ate :	V($\uparrow$ PRED)	=	‘eat $< (\uparrow \text{SUBJ})(\uparrow \text{OBJ}) >$’
	($\uparrow_{\mu}$ AUX)	=	–
	($\uparrow_{\mu}$ FIN)	=	+
	($\uparrow$ TERM)	=	
	($\uparrow$ PERFECT)	=	–
	($\uparrow$ TIME_REF)	=	PAST
	($\uparrow$ VIEWPOINT)	=	PERFECTIVE
	($\uparrow$ PROG)	=	–
	($\uparrow$ STATIVE)	=	–

Note that the lack of a value for particular feature means that the feature value is unspecified.

[10] I use LFG terminology employing a level of morphological representation (m-structure) here, following Butt et al 1996.

stroked:	V$(\uparrow \text{PRED})$	$= \text{'stroke} < (\uparrow \text{SUBJ})(\uparrow \text{OBJ}) >\text{'}$
	$(\uparrow_{\mu} \text{AUX})$	$= -$
	$(\uparrow_{\mu} \text{FIN})$	$= +$
	$(\uparrow \text{TERM})$	$=$
	$(\uparrow \text{PERFECT})$	$= -$
	$(\uparrow \text{TIME\_REF})$	$= \text{PAST}$
	$(\uparrow \text{VIEWPOINT})$	$= \text{PERFECTIVE}$
	$(\uparrow \text{PROG})$	$= -$
	$(\uparrow \text{STATIVE})$	$= -$
eating:	V$(\uparrow \text{PRED})$	$= \text{'eat} < (\uparrow \text{SUBJ})(\uparrow \text{OBJ}) >\text{'}$
	$(\uparrow_{\mu} \text{AUX})$	$= +$
	$(\uparrow_{\mu} \text{FIN})$	$= +$
	$(\uparrow \text{TERM})$	$=$
	$(\uparrow \text{PERFECT})$	$=$
	$(\uparrow \text{TIME\_REF})$	$=$
	$(\uparrow \text{VIEWPOINT})$	$= \text{IMPERFECTIVE}$
	$(\uparrow \text{PROG})$	$= +$
	$(\uparrow \text{STATIVE})$	$= -$
was: V_{aux}	$(\uparrow_{\mu} \text{AUX})$	$= +$
	$(\uparrow_{\mu} \text{FIN})$	$= +$
	$(\uparrow_{\mu} \text{DEP AUX})$	$= -$
	$(\uparrow_{\mu} \text{DEP VFORM})$	$= \text{PROG}$
	$(\uparrow \text{PERFECT})$	$= -$
	$(\uparrow \text{TIME\_REF})$	$= \text{PAST}$
	$(\uparrow \text{VIEWPOINT})$	$= \text{IMPERFECTIVE}$
	$(\uparrow \text{PROG})$	$= +$

F-structure for ‘Mary was eating a cake’:

$$\left[\begin{array}{ll}
\text{PRED} & \text{‘eat}\langle\text{SUBJ, OBJ}\rangle\text{’} \\
\text{TIME_REF} & \text{PAST} \\
\text{STATIVE} & - \\
\text{PROG} & + \\
\text{TERM} & + \\
\text{VIEWPOINT} & \text{IMPERFECTIVE} \\
 & \\
\text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘Mary’} \\ \text{TERM} & + \\ \ldots\ldots\ldots & \end{array}\right] \\
 & \\
\text{OBJ} & \left[\begin{array}{ll} \text{PRED} & \text{‘cake’} \\ \text{TERM} & + \\ \ldots\ldots\ldots. & \end{array}\right]
\end{array}\right]$$

3. Additional TA-related features

I have sketched a set of possible TA-related features and shown how they might be used to obtain compositionally-derived f-structures. Before closing, I would like to consider whether there are any additional TA-related features that should be included.

One clear case seems to be the need to encode in some form the distinction between accomplishments and achievements (or some closely-related distinction, as explained below). This need is evident when we look at English sentences containing ‘when’ constructions. Sandström (1993) draws attention to the following type of example:

(12) When Mary arrived at the station she bought her ticket.

(13) ?When Mary drove to the station she bought her ticket.

Sandström points out that in (12), the relation between the arriving event and the ticket-buying event is one of temporal succession. The effect is similar to that conveyed by the two-sentence sequence:

(14) Mary arrived at the station. She bought her ticket.

In (13), however, it is very difficult, if not impossible, to get this interpretation, and (13) sounds rather strange as a result. (Another reading, perhaps best described as an "occasion" reading[11], can be obtained by some speakers, but the temporal succession reading is clearly absent.)

Sandström identifies the difference between (12) and (13) as being due to the fact that 'Mary arrived at the station' is an achievement while 'Mary drove to the station' is an accomplishment.

Glasbey (1995, 2001) argues that the crucial distinction is not one of "punctuality" (the standard difference between achievements and accomplishments) but rather involves properties of the subject NP. Glasbey's arguments are based in part on examples such as:

(15) When the soup cooled down we drank it.

And:

(16) When the snow melted away we went for a walk.

The standard tests for aspectual class (see Dowty 1979) classify 'the soup cooled down' and 'the snow melted away' as accomplishments – but the temporal succession reading is readily available. Glasbey (2001) proposes that the criterion for the temporal succession reading in this construction is a "thematic" property of verb's subject NP argument, and concerns whether the subject is described as undergoing a change of state. This property, which she calls SUBJECT STATE TRANSITION or SST, can be seen as a feature of the subject NP, which is placed there by a property of the verb which is clearly distinct from considerations of punctuality.

I therefore propose an additional feature, called SST {+/–}, to be placed on the subject NP of the sentence, the value of the feature being derived from a lexical feature of the verb. The verb 'arrive', for example, which has the value +SST, would have the following feature specification as part of its lexical entry:

$(\uparrow \text{SST}) \;=\; +$

while, in contrast, a verb like 'eat', which has the value –SST, would have the feature specification:

$(\uparrow \text{SST}) \;=\; -$

The value of the SST feature would then be carried up to clause level.

It is interesting to note that here we have a feature which intuitively belongs to the subject NP and yet affects the TA properties of the clause as a whole. This is reminiscent in some

[11] A paraphrase of this reading would be 'On the occasion that Mary drove to the station she bought her ticket'.

ways of the way that tense information in some languages is conveyed (at least in part) by tense markers placed on NPs.[12] One noteworthy difference, however, is that in the SST case, the value of the SST feature for the clause is derived from the SST value of the verb, rather than some feature marked on the subject NP. It would be interesting to see if there are any languages that mark SST (or similar) information either fully or partially by inflection of the subject NP.

Again, further work is needed, but we have shown, at least, that the list of TA-related features suggested above is incomplete. It seems highly likely that more such TA-related features will be uncovered. The challenge is, as I said earlier, to make the set of TA-related features as language-independent as possible, while, of course, encompassing all the data.

4. Conclusion

I have shown that in a feature-based linguistic framework such as LFG, in order to determine the temporal structure of discourse we need to identify a set of cross-linguistically valid TA-related features.

I have suggested that in LFG it may be appropriate for such a set of features to be encoded at f-structure. This proposal is, however, based on the assumption, stated earlier, that f-structure constitutes the sole input to semantic interpretation. This assumption appears to be somewhat controversial in the LFG community, and clearly further work is needed to clarify and resolve the matter.

Having reviewed the effects of a range of TA-related features on the temporal structure of discourse, I have proposed a set of such features that aims as far as possible to be language-independent (and unquestionably falls short of that aim). I have also outlined briefly how the required f-structures could be derived from lexical entries for verbs and other constituents, in an information-preserving manner.

Much further research is needed, of course, both to work out the fine details of the above sketch, and to establish, as far as possible, a fully language-independent set of TA-related features that will do the work required for discourse interpretation, translation between languages, and whatever else we may ask of it.

[12] An example is the Arawak language Chamicuro, spoken in Peru (discussed in Nordlinger and Sadler 2000), where clausal tense information is encoded on the definite article.

References

Butt, M., Niño, María-Eugenia, Segond, Frédérique (1996). Multilingual processing of auxiliaries within LFG. *In* D. Gibson (ed) *Natural Language Processing and Speech Technology: Results of the 3rd Konyens Conference*, pp 111-122. Mouton de Gruyter.

Dowty, D. (1979). *Word Meaning and Montague Grammar.* Dordrecht : D. Reidel.

Frank, A., Zaenen, A. (draft). Tense in LFG: Syntax and morphology. Xerox Research Centre Europe, 6 chemin de Maupertuis, 38240 Meylan, France.

Glasbey, S. (1990). *Tense and Aspect in Natural Language Processing – A Situation Theory Approach.* MSc thesis, Department of AI, University of Edinburgh.

Glasbey, S. (1994). *Event Structure in Natural Language Discourse.* PhD Thesis, Centre for Cognitive Science, University of Edinburgh.

Glasbey, S. (1995). 'When', discourse relations and the thematic structure of events. Workshop notes of the 5th International Workshop TSM'95, Château de Bonas 23-27 June 1995. Institut de Recherche en Informatique de Toulouse – Université Paul Sabatier, INPT and CNRS. Equipe de Recherche en Syntaxe et Sémantique – Université Toulouse-Le-Mirail and CNRS.

Glasbey, S. (2001) Event structure, punctuality and 'when'. Draft, University of Dundee.

Glasbey, S. (1998). Progressives, states and backgrounding. In S. Rothstein, ed, *Events and Grammar*, pp. 105-124. Kluwer Academic Publishers, Dordrecht/Boston/London.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic: Introduction to the Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Dordrecht: Kluwer.

Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In Sag, I., Szabolsci, A. (eds), *Lexical Matters.* CSLI publications, Stanford, Ca, USA.

Lascarides, A., Asher, N. (1993). Temporal interpretation, discourse structure and commonsense entailment. *Linguistics and Philosophy* 16 pp 437-493.

Moens, M. and M. J. Steedman (1988). Temporal ontology and temporal reference. *Computational Linguistics* 14(2), 15-28.

Nordlinger, R., Sadler, L. (2000). Tense as a nominal category. *In* Butt, M. and Holloway King, T. (eds), *Proceedings of the LFG00 Conference,* University of California, Berkeley: CSLI Publications.

Sandström, G. (1993). When-clauses and the temporal interpretation of narrative discourse. *Technical Report 34, Department of General Linguistics, University of Umeå.*

Smith, C. (1991). *The Parameter of Aspect.* Kluwer Academic Publishers, Dordrecht.

Van Genabith, J., Crouch, R. (1997). On interpreting F-structures as UDRSs. In: ACL-EACL-97, Madrid, Spain, Proceedings of the Conference, 1997, pp. 402--409.

Vendler, Z. (1967). Verbs and Times. In *Linguistics in Philosophy*, pp 97-121. Ithaca, N.Y: Cornell University Press.

Verkuyl, H. J. (1972). *On the Compositional Nature of the Aspects.* Dordrecht: D. Reidel.

Verkuyl, H. J. (1993). *A Theory of Aspectuality. The Interaction between Temporal and Atemporal Structure.* Cambridge Studies in Linguistics, Vol 64. Cambridge: Cambridge University Press.

Vlach, F. (1981). The semantics of the progressive. In P. J. Tedeschi and A. E. Zaenen, eds., *Syntax and Semantics,* Vol 14: *Tense and Aspect*, pp 271-292. New York: academic Press.