Filling the Gap

Agnieszka Patejuk
Institute of Computer Science, Polish Academy of Sciences

Adam Przepiórkowski
Institute of Computer Science, Polish Academy of Sciences
Institute of Philosophy, University of Warsaw

Proceedings of the LFG'17 Conference

University of Konstanz

Miriam Butt, Tracy Holloway King (Editors)

2017

CSLI Publications

pages 327–347

http://csli-publications.stanford.edu/LFG/2017

Keywords: gapping, English, Polish, agreement

Patejuk, Agnieszka, & Przepiórkowski, Adam. (2017). Filling the Gap. In Butt, Miriam, & King, Tracy Holloway (Eds.): *Proceedings of the LFG'17 Conference, University of Konstanz* (pp. 327–347). Stanford, CA: CSLI Publications.

Abstract

The paper presents an LFG account of Gapping based on data from English and Polish. The analysis accounts for the requirement that the elided verb be understood as having the same semantic – but not necessarily agreement – features as the overt verb, and for the possibility of mismatches between arguments of the overt and the elided verb.

1 Introduction

There is no worked-out LFG analysis of Gapping. It is – by design – not dealt with in Maxwell and Manning 1996, which provides an analysis of some other kinds of non-constituent coordination (using mechanisms which go beyond standard LFG), and just a very general basic idea is suggested in Kaplan 1987, 1995 based on a new operator – priority union. The aim of this paper is to fill this gap. The proposed analysis does not assume any non-standard mechanisms and it uniformly deals with standard cases of Gapping (as in (1)–(2)), as well as other phenomena sometimes analysed as subspecies of Gapping, including Conjunction Reduction (as in (3)).

(1) Marge gave an apple to Lisa, and Homer a donut to Bart.
(2) Marge gave an apple to Lisa, and Homer to Bart.
(3) Marge gave an apple to Lisa and a banana to Bart.

The main analysis is developed on the basis of English, but more complex interactions are illustrated on the basis of Polish, a free word order language with interesting valency and case assignment phenomena.

2 Basics

2.1 Data

A property distinguishing Gapping from many other ellipsis-like phenomena is that a verb is completely elided, rather than being replaced by a verbal pro-form like *do (so)*. This property also sets apart the examples in (1)–(3) above from, e.g., Right Node Raising (cf. (4)) and VP Ellipsis (cf. (5)):

(4) Marge gave an apple, and Homer wanted to give a donut, to their daughter Lisa. (RNR)
(5) Marge has just given Lisa a kiss, and Homer will, too. (VPE)

Another property of Gapping is that the elided verb must be understood as having the same semantic and information-structural features as the overt verb, including tense and voice. No such constraints hold, e.g., in the case of VP Ellipsis: (5) above involves different tense values, and (6) below (cited as attested in Dalrymple et al. 1991: 440) – different voice values.

[†]We are grateful for comments to Emily Bender, Dan Flickinger and Ron Kaplan, as well as the two anonymous reviewers – they led to various improvements in the form and content of this paper. The research reported here is partially supported by the Polish Ministry of Science and Higher Education within the CLARIN ERIC programme 2016–2018 (http://clarin.eu/).

(6) A lot of this material can be presented in a fairly informal and accessible fashion, and often I do.

Attempts to violate these constraints on Gapping result in ungrammaticality:[1]

(7) Marge gave an apple to Lisa yesterday, and Homer *(will give) a banana to Bart tomorrow.

(8) Lisa was given an apple by Marge and Homer *(gave) a banana to Bart.

This does not, however, mean that the elided verb, if present, must have the same form as the overt verb – they may differ in agreement features, as in the following example from Johnson 2014: 4:

(9) He likes beans and you (like) rice.

More examples of this kind will be given in §4 below, in the context of Polish.

2.2 Analysis

A rule like (10) is usually assumed for sentential coordination in LFG (Kaplan and Maxwell 1988: 304, Dalrymple 2001: 362, Peterson 2004: 652, etc.).

(10) IP → IP Conj IP
$\downarrow \in \uparrow$ $\downarrow \in \uparrow$

At the functional level it creates a set with two f-structures corresponding to the two conjoined sentences. Given that in the second IP on the right hand side of this rule the verb is missing in Gapping constructions, the pleasing symmetry between the two IP constituents in (10) cannot be maintained: the value of PRED in the second f-structure must originate in the first f-structure.

The analysis proposed here generalises and appropriately encodes the intuition that functional annotations pertaining to the first conjunct may optionally hold for the whole set and, hence, spread to other conjuncts. A coordination rule that handles Gapping is given in (11).

(11) IP → IP1 [Comma IP]* Conj IP
$\uparrow = \downarrow$ $\downarrow \in \uparrow$ $\downarrow \in \uparrow$
$(\downarrow \text{LOCAL}) \in \uparrow$

This rule trivially extends the previous version to the case of possibly more than two conjuncts and – less trivially – treats the first of these (IP1) in a special way to be explained briefly.

As more complex interactions will be illustrated with Polish, we assume here the following c-structure rules which are a part of a large implemented grammar of Polish (Patejuk and Przepiórkowski 2012b, 2014c, 2017).

(12) IP → DEP*, (I)

(13) DEP ≡ NP | InfP | PP | ...
$(\uparrow \text{SUBJ}) = \downarrow$ $(\uparrow \text{XCOMP}) = \downarrow$ $\downarrow \in (\uparrow \text{ADJ})$
$(\downarrow \text{CASE}) =_c \text{NOM}$ ... ...

[1] As usual, an asterisk in front of parentheses means that optionality is ruled out, i.e., that only a version of the sentence with the material in the parentheses is grammatical.

As usual, the lack of any functional schemata below a right hand side (RHS) category is understood as a shorthand for the single "head equation" $\uparrow{=}\downarrow$. This means that all constituents in (12) are co-heads of the IP: this holds for the optional verb, I, and any number of dependents. Note also the comma in the RHS of (12), which signals an arbitrary order of the constituents. Since Polish is a so-called free word order language, i.e., a language where word order is to a large extent determined by information structure, the subject is treated just like any other dependent – there is no need to split an IP first into the subject and an I′, and then the I′ into an I and any postverbal dependents, as is normally done in English. But nothing in the basic analysis depends on the flat structure of clauses produced by (12), and the proposed analysis is compatible with a more hierarchical structure and strict linearisation, as usually assumed for English (but see §5.1 for an aspect of Gapping where the flat structure becomes crucial).

Possible dependents are specified in (13) in the usual way: a nominative NP is a possible subject, an infinitival phrase is a candidate for an XCOMP value, etc. What is perhaps slightly unusual is the splitting of the IP rule into the two rules (12)–(13), which jointly have the same effect as the following single rule:

(14) $\begin{array}{lcccccc} \text{IP} & \rightarrow & \text{NP}, & \text{InfP}, & \text{PP*}, & \ldots, & \text{(I)} \\ & & (\uparrow \textsc{subj}) = \downarrow & (\uparrow \textsc{xcomp}) = \downarrow & \downarrow \in (\uparrow \textsc{adj}) & & \\ & & (\downarrow \textsc{case}) =_c \textsc{nom} & \ldots & \ldots & & \end{array}$

The advantage of splitting this rule into two is that the definition of possible dependents in (13) may be reused, in particular in the following rule for IP1:

(15) $\begin{array}{lcll} \text{IP1} & \rightarrow & \text{DEP*}, & \text{I} \\ & & (\uparrow (\textsc{local})) = \downarrow & \end{array}$

One simple difference between the rule for the initial IP1 in (15) and the rule for any non-initial IP in (12) is that the I, i.e., the finite verb, is obligatory in IP1, while it is optional in further IPs, which may be gapped. The other difference is the presence of the optional attribute LOCAL in the functional annotation on DEP. The effect of this optional attribute in (15) is that each dependent either contributes to the top f-structure of IP1 (if the annotation $(\uparrow (\textsc{local})) = \downarrow$ is resolved to $\uparrow{=}\downarrow$) or to the value of LOCAL, assumed to be a non-distributive feature (if the annotation is resolved to $(\uparrow \textsc{local}) = \downarrow$). The procedural intuition is that LOCAL at this stage contains exactly those pieces of the first conjunct which do not distribute to other conjuncts. If this analysis were to be applied to example (1), repeated below, the f-structure for the first conjunct would at this stage look as in (16), on the assumption that the functional annotation resolves to $(\uparrow \textsc{local}) = \downarrow$ for each of the three dependents.[2]

(1) Marge gave an apple to Lisa, and Homer a donut to Bart.

[2] In all AVMs corresponding to the running example, the values of LOCAL are marked as $\boxed{0}$.

(16)
$$\begin{bmatrix} \textsc{pred} & \text{`}\textsc{give}\text{<}\textsc{subj},\textsc{obj},\textsc{obl}\text{>'} \\ \textsc{tense} & \textsc{past} \\ \textsc{voice} & \textsc{active} \\ \textsc{local} & \boxed{0}\begin{bmatrix} \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Marge}\text{'} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{apple}\text{'} \end{bmatrix} \\ \textsc{obl} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Lisa}\text{'} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Given the three dependents, there are $2^3 = 8$ possibilities of resolving the three $(\uparrow\ (\textsc{local})) = \downarrow$ functional equations, of which (16) is just one. Another one could potentially be (17), but it would not lead to a successful analysis of (1).

(17)
$$\begin{bmatrix} \textsc{pred} & \text{`}\textsc{give}\text{<}\textsc{subj},\textsc{obj},\textsc{obl}\text{>'} \\ \textsc{tense} & \textsc{past} \\ \textsc{voice} & \textsc{active} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{apple}\text{'} \end{bmatrix} \\ \textsc{local} & \begin{bmatrix} \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Marge}\text{'} \end{bmatrix} \\ \textsc{obl} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Lisa}\text{'} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

What is important is that only the DEP constituents may have their f-structures put into the value of LOCAL, while the verb in I is implicitly annotated with the head equation $\uparrow = \downarrow$, i.e., necessarily contributes to the top level of the structure. Moreover, it contributes all the features – not just PRED, but also other attributes defined in the lexical entry of the verb, including TENSE and VOICE (or similar attributes). This naturally accounts for the fact mentioned above that the elided verb not only takes over the basic meaning of the overt verb (i.e., its PRED), but also any other semantic and information-structural features of a given verb form, including its tense and voice.

Let us now consider the coordination rule (11), repeated below:

(11) IP → IP1 [Comma IP]* Conj IP
$\uparrow = \downarrow$ $\downarrow \in \uparrow$ $\downarrow \in \uparrow$
$(\downarrow \textsc{local}) \in \uparrow$

While all non-initial IPs contribute their f-structures to the resulting set, as standard in LFG, in the case of IP1 it is the LOCAL value which is defined as the member of the whole set, via the $(\downarrow \textsc{local}) \in \uparrow$ functional annotation.

In the case of the running example (1), speaking procedurally again, this results in the set in (18), where the first element is the value of LOCAL of the first conjunct (cf. (16) above) and the second element is the f-structure of the second conjunct.

(18)
$$\left\{ \boxed{0}\begin{bmatrix} \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Marge}\text{'} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{apple}\text{'} \end{bmatrix} \\ \textsc{obl} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Lisa}\text{'} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Homer}\text{'} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{donut}\text{'} \end{bmatrix} \\ \textsc{obl} & \begin{bmatrix} \textsc{pred} & \text{`}\textsc{Bart}\text{'} \end{bmatrix} \end{bmatrix} \right\}$$

Further, according to the other annotation on IP1, ↑=↓, this set is unified with the whole f-structure of IP1, in effect distributing all distributable features of IP1 to all elements of the set. In the case of the running example, the unification of the set in (18) with the top f-structure for the first conjunct in (16) results in the final hybrid f-structure for the coordinated IP shown in (19).

(19)
$$\left[\begin{array}{l}\left\{\left(\boxed{0}\left[\begin{array}{ll}\textsc{pred} & \text{`}\textsc{give}\text{<}\boxed{1},\boxed{2},\boxed{3}\text{>'}\\ \textsc{subj} & \boxed{1}\left[\textsc{pred}\quad \text{`}\textsc{Marge}\text{'}\right]\\ \textsc{obj} & \boxed{2}\left[\textsc{pred}\quad \text{`}\textsc{apple}\text{'}\right]\\ \textsc{obl} & \boxed{3}\left[\textsc{pred}\quad \text{`}\textsc{Lisa}\text{'}\right]\\ \textsc{tense} & \textsc{past}\\ \textsc{voice} & \textsc{active}\end{array}\right],\left[\begin{array}{ll}\textsc{pred} & \text{`}\textsc{give}\text{<}\boxed{4},\boxed{5},\boxed{6}\text{>'}\\ \textsc{subj} & \boxed{4}\left[\textsc{pred}\quad \text{`}\textsc{Homer}\text{'}\right]\\ \textsc{obj} & \boxed{5}\left[\textsc{pred}\quad \text{`}\textsc{banana}\text{'}\right]\\ \textsc{obl} & \boxed{6}\left[\textsc{pred}\quad \text{`}\textsc{Bart}\text{'}\right]\\ \textsc{tense} & \textsc{past}\\ \textsc{voice} & \textsc{active}\end{array}\right]\right)\right\}\\ \textsc{local}\ \boxed{0}\end{array}\right]$$

Apart from the presence of the technical attribute LOCAL, this is the desired representation of (1). It would be easy to get rid of this attribute by using the restriction operator and replacing the functional equation ↑=↓ under IP1 in (11) with ↑=↓/LOCAL, but this attribute will play a role in the analysis of agreement, so we leave it there.

An important aspect of this analysis is that the information introduced by the verb in the first conjunct is distributed to all conjuncts simultaneously. For this reason, different values of PRED in particular conjuncts – namely, values differing in what arguments fill particular argument positions – are fine, as long as they are all subsumed (see, e.g., Maxwell and Manning 1996: 11) by the value specified in the lexical entry of the verb (e.g., by (↑ PRED) = 'GIVE<SUBJ,OBJ,OBL>').

Out of the 8 possibilities of analysing the first conjunct of (1), where each of SUBJ, OBJ and OBL is present either at the top level or within the value of LOCAL, only the analysis illustrated in (16), where all of them are put within LOCAL, guarantees the successful analysis of the whole sentence. If any of these three attributes were instead present at the top level, as is OBJ in (17), they would distribute to the other conjunct, which would result in a feature clash, as the other conjunct already contains the values of all three attributes (see the second element of the set in (18)). However, (17) is exactly the right analysis of the first conjunct in the case of example (2), repeated below, where not only the verb is elided, but also the direct object:

(2) Marge gave an apple to Lisa, and Homer to Bart.

In this case, the following two elements will initially be contributed to the coordination set by the membership (∈) statements in the coordination rule (11):

(20)
$$\left\{\left(\boxed{0}\left[\begin{array}{ll}\textsc{subj} & \left[\textsc{pred}\quad \text{`}\textsc{Marge}\text{'}\right]\\ \textsc{obl} & \left[\textsc{pred}\quad \text{`}\textsc{Lisa}\text{'}\right]\end{array}\right]\right),\left[\begin{array}{ll}\textsc{subj} & \left[\textsc{pred}\quad \text{`}\textsc{Homer}\text{'}\right]\\ \textsc{obl} & \left[\textsc{pred}\quad \text{`}\textsc{Bart}\text{'}\right]\end{array}\right]\right\}$$

The effect of the unification of this set with the full f-structure for IP1 in (17) will then have the result of distributing not only the values of PRED, TENSE and VOICE, but also the value of OBJ:

(21)
$$\left[\begin{array}{l}\left\{\left(\boxed{0}\left[\begin{array}{ll}\text{PRED} & \text{`GIVE<}\boxed{1}\text{,}\boxed{2}\text{,}\boxed{3}\text{>'}\\ \text{SUBJ} & \boxed{1}\left[\text{PRED}\quad \text{`MARGE'}\right]\\ \text{OBJ} & \boxed{2}\left[\text{PRED}\quad \text{`APPLE'}\right]\\ \text{OBL} & \boxed{3}\left[\text{PRED}\quad \text{`LISA'}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right),\left[\begin{array}{ll}\text{PRED} & \text{`GIVE<}\boxed{4}\text{,}\boxed{2}\text{,}\boxed{5}\text{>'}\\ \text{SUBJ} & \boxed{4}\left[\text{PRED}\quad \text{`HOMER'}\right]\\ \text{OBJ} & \boxed{2}\\ \text{OBL} & \boxed{5}\left[\text{PRED}\quad \text{`BART'}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\right)\right\}\\ \text{LOCAL}\ \boxed{0}\end{array}\right]$$

In a similar fashion, the analysis also deals with Conjunction Reduction, as in (3), repeated below.

(3) Marge gave an apple to Lisa and a banana to Bart.

In this case, the annotation on the DEP corresponding to *Marge* will resolve to ↑=↓, with the effect of the SUBJ value spreading to other conjuncts. This will result in a structure like (22).

(22)
$$\left[\begin{array}{l}\left\{\left(\boxed{0}\left[\begin{array}{ll}\text{PRED} & \text{`GIVE<}\boxed{1}\text{,}\boxed{2}\text{,}\boxed{3}\text{>'}\\ \text{SUBJ} & \boxed{1}\left[\text{PRED}\quad \text{`MARGE'}\right]\\ \text{OBJ} & \boxed{2}\left[\text{PRED}\quad \text{`APPLE'}\right]\\ \text{OBL} & \boxed{3}\left[\text{PRED}\quad \text{`LISA'}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right),\left[\begin{array}{ll}\text{PRED} & \text{`GIVE<}\boxed{1}\text{,}\boxed{4}\text{,}\boxed{5}\text{>'}\\ \text{SUBJ} & \boxed{1}\\ \text{OBJ} & \boxed{4}\left[\text{PRED}\quad \text{`BANANA'}\right]\\ \text{OBL} & \boxed{5}\left[\text{PRED}\quad \text{`BART'}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\right)\right\}\\ \text{LOCAL}\ \boxed{0}\end{array}\right]$$

3 Delimiting Gapping

As demonstrated above, the analysis proposed here does not only deal with the most typical cases of Gapping, such as (1)–(2) above, but also with Conjunction Reduction, sometimes analysed as a special case of Gapping (Neijt 1979, Johnson 2014). On the other hand, Right Node Raising and VP Ellipsis have properties which set them apart from Gapping, including the lack of any constraints on the identity – across conjuncts – of semantic features such as tense and voice. One property encoded in the proposed analysis which distinguishes standard Gapping and Conjunction Reduction on one hand and RNR and VP Ellipsis on the other is that the former – but not the latter – require the absence of the main verb in the non-initial conjunct(s). This follows from the rule (15) for IP1, repeated below, namely, from the fact that the verb, I, is implicitly annotated with the head equation ↑=↓ and, hence, must distribute to other conjuncts. This is only possible when the other conjuncts do not contain their own main verbs.

(15) IP1 → DEP*, I
$(\uparrow (\textsc{local})) = \downarrow$

If the constituent I in (15) were annotated just like DEP, i.e., with the equation $(\uparrow (\textsc{local})) = \downarrow$, the same rule (15) could be used for ordinary sentential coordination: all constituents within IP1 (also the verb) could be analysed as elements of LOCAL, i.e., none would distribute to other conjuncts. However, this positive aspect of this more general analysis would be marred by the fact that the analysis would severely overgenerate. For example, if the verb and almost all dependents were analysed as local to the first conjunct, and only, say, the object were to be distributed across the coordinate structure, the result would be unacceptable:

(23) Marge gave an apple to Lisa and Homer took *(an apple) from Bart.

Hence, as it stands, the IP coordination rule (11) only deals with Gapping, etc., but not with ordinary sentential coordination. However, it is not necessary to formulate a completely separate rule for ordinary coordination; rather, what is needed is an extension of (11) which makes it possible to have an ordinary IP as the first constituent, rather than the Gapping-specific IP1:

(24) IP → (IP | IP1) [Comma IP]* Conj IP
$\downarrow \in \uparrow$ $\downarrow = \uparrow$ $\downarrow \in \uparrow$ $\downarrow \in \uparrow$
$(\downarrow \textsc{local}) \in \uparrow$

Despite the fact that the proposed analysis makes a clear distinction between Gapping (now also including Conjunction Reduction) and other superficially similar phenomena, there are cases where it seems to introduce spurious ambiguities, as in the following example:

(25) Betsy has read a book or a magazine.

The obvious analysis of (25) is as involving a coordination of nominal phrases within the object position of *read*. However, a Gapping analysis of this sentence is also available, on which what is elided is both the verbal material (*has read*) and the subject (*Betsy*).

Johnson (2014) – citing Schwarz 1999 – argues that a Gapping analysis of such an apparent NP-coordination should be made available by the grammar. This claim is based on examples like the following (Johnson 2014: 6):

(26) Either Betsy wanted to read a book or a magazine.

Accepting the analysis of Schwarz 1999, on which *either* marks the start of the first conjunct rather than being displaced from its original location (as previously analysed in Larson 1985), Johnson (2014) concludes that (26) involves sentential coordination, so the second conjunct, *a magazine*, represents the sentence *Betsy wanted to read a magazine*, with the subject and the verbal material elided. If so, also the following sentence, which differs from (26) only in not marking the beginning of the first conjunct, should also have a Gapping analysis in addition to the NP-coordination one (and similarly for (25)):

(27) Betsy wanted to read a book or a magazine.

According to Johnson (2014: 6), there is a semantic reflex of the structural ambiguity of (27). On the NP-coordination analysis, the sentence "describes a particular desire that Betsy has: she desires to read something that is either a book or a magazine". But on the Gapping analysis, (27) is synonymous with the unambiguous (26), "which claims that Betsy has one or the other of two desires… Betsy wanted to read a book or Betsy wanted to read a magazine".

In summary, the proposed account seems to properly delineate Gapping: it only deals with cases where non-initial conjuncts lack the main verb and it includes Conjunction Reduction and – although it may be surprising that this is an advantage of the analysis – also cases which superficially look like instances of coordination within a dependent of the main verb, as in (25).

4 Interactions and extensions

4.1 Agreement

The analysis proposed above naturally accounts for the fact that the elided verb must be understood as having the same semantic features – such as tense – as the overt verb, but it also seems to require that it have the same agreement features, i.e., it seems to exclude examples such as (9), repeated below, where the overt verb occurs in the singular and the elided verb, if present, would occur in the plural:

(9) Maggie only drinks milk, while Carl and Lennie (only drink) beer.

In languages in which subject–verb agreement involves grammatical gender, this morphosyntactic feature also does not have to be shared in Gapping, as illustrated by the Polish example (28) and its intended f-structure (29):

(28) Lisa lubiła Nelsona, a Nelson (lubił) Lisę.
Lisa.NOM.F liked.F Nelson.ACC.M and Nelson.NOM.M liked.M Lisa.ACC.F
'Lisa liked Nelson and Nelson (liked) Lisa.' (Polish)

(29)
$$\left[\begin{array}{l}\text{LOCAL } \boxed{0}\\ \left\{\left[\begin{array}{ll}\text{PRED} & \text{'LIKE<}\boxed{1}\text{,}\boxed{2}\text{>'}\\ \text{SUBJ} & \boxed{1}\left[\begin{array}{ll}\text{PRED} & \text{'LISA'}\\ \text{CASE} & \text{NOM}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{F}\end{array}\right]\\ \text{OBJ} & \boxed{2}\left[\begin{array}{ll}\text{PRED} & \text{'NELSON'}\\ \text{CASE} & \text{ACC}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{M}\end{array}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\boxed{0}, \left[\begin{array}{ll}\text{PRED} & \text{'LIKE<}\boxed{3}\text{,}\boxed{4}\text{>'}\\ \text{SUBJ} & \boxed{3}\left[\begin{array}{ll}\text{PRED} & \text{'NELSON'}\\ \text{CASE} & \text{NOM}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{M}\end{array}\right]\\ \text{OBJ} & \boxed{4}\left[\begin{array}{ll}\text{PRED} & \text{'LISA'}\\ \text{CASE} & \text{ACC}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{F}\end{array}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\right\}\end{array}\right]$$

In the usual LFG analysis of subject–verb agreement, only the subject bears explicit morphosyntactic features such as NUMB(er) or GEND(er), while the verb is, say,

singular and masculine only in the sense that it requires its subject to have appropriate values of NUMB (say, SG) and GEND (say, F(eminine)). So, in this approach, lexical entries of nouns will include defining equations such as (30), while lexical entries of verbs will include constraining equations such as (31).

(30) a. $(\uparrow \text{NUMB}) = \text{SG}$
b. $(\uparrow \text{GEND}) = \text{F}$

(31) a. $(\uparrow \text{SUBJ NUMB}) =_c \text{SG}$
b. $(\uparrow \text{SUBJ GEND}) =_c \text{F}$

In the case of many languages, the above constraining equations are not sufficient, as the verb does not necessarily agree with the subject as a whole, but may agree with just one – usually the closest – conjunct within the subject, when the subject is a coordinate structure (cf., e.g., Sadler 1999, Kuhn and Sadler 2007 and Dalrymple and Hristov 2010). This means that, at least in such languages, the constraining equations in (31) could be replaced with more general equations in (32), where the local name %s, representing the agreement target, is resolved either to the value of SUBJ, in effect giving the equations in (31), or to the value of an appropriate conjunct within SUBJ (cf., e.g., Falk 2006: 198–199 and Patejuk 2015: §3.1.3).

(32) a. $(\uparrow \%\text{S NUMB}) =_c \text{SG}$
b. $(\uparrow \%\text{S GEND}) =_c \text{F}$

An extension of such an analysis of subject–verb agreement to Gapping is immediate: a third – or second, in the case of languages which do not display single conjunct agreement – possible agreement target needs to be added, namely, the subject within LOCAL.[3] In the case of English, where single conjunct agreement is not observed, subject–verb agreement would then boil down to the following statements in the lexical entries of finite verbs:[4]

(33) a. $\%\text{S} = (\uparrow \text{SUBJ}) \mid \%\text{S} = (\uparrow \text{LOCAL SUBJ})$
b. $(\uparrow \%\text{S NUMB}) =_c \text{SG}$
c. $(\uparrow \%\text{S GEND}) =_c \text{F}$

Since such statements are effectively parts of lexical entries of overt finite verbs, the $\uparrow$ metavariable points to the f-structure corresponding to the verb. According to the current analysis of Gapping, the f-structure of the overt verb is at the same time the f-structure of IP1 and, hence, the f-structure of the whole coordinate structure. As the whole structure contains the attribute LOCAL, whose value is the first conjunct, $(\uparrow \text{LOCAL SUBJ})$ exists and the second disjunct in (33) may be selected as defining the target of agreement.[5] This means that there is no sense in which the elided verb

[3] In the case of languages allowing for single conjunct agreement, one more target must be specified, namely, a distinguished conjunct of the LOCAL subject.

[4] Appropriate sets of such statements can be defined as templates, called from lexical entries (Dalrymple et al. 2004).

[5] Given that SUBJ is a distributive feature, this simplified analysis produces spurious ambiguities when two subjects in a Gapping construction bear the same morphosyntactic features. The XLE implementation of this analysis avoids such ambiguities by making sure that the first option, $\%\text{S} = (\uparrow \text{SUBJ})$, is realised only in the absence of Gapping (i.e., in the absence of the LOCAL feature). This

in the second conjunct agrees with the subject within this second conjunct. The only subject–verb agreement in Gapping is that between the overt verb and its (i.e., LOCAL) subject.

4.2 Case assignment

Consider the following slightly more complex Polish sentence and its intended f-structure:

(34) Lisa lubiła Nelsona, a Nelson Lisy nie.
Lisa.NOM.F liked.F Nelson.ACC.M but Nelson.NOM.M Lisa.GEN.F NEG
'Lisa liked Nelson, but Nelson didn't like Lisa.' (Polish)

(35)
$$\left[\begin{array}{l}\left\{\left(\begin{array}{l}\boxed{0}\left[\begin{array}{ll}\text{PRED} & \text{'LIKE<}\boxed{1}\text{,}\boxed{2}\text{>'}\\ \text{SUBJ} & \boxed{1}\left[\begin{array}{ll}\text{PRED} & \text{'LISA'}\\ \text{CASE} & \text{NOM}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{F}\end{array}\right]\\ \text{OBJ} & \boxed{2}\left[\begin{array}{ll}\text{PRED} & \text{'NELSON'}\\ \text{CASE} & \text{ACC}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{M}\end{array}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\end{array}\right), \left[\begin{array}{ll}\text{PRED} & \text{'LIKE<}\boxed{3}\text{,}\boxed{4}\text{>'}\\ \text{SUBJ} & \boxed{3}\left[\begin{array}{ll}\text{PRED} & \text{'NELSON'}\\ \text{CASE} & \text{NOM}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{M}\end{array}\right]\\ \text{OBJ} & \boxed{4}\left[\begin{array}{ll}\text{PRED} & \text{'LISA'}\\ \text{CASE} & \text{GEN}\\ \text{NUMB} & \text{SG}\\ \text{GEND} & \text{F}\end{array}\right]\\ \text{NEG} & +\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\right\}\\ \text{LOCAL}\ \boxed{0}\end{array}\right]$$

Apart from the issue of different agreement features, dealt with in the previous subsection, this example also illustrates a potential problem of different case assignment properties of the overt and the elided verb: while the direct object, *Nelsona*, of the affirmative overt verb occurs in the accusative case, the direct object in the second conjunct, *Lisy*, must occur in the genitive case, due to the presence of the verbal negation marker, *nie*.

Interestingly, to account for different case assignment to the objects in the two conjuncts of (34), no modification is needed to the account of structural case assignment to objects presented in Patejuk and Przepiórkowski 2014b. Simplifying, whether the structurally-cased object receives the accusative or the genitive depends on the presence of a +-valued NEG feature, supplied by the negative marker *nie*, which must occur in the final position when stranded from the verb; cf. the extended rule for IP in (36), which should replace the previous version in (12).

(36) IP → [DEP*, (I)] (NEG)

analysis of agreement in Gapping bears a striking resemblance to Ash Asudeh's unpublished analysis of closest conjunct agreement, as described in Kuhn and Sadler 2007: 314–315; this fact has come to our attention only after developing our analysis in the first version of this paper.

Such a [NEG +] feature is absent on the first conjunct but it is contributed by the final *nie* in the second conjunct, so only the second conjunct is understood as negated.

Crucially, case assignment possibilities are formulated in Patejuk and Przepiórkowski 2014b using off-path constraints, in such a way that they are evaluated separately for each conjunct. More specifically, lexical entries of forms of verbs such as LUBIĆ 'like' call a template whose effect is the same as the following existential constraint:

(37) $(\uparrow \textsc{obj} \quad \textsc{pred} \quad)$
$[\neg((\textsc{obj} \leftarrow) \textsc{neg}) \wedge (\leftarrow \textsc{case}) =_c \textsc{acc}] \quad \vee$
$[((\textsc{xcomp}^* \textsc{obj} \leftarrow) \textsc{neg}) =_c + \wedge (\leftarrow \textsc{case}) =_c \textsc{gen}]$

The basic constraint, stating that there is an object which has a PRED value, is trivially true, but what is less trivial are the off-path specifications attached to this PRED. What they are saying is that either there is no NEG feature on the verb and the object bears the accusative case, or there is a +-valued NEG feature on this or a higher (but only across XCOMP boundaries) verb and the object bears the genitive case. Since these off-path constraints are attached to PRED, they are distributed with it in coordinate structures and they are evaluated independently in each conjunct. So, in the case there is no NEG in one conjunct, the object must bear the accusative case (cf. the first conjunct in (35)), and in the case there is NEG and its value is +, the object must occur in the genitive (cf. the second conjunct in (35)).

This rather complex analysis of structural case assignment is needed independently of Gapping cases such as (34), namely, to account for unlike category coordination, where perhaps only some of the conjuncts receive case at all (Przepiórkowski and Patejuk 2012, Patejuk 2015), as in:

(38) Lisa chciała książkę i żeby ktoś ją przytulił.
Lisa.NOM wanted book.ACC and that somebody.NOM she.ACC hug
'Lisa wanted a book and that somebody hug her.' (Polish)

Hence, the only extension of the basic analysis that is required to handle case assignment facts such as those illustrated by (34) consists in adding the possibility of the final stranded negation to the main IP rule, as done in (36).

4.3 Unlike categories

Similarly, an account of subcategorisation and coordination in Patejuk and Przepiórkowski 2012a and Patejuk 2015, in which alternative ways of categorial realisation of a given grammatical function are evaluated separately for each conjunct (thus making unlike category coordination possible), allows for Gapping examples such as (39), where the object of the gapped verb is realised as different categories in different conjuncts, by analogy to the unlike coordination example in (38):

(39) Lisa chciała książkę, a Maggie żeby ktoś ją
Lisa.NOM wanted book.ACC and Maggie.NOM that somebody.NOM she.ACC
przytulił.
hug
'Lisa wanted a book and Maggie wanted someone to hug her.' (Polish)

Verbs which allow different categories (and hence also their coordination) in a single position specify the range of possibilities in a way analogous to the specification of case possibilities in (37) above. In the specific case of CHCIEĆ 'want' the object is specified as either case-bearing (with case resolved to accusative or genitive as above) or sentential, with the complementiser ŻEBY 'that' (among other possibilities):

(40) $(\uparrow \text{OBJ} \quad \text{PRED} \quad)$
$[\neg((\text{OBJ} \leftarrow) \text{ NEG}) \wedge (\leftarrow \text{CASE}) =_c \text{ ACC}] \quad \vee$
$[((\text{XCOMP}^* \text{ OBJ} \leftarrow) \text{ NEG}) =_c + \wedge (\leftarrow \text{CASE}) =_c \text{ GEN}] \quad \vee$
$(\leftarrow \text{COMP-FORM}) =_c \text{ŻEBY}$

Again, this kind of specification is needed – and was proposed – independently of Gapping examples such as (39), namely, in order to account for unlike category coordination cases such as (38) above. In the case of the Gapping example (39), the effect of the interaction of the above specification with the analysis of Gapping proposed in the previous section is the following f-structure:

(41)
$$\left[\begin{array}{l} \left\{\left(\begin{array}{l} \boxed{0}\left[\begin{array}{ll} \text{PRED} & \text{'WANT<}\boxed{1}\text{,}\boxed{2}\text{>'} \\ \text{SUBJ} & \boxed{1}\left[\begin{array}{ll} \text{PRED} & \text{'LISA'} \\ \text{CASE} & \text{NOM} \end{array}\right] \\ \text{OBJ} & \boxed{2}\left[\begin{array}{ll} \text{PRED} & \text{'BOOK'} \\ \text{CASE} & \text{ACC} \end{array}\right] \\ \text{TENSE} & \text{PAST} \\ \text{VOICE} & \text{ACTIVE} \end{array}\right], \\ \left[\begin{array}{ll} \text{PRED} & \text{'WANT<}\boxed{3}\text{,}\boxed{4}\text{>'} \\ \text{SUBJ} & \boxed{3}\left[\begin{array}{ll} \text{PRED} & \text{'MAGGIE'} \\ \text{CASE} & \text{NOM} \end{array}\right] \\ \text{OBJ} & \boxed{4}\left[\begin{array}{ll} \text{PRED} & \text{'HUG<}\boxed{5}\text{,}\boxed{6}\text{>'} \\ \text{SUBJ} & \boxed{5}\left[\begin{array}{ll} \text{PRED} & \text{'SB'} \\ \text{CASE} & \text{NOM} \end{array}\right] \\ \text{OBJ} & \boxed{6}\left[\begin{array}{ll} \text{PRED} & \text{'SHE'} \\ \text{CASE} & \text{ACC} \end{array}\right] \\ \text{COMP-FORM} & \text{THAT} \end{array}\right] \\ \text{TENSE} & \text{PAST} \\ \text{VOICE} & \text{ACTIVE} \end{array}\right] \end{array}\right)\right\} \\ \text{LOCAL} \quad \boxed{0} \end{array}\right]$$

5 Limitations

5.1 Verb clusters

Let us examine the following sentences, similar to those given in §3:

(42) Lisa has read a book and Marge a magazine.

(43) Lisa wanted to read a book and Marge a magazine.

In both cases, there is more than just a single verb form which is elided in the gapped sentence: in the case of (42) it is *has* and *read*, and in the case of (43) it is *wanted*, *to* and *read*. On standard LFG assumptions, the first sentence is not problematic for our analysis, as *has* and *read* would be treated as co-heads within the I constituent and would share a single f-structure. However, the second sentence involves two contentful verbs which are elided, each projecting a separate f-structure: *wanted*

and *(to) read*.

Such examples belong to the core cases of Gapping and were systematically discussed in Ross 1970. On the basis of examples adduced there, Johnson (2014: 2) provides the following examples illustrating the possibility of eliding a varying amount of verbal material:

(44) I want to try to begin to write a novel, and Mary (wants) to get ready to set out to review a play.

(45) I want to try to begin to write a novel, and Mary (wants to try) to set out to review a play.

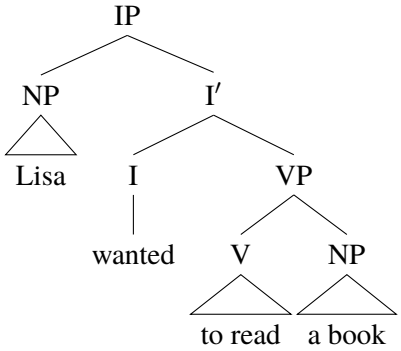
(46) I want to try to begin to write a novel, and Mary (wants to try to begin) to review a play.

(47) I want to try to begin to write a novel, and Mary (wants to try to begin to review) a play.

He notes that the more material is elided, the easier it is to process such sentences, probably due to the fact that fewer constituents are contrasted then.

Given the standard assumption – supported by constituency tests – that infinitival phrases form constituents, i.e., that a sentence like the first conjunct in (43) should have a c-structure like that given schematically in (48), it is not clear to us how to extend the account proposed here to deal with such cases.

(48)



On the other hand, the current account may be straightforwardly extended to deal with similar facts in Polish. The crucial property of Polish which differentiates it from English is its free word order, also in such 'verb cluster' environments. For example, while the most obvious translation of the first conjunct of (43) is given in (49), other word orders are also possible, given the right information structure requirements, including (50)–(53):

(49) Lisa chciała przeczytać książkę.
Lisa.NOM wanted read.INF book.ACC
'Lisa wanted to read a book.' (Polish)

(50) Przeczytać książkę chciała Lisa.
read.INF book.ACC wanted Lisa.NOM

(51) Lisa książkę chciała przeczytać.
Lisa.NOM book.ACC wanted read.INF

(52) Książkę chciała przeczytać Lisa.
book.ACC wanted read.INF Lisa.NOM

(53) Przeczytać chciała Lisa książkę.
read.INF wanted Lisa.NOM book.ACC

This word order freedom suggests a sentential rule which does not treat infinitival dependents as single constituents but rather separates infinitival verbs from their dependents. That means that, in the case of Polish, rule (13), repeated below, which defines possible dependents of verbs, should be replaced with a rule like (54), where single infinitival verbs – rather than whole infinitival phrases – may be (possibly indirect) dependents, and where other constituents – apart from subjects – are interpreted as dependents of either the main verb or one of the infinitival verbs:[6]

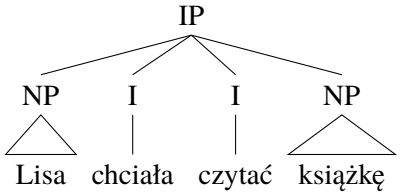
(13) DEP ≡ NP | InfP | PP | ...
$(\uparrow \text{SUBJ}) = \downarrow$ $(\uparrow \text{XCOMP}) = \downarrow$ $\downarrow \in (\uparrow \text{ADJ})$
$(\downarrow \text{CASE}) =_c \text{NOM}$ ... ...

(54) DEP ≡ NP | I | PP |
$(\uparrow \text{SUBJ}) = \downarrow$ $(\uparrow \text{XCOMP}^+) = \downarrow$ $\downarrow \in (\uparrow \text{XCOMP}^* \text{ ADJ})$
$(\downarrow \text{CASE}) =_c \text{NOM}$ $(\downarrow \text{VFORM}) =_c \text{INF}$ ...

NP | ...
$(\uparrow \text{XCOMP}^* \text{ OBJ}) = \downarrow$
$(\downarrow \text{CASE}) =_c \text{ACC}$

According to this rule, (49) receives the c-structure in (55).

(55)



Now, given the above c-structure rule independently motivated by word order facts, it is easy to analyse sentences such as (56) (a Polish translation of (43)) in a way that assigns them an f-structure such as (57).[7]

(56) Lisa chciała przeczytać książkę, a Marge czasopismo.
Lisa.NOM wanted read.INF book.ACC and Marge.NOM magazine.ACC
'Lisa wanted to read a book and Marge a magazine.' (Polish)

For such an analysis to work, out of the three dependents in the first conjunct, namely, the subject of the main verb *Lisa*, the lower verb *czytać* 'read' and the object of the lower verb *książkę* 'book', the subject and the object are put into LOCAL and, hence, will not distribute to other conjuncts, while the lower verb, providing

[6] See Zaenen and Kaplan 1995 for a similar use of functional uncertainty in an analysis of West Germanic verb clusters.

[7] The actual account is a little more complicated as the infinitival argument of CHCIEĆ 'want' is, arguably, its OBJECT rather than XCOMP (Patejuk and Przepiórkowski 2014a).

the value of XCOMP, will contribute to the top level of IP1 and, hence, will distribute to other conjuncts, cf. (58).

(57)
$$\left[\begin{array}{l}\left\{\left(\boxed{0}\left[\begin{array}{ll}\text{PRED} & \text{`WANT<}\boxed{1}\text{,}\boxed{2}\text{>'}\\ \text{SUBJ} & \boxed{1}\left[\begin{array}{ll}\text{PRED} & \text{`LISA'}\\ \text{CASE} & \text{NOM}\end{array}\right]\\ \text{XCOMP} & \boxed{2}\left[\begin{array}{ll}\text{PRED} & \text{`READ<}\boxed{1}\text{,}\boxed{3}\text{>'}\\ \text{SUBJ} & \boxed{1}\\ \text{OBJ} & \boxed{3}\left[\begin{array}{ll}\text{PRED} & \text{`BOOK'}\\ \text{CASE} & \text{ACC}\end{array}\right]\end{array}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\right), \left[\begin{array}{ll}\text{PRED} & \text{`WANT<}\boxed{4}\text{,}\boxed{5}\text{>'}\\ \text{SUBJ} & \boxed{4}\left[\begin{array}{ll}\text{PRED} & \text{`MARGE'}\\ \text{CASE} & \text{NOM}\end{array}\right]\\ \text{XCOMP} & \boxed{5}\left[\begin{array}{ll}\text{PRED} & \text{`READ<}\boxed{4}\text{,}\boxed{6}\text{>'}\\ \text{SUBJ} & \boxed{4}\\ \text{OBJ} & \boxed{6}\left[\begin{array}{ll}\text{PRED} & \text{`MAGAZINE'}\\ \text{CASE} & \text{ACC}\end{array}\right]\end{array}\right]\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\end{array}\right]\right\}\\ \text{LOCAL}\quad \boxed{0}\end{array}\right]$$

(58)
$$\left[\begin{array}{ll}\text{PRED} & \text{`WANT<SUBJ,XCOMP>'}\\ \text{TENSE} & \text{PAST}\\ \text{VOICE} & \text{ACTIVE}\\ \text{XCOMP} & \left[\text{PRED}\quad \text{`READ<SUBJ,OBJ>'}\right]\\ \text{LOCAL} & \left[\begin{array}{ll}\text{SUBJ} & \left[\text{PRED}\quad \text{`LISA'}\right]\\ \text{XCOMP} & \left[\text{OBJ}\quad \left[\text{PRED}\quad \text{`BOOK'}\right]\right]\end{array}\right]\end{array}\right]$$

While this analysis also successfully handles the more complex Polish facts analogous to those in (44)–(47), it crucially relies on the fact that Polish is a free word language and that infinitival environments display 'clause union' properties in this language, so it is unfortunately not directly applicable to languages such as English, where similar environments have a more hierarchical c-structure.

5.2 Other loose ends

Apart from the possibility of eliding more verbal material, which is problematic for the current account (and many other accounts) in the case of languages such as English, there are other issues left for future research, which may or may not turn out to be problematic.

First of all, the analysis is currently limited to coordinate constructions, with the first conjunct providing interpretation for the constituents elided in subsequent conjuncts. But, as is well known, at least two other constructions are similar to Gapping, namely, list-like answers to questions and comparative constructions (Johnson 2014: 1):

(59) Q. Who met who?
A. Jerry, Sarah; Sally, Mark; Trish, Betsy

(60) Sally met more parents than Tom, kids.

We believe that the analysis proposed here may be extended to such cases by defin-

ing other sources of material to be distributed than just IP1 in coordinate structures, but we have not yet attempted to do so.

What is potentially more problematic for the current analysis are examples like the following:

(61) Homer and Bart like donuts, Maggie likes yogurt and Lisa apples.

(62) Either Homer likes donuts and Bart bananas, or Bart donuts and Homer bananas.

The analysis fails in the case of (61) because the verbal information from the first (or the second) conjunct that must be distributed to the last conjunct is distributed to all conjuncts, including to the second (or the first) conjunct, which already contains such information. This problem could perhaps be avoided on the assumption that (61) involves two binary coordinate structures: asyndetic, with *Homer and Bart like donuts* as the first conjunct and the rest of the sentence as the second, and syndetic, with *Maggie likes yogurt* as the first conjunct and *Lisa apples* as the second. But in the case of (62) it is exactly the recursive structure of coordination that is problematic: the material contributed by *likes* in the first conjunct only spreads to other elements of the same inner set, i.e., to the second conjunct, but not to the third or fourth conjuncts, which form their own set. To the best of our knowledge, such examples, to the extent that they are grammatical, represent a challenge to all theories of Gapping. In any case, the issues discussed in this section, as well as interactions of the proposed analysis with extraction, scope of negation, etc., have to be postponed to future research.

6 Previous LFG work

As mentioned in the introduction, there is no previous comprehensive LFG analysis of Gapping, but there are at least two suggestions. One is made in passing in Kaplan 1987, 1995: 365–366, where a new operator is proposed, called 'priority union' (see also Dalrymple 2001: 175–177). Given two f-structures, f and g, their priority union f/g is that feature structure which contains all attribute–value pairs which constitute f (i.e., all information present in f), as well as those attribute–value pairs $\langle a, v\rangle$ in g for which a does not have a value in f. The intuition is that, if f is the f-structure corresponding directly to a gapped conjunct and g is the first conjunct, then f/g is the full f-structure of the gapped conjunct, with information missing in f filled by g.

To the best of our knowledge, this intuitive idea has never been turned into a worked-out analysis of Gapping. While it is possible that a successful account of Gapping may be based on this priority union operator, the current paper proposes an analysis which does not assume any non-standard machinery, one that can be directly implemented in XLE.[8]

[8] XLE (Crouch et al. 2011) is a comprehensive platform for turning LFG grammars into parsers, which implements various mechanisms proposed throughout the history of LFG (including off-path constraints, hybrid feature structures and the restriction operator), but which lacks an implementation of priority union.

Another idea is presented in Alzaidi 2010 and is based on an earlier analysis of Frank 2002, which also inspired the current account. The target of Frank's analysis is the so-called SGF coordination in German, illustrated by (63) (Höhle 1983), where the subject within the first conjunct is shared with the second conjunct; the simplified intended f-structure for this example is given in (64).

(63) In den Wald ging der Jäger und fing einen Hasen.
into the forest went the.NOM hunter.NOM and caught a.ACC rabbit.ACC
'The hunter went into the forest and caught a rabbit.' (German)

(64) $\left\{ \begin{bmatrix} \textsc{pred} & \text{'GO<}\boxed{1}\text{,}\boxed{2}\text{>'} \\ \textsc{subj} & \boxed{1}\begin{bmatrix} \textsc{pred} & \text{'HUNTER'} \end{bmatrix} \\ \textsc{obl} & \boxed{2}\text{"INTO THE FOREST"} \end{bmatrix}, \begin{bmatrix} \textsc{pred} & \text{'CATCH<}\boxed{1}\text{,}\boxed{3}\text{>'} \\ \textsc{subj} & \boxed{1} \\ \textsc{obj} & \boxed{3}\begin{bmatrix} \textsc{pred} & \text{'RABBIT'} \end{bmatrix} \end{bmatrix} \right\}$

Since the subject does not occur in the sentence-initial position in the first conjunct, the sharing of the subject cannot be interpreted as resulting from VP-coordination (as it could in the case of the English translation).

In order to account for SGF coordination, Frank 2002: 188 breaks the symmetry of the basic coordination rule (10), repeated below, and replaces it with (65).

(10) $\text{IP} \rightarrow \underset{\downarrow\in\uparrow}{\text{IP}}\ \text{Conj}\ \underset{\downarrow\in\uparrow}{\text{IP}}$

(65) $\text{IP} \rightarrow \underset{\substack{\downarrow\in\uparrow \\ ((\uparrow \textsc{gdf}) = (\downarrow \textsc{gdf}))}}{\text{IP}}\ \text{Conj}\ \underset{\downarrow\in\uparrow}{\text{IP}}$

GDF stands here for a grammatical discourse function, i.e., TOPIC, FOCUS or – crucially for the example (63) – SUBJ. The effect of this additional annotation is that the subject of the first conjunct (i.e., *der Jäger* 'the hunter') optionally becomes the subject of the whole set and, hence, if it does, it distributes to both conjuncts: spuriously to the first conjunct and crucially to the second.

Alzaidi 2010: 81 proposes applying the same idea to Gapping (in Hejazi Arabic), where what is distributed is not a grammatical function, but PRED and TENSE:

(66) $\text{IP} \rightarrow \underset{\substack{\downarrow\in\uparrow \\ (\uparrow \textsc{pred}) = (\downarrow \textsc{pred}) \\ (\uparrow \textsc{tense}) = (\downarrow \textsc{tense})}}{\text{IP}}\ \text{Conj}\ \underset{\downarrow\in\uparrow}{\text{IP}}$

There are various problems with this idea. The main formal problem is that the explicit equation $(\uparrow \textsc{pred}) = (\downarrow \textsc{pred})$ in (66) works differently than the distribution of PRED proposed above. In (66), the value of $(\downarrow \textsc{pred})$ is the specific value of this feature as instantiated in the first conjunct, i.e., together with specific values of any arguments mentioned in the semantic form. This means that the same first-conjunct arguments are present in the PRED value of the second conjunct, i.e., that these arguments cannot be instantiated to whatever values of SUBJ, OBJ, etc., are introduced in the second conjunct. For this reason the analysis crashes (as verified in an attempt at implementing it in XLE). This differs from the analysis proposed

here, where the information introduced by the verb in the first conjunct is distributed to all conjuncts simultaneously.

There is also a conceptual problem with the analysis proposed in Alzaidi 2010, namely, that the identity of semantic features must be stipulated by explicit statements such as $(\uparrow \textsc{tense}) = (\downarrow \textsc{tense})$. So it is conceivable that there are languages where only some such identities hold, but not all – say only voice must be equal, but not tense. As we are not aware of languages in which Gapping would not involve the same conditions on the identity of all semantic features, an analysis like the one proposed in the current paper – one where all features defined in verbal lexical entries go together – should be preferred. Note also that – since not only the verb but also various dependents may be elided in Gapping – the first conjunct in (66) would also have to be annotated with optional equations such as $((\uparrow \textsc{gf}) = (\downarrow \textsc{gf}))$, for any non-adjunct grammatical function, and an equation such as $((\uparrow \textsc{adj} \in) = (\downarrow \textsc{adj} \in))*$ (and similarly for $\textsc{xadj}$), to account for the possibility of eliding any number (hence the Kleene star) of adjuncts introduced in the first conjunct.

Finally, it is not clear how the proposal of Alzaidi 2010 could be extended to deal with different case values or different syntactic categories of the same grammatical functions (see §§4.2–4.3). We conclude, then, that the current proposal is the first working analysis of a reasonable subset of Gapping phenomena in LFG.

7 Summary

Several types of non-constituent coordinations, including Conjunction Reduction and Right Node Raising, received an interesting uniform LFG(ish) account in Maxwell and Manning 1996. Unfortunately, Gapping falls outside of the scope of that analysis. In general, Kaplan 1987, 1995 and Alzaidi 2010 notwithstanding, no working analysis of gapping has been proposed in LFG. We hope to have just bridged this gap. The account proposed here does not use any non-standard mechanism and it correctly interacts with agreement, case assignment and diverse morpho-syntactic realisations of arguments; all these interactions were verified in the implementation of the current analysis within the large-scale XLE grammar of Polish mentioned above. The analysis naturally captures the observation that the elided verb must be understood as bearing the same semantic and information-structural features as the overt verb, and – given the independently motivated flat structure approach to 'verb clusters' in Polish – accounts for the possibility of eliding larger chunks of verbal material in this language.

One of the already known limitations of this analysis, however, is that it is not clear how to extend the analysis of this last aspect of Gapping – the possibility to elide more verbal material – to languages such as English, with more hierarchical c-structure. It is also not immediately clear how to extend the analysis to coordinate structures in which more than one conjunct contains the overt main verb or to cases of recursively embedded coordination. Such cases, and various interactions of the proposed analysis with other phenomena, should be a matter of future research.

References

Alzaidi, Muhammad Swaileh. 2010. *Gapping and Right Node Raising: An LFG Approach*. Masters Thesis, University of Essex.

Butt, Miriam and King, Tracy Holloway (eds.). 2014. *The Proceedings of the LFG'14 Conference*, Stanford, CA, CSLI Publications.

Crouch, Dick, Dalrymple, Mary, Kaplan, Ron, King, Tracy, Maxwell, John and Newman, Paula. 2011. XLE Documentation. http://www2.parc.com/isl/groups/nltt/xle/doc/xle_toc.html.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Dalrymple, Mary and Hristov, Bozhil. 2010. Agreement Patterns and Coordination in Lexical Functional Grammar. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'10 Conference*, pages 186–206, Ottawa, Canada: CSLI Publications.

Dalrymple, Mary, Kaplan, Ronald M. and King, Tracy Holloway. 2004. Linguistic Generalizations over Descriptions. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'04 Conference*, pages 199–208, Stanford, CA: CSLI Publications.

Dalrymple, Mary, Kaplan, Ronald M., Maxwell, III, John T. and Zaenen, Annie (eds.). 1995. *Formal Issues in Lexical-Functional Grammar*. CSLI Lecture Notes, No. 47, Stanford, CA: CSLI Publications.

Dalrymple, Mary, Shieber, Stuart M. and Pereira, Fernando C. N. 1991. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy* 14, 399–452.

Falk, Yehuda N. 2006. On the Representation of Case and Agreement. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'06 Conference*, pages 184–201, Universität Konstanz, Germany: CSLI Publications.

Frank, Annette. 2002. A (Discourse) Functional Analysis of Assymmetric Coordination. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'02 Conference*, pages 174–196, National Technical University of Athens: CSLI Publications.

Höhle, Tilman N. 1983. Subjektlücken in Koordinationen, unpublished manuscript, Köln.

Johnson, Kyle. 2014. Gapping, http://people.umass.edu/kbj/homepage/Content/gapping.pdf.

Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. In P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson and P. Bennett (eds.), *Linguistic Theory and Computer Applications*, pages 149–188, London: Academic Press.

Kaplan, Ronald M. 1995. Three Seductions of Computational Psycholinguistics. In Dalrymple et al. (1995), pages 339–367.

Kaplan, Ronald M. and Maxwell, III, John T. 1988. Constituent Coordination in Lexical-Functional Grammar. In *Proceedings of the 12th International Conference on Computational Linguistics (COLING 1988)*, pages 303–305, Budapest.

Kuhn, Jonas and Sadler, Louisa. 2007. Single Conjunct Agreement and the Formal Treatment of Coordination in LFG. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'07 Conference*, pages 302–322, University of Stanford, Cali-

fornia, USA: CSLI Publications.
Larson, Richard. 1985. On the Syntax of Disjunction Scope. *Natural Language and Linguistic Theory* 3, 217–264.
Maxwell, III, John T. and Manning, Christopher D. 1996. A Theory of Non-constituent Coordination based on Finite-State Rules. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'96 Conference*, Grenoble, France: CSLI Publications.
Neijt, Anneke. 1979. *Gapping: A Contribution to Sentence Grammar*. Dordrecht: Foris Publications.
Patejuk, Agnieszka. 2015. *Unlike coordination in Polish: an LFG account*. Ph.D. Thesis, Instytut Języka Polskiego PAN, Cracow.
Patejuk, Agnieszka and Przepiórkowski, Adam. 2012a. A Comprehensive Analysis of Constituent Coordination for Grammar Engineering. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012)*, pages 2191–2207, Mumbai, India.
Patejuk, Agnieszka and Przepiórkowski, Adam. 2012b. Towards an LFG Parser for Polish: An Exercise in Parasitic Grammar Development. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012*, pages 3849–3852, ELRA, Istanbul, Turkey.
Patejuk, Agnieszka and Przepiórkowski, Adam. 2014a. Control into Selected Conjuncts. In Butt and King (2014), pages 448–460.
Patejuk, Agnieszka and Przepiórkowski, Adam. 2014b. Structural Case Assignment to Objects in Polish. In Butt and King (2014), pages 429–447.
Patejuk, Agnieszka and Przepiórkowski, Adam. 2014c. Synergistic Development of Grammatical Resources: A Valence Dictionary, an LFG Grammar, and an LFG Structure Bank for Polish. In V. Henrich, E. Hinrichs, D. de Kok, P. Osenova and A. Przepiórkowski (eds.), *Proceedings of the Thirteenth International Workshop on Treebanks and Linguistic Theories (TLT 13)*, pages 113–126, Tübingen: Department of Linguistics (SfS), University of Tübingen.
Patejuk, Agnieszka and Przepiórkowski, Adam. 2017. POLFIE: współczesna gramatyka formalna języka polskiego. *Język Polski* XCVII(1), 48–64.
Peterson, Peter G. 2004. Coordination: Consequences of a Lexical-Functional Account. *Natural Language and Linguistic Theory* 22, 643–679.
Przepiórkowski, Adam and Patejuk, Agnieszka. 2012. On case assignment and the coordination of unlikes: The limits of distributive features. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'12 Conference*, pages 479–489, Stanford, CA: CSLI Publications.
Ross, John. 1970. Gapping and the Order of Constituents. In M. Bierwisch and K. E. Heidolph (eds.), *Progress in Linguistics*, The Hague: Mouton.
Sadler, Louisa. 1999. Non-Distributive Features in Welsh Coordination. In M. Butt and T. H. King (eds.), *The Proceedings of the LFG'99 Conference*, University of Queensland, Brisbane: CSLI Publications.
Schwarz, Bernhard. 1999. On the Syntax of *either* ... *or*. *Natural Language and Linguistic Theory* 17, 339–370.
Zaenen, Annie and Kaplan, Ronald M. 1995. Formal Devices for Linguistic Generalizations: West Germanic Word Order in LFG. In Dalrymple et al. (1995), pages 215–239.