

COPULAR COMPLEMENTS: CLOSED OR OPEN?

Mary Dalrymple, Helge Dyvik, and Tracy Holloway King
Oxford University, University of Bergen, Palo Alto Research Center

Proceedings of the LFG04 Conference
University of Canterbury
Miriam Butt and Tracy Holloway King (Editors)

2004

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

Abstract

In this paper, we present some preliminary observations on the syntactic analysis of copular constructions in LFG. We suggest several conclusions and directions for future research. First, the open analysis is appropriate for some copular constructions but not for others, even within the same language. Second, the same construction can be open in some languages and closed in others. Third, we can determine whether a particular copular construction is open or closed by examining phenomena like agreement and the obligatory presence of the copula. Fourth, we do not expect open postcopular complements always to be open in other environments.

1 Introduction

In this paper, we explore some questions in the syntactic analysis of copular constructions in LFG. Examples from English are shown in (1).

- (1) a. They are books.
b. The books are flimsy.
c. The books are on the table.
d. The problem is that they appear.
e. The problem is their appearing.
f. The problem is to leave before 6:00.

We focus on the grammatical function of the post-copular element. In particular, is it an open complement (XCOMP, shown in (2a)) or a closed complement (PREDLINK, as proposed by Butt et al. 1999, shown in (2b))? Or is it open for some categories and closed for others? And how does this vary across languages? Do all languages have XCOMPS with copulas? Or PREDLINKS? Or do some have one and some the other?

(2) a. Open Complement

$$\left[\begin{array}{ll} \text{PRED} & \text{'be<XCOMP>SUBJ'} \\ \text{SUBJ} & [\dots]1 \\ \text{XCOMP} & \left[\begin{array}{ll} \text{PRED} & \text{'... <SUBJ>'} \\ \text{SUBJ} & []1 \end{array} \right] \end{array} \right]$$

b. Closed Complement

$$\left[\begin{array}{ll} \text{PRED} & \text{'be<SUBJ, PREDLINK>'} \\ \text{SUBJ} & [\dots] \\ \text{PREDLINK} & \left[\begin{array}{ll} \text{PRED} & \text{'...'} \end{array} \right] \end{array} \right]$$

Although the difference between open and closed complements has semantic reflexes which have been examined extensively in the semantic literature, we focus here on syntactic issues. Positing a SUBJ in the f-structure should be supported by syntactic and not semantic criteria, since otherwise the proper distinction between grammatical functions and semantic arguments is not maintained.

In examining this issue, we first make our notion of 'syntactic criteria' more precise, before looking at constructions in which the copula may or must be absent. We then turn to constructions in which the post-copular constituent contains its own subject, separate from the subject in the matrix clause. Next, we look at control relations and agreement involving the subject and the post-copular constituent. Finally, we provide some discussion.

1.1 Subject Criteria

If f-structure analyses are to serve as a basis for typological comparison of languages, they must reflect grammatical rather than purely semantic properties of the expressions analyzed. Similarity in argument structure or semantic content shows only that the same content can be expressed in the languages compared, not that the syntactic expression of this content is parallel: however, the latter is necessary to justify a syntactic generalization. Therefore, the observation that a predicate takes a semantic argument is not sufficient to justify the conclusion that the lexeme expressing the predicate subcategorizes for a corresponding syntactic function. Such a conclusion needs further support from grammatical properties.

Here we restrict ourselves to presenting our basic assumptions about the definition of subject, which are the following. Universal characterizations of categories like ‘subject’ refer to prototypes. A language has subjects only if it has prototypical subjects, but then it may also have non-prototypical subjects. We take a prototypical subject to be a grammatically isolable nominal constituent which may carry the roles ‘volitional’ and ‘agentive’ with appropriate predicates, and which expresses wide scope in basic sentences with unmarked intonation.

A central point here is that a prototypical subject has a c-structure realization: prototypical subjects are overtly expressed. If we discard this criterion, i.e. envisage operating with the term ‘subject’ in cases where there is never any c-structure realization of this function, then ‘subject’ ceases to be a grammatical term and becomes a purely semantic term, denoting a subtype of semantic argument. On the other hand, this criterion does not forbid the occurrence of nonprototypical, (locally) non-overt subjects. We can argue for the presence of non-overt subjects in the following way: Predicates which appear with overt subjects must subcategorize for ‘subject’ in order to achieve the appropriate role assignment. This leads to the assumption of non-prototypical functional subjects when these forms occur as predicates in constructions where no (local) constituent bears this function. This motivates, for instance, the positing of controlled subjects in the case of infinitives, (3a), and of null subjects in pro-drop constructions, (3b), since these verbs also occur as the syntactic heads of predicates that combine with prototypical subjects.

- (3) a. They want to leave.
b. Katalavéno elliniká.
 understand-1SG Greek
 ‘I understand Greek.’ (Modern Greek)

In order to investigate which lexical categories take subjects and hence may occur in open complements, we need to consider the motivation for assuming that a given expression E has a subject. We take the primary criterion to be E’s status as syntactic head of a predicate phrase. Another potentially relevant property is agreement, if we (with, e.g., Keenan 1976) include among our prototypical subject criteria that the subject is always among the controllers of agreement in languages with agreement. Agreement would then be a possible basis for saying that the ‘subject-of’ relation obtains between the subject and the expression showing agreement.

1.2 Copular Occurrence Governed by Category

Some languages are unlike English in that the copula is not required, or not permitted, in certain constructions involving nonverbal predication (Rosén 1996). For example, Japanese adjectives do not require the copula, as seen in (4):

- (4) a. hon wa akai
 book red
 ‘The book is red.’ (Japanese)

- b.
$$\left[\begin{array}{l} \text{PRED} \quad \text{'red<SUBJ>'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad \text{'book'} \right] \end{array} \right]$$

In these cases, it is plausible to assume that the adjective provides the main PRED for the clause. This is based on the reasoning in section 1.1: the adjective is the syntactic head of the predicate phrase. If this is not considered a sufficient criterion for assuming that it subcategorizes for the (prototypical) subject of the sentence, then even the assumption that ordinary verbs subcategorize for subjects may be called into question. Thus, in order to satisfy Coherence, the predicate must be open, subcategorizing for a SUBJ (Andrews 1982). In (4b), the adjective *akai* 'red' contributes the main PRED of the f-structure. It takes a SUBJ which licenses the occurrence of *hon* 'book'. Thus, in languages like Japanese, adjectives take overt SUBJs with which they combine directly in the syntax.

We propose that this analysis is also the correct one for cases in which Japanese adjectival predicates occur with an overt copula: a complement should be treated as open, subcategorizing for a SUBJ, if it can ever be used without a copula. Thus, in (5), the adjective is also an open function, and also subcategorizes for a SUBJ:

- (5) sono hon wa akai desu
 this book red is
 'This book is red.' (Japanese)

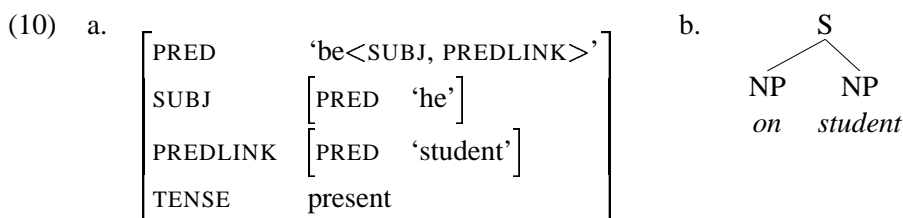
In contrast to the examples above with adjectival complements, in Japanese copular sentences with nominal complements, the copula cannot be omitted, as in (6).

- (6) a. sono hon wa syousetsu desu
 this book novel is
 'This book is a novel.' (Japanese)
 b. *sono hon wa syousetsu

This suggests that Japanese nouns are not open complements and do not subcategorize for a SUBJ. We conclude that the category of the constituent – whether it is an adjective or a noun – can affect whether it can be a predicate on its own, licensing its own subject, or whether it must occur with a copula. Even within the same language, different constituents can behave differently in copular constructions. See Falk (2004) for a similar conclusion for Hebrew. Interestingly, Falk argues that some adjectival “copular” constructions in Hebrew take closed PREDLINK complements, while others are like the Japanese adjectives in (5) in which the adjective provides the main clausal predicate. The fact that different constituents can behave differently in copular constructions means that the full range of copular constructions must be examined within a language in order to analyze it completely. That is, the fact that one type of constituent requires a certain analysis of copular constructions does not guarantee that other, superficially similar constructions will be amenable to the same analysis.

1.3 Copula Occurrence Governed by Tense

In some languages, the copula is required in some situations but forbidden in others. One crosslinguistically common governing factor for this is tense. For example, in Russian (Chvany 1975) and Arabic (Shlonsky 1997), the copula is null in the present tense but overt in the past and future tenses. This is shown for Russian in (7).



Given this rule, the f-structure in (10a) is identical to those associated with overt copular constructions, except that the value of tense is present instead of past or future. However, the predicate and tense information come from the annotated phrase structure rules instead of the overt copular form. This is due to the presence of the ϵ category in rule (9), which amounts to an instruction to introduce the associated equations if no VCop constituent is present. It is important to note that the ϵ in (10) is not an empty category; that is, no ϵ will appear in the c-structure in (10b). The only effect is on the f-structure: the associated equations will be introduced as if they were associated with the other categories in the rule (Kaplan and Maxwell 1996).

Here we do not argue for one or the other of these analyses for Russian because there are a number of other Russian copular constructions whose analysis may bear on the canonical constructions in this section. Note that certain impersonal adverbial predicates use the forms of *byt'* as tense markers with the same distribution as copular *byt'*. Regardless of the chosen analysis for the Russian data, the possibility of finding both analyses crosslinguistically remains. For example, the Arabic data, which was not presented here, would need to be considered in its own light.

1.4 The Copula as Grammatical Prothesis

In English, in contrast to languages like Japanese, an adjective cannot occur on its own as the syntactic head of a predicate; a copula is always required. This provides a functionally-motivated account of the existence of the copula: it is needed because the adjectives themselves are unable to combine directly with overt SUBJs, unlike Japanese adjectives.

Given this, the copula can be seen as giving to the adjective a needed grammatical prothesis: a SUBJ argument to which to link the adjective's semantic role. This analysis entails that the syntactic head of the predicate is the copula, not the adjective; syntactically, the overt subject is SUBJ of the copula. For constructions where the copula is obligatory, then, the conclusion would be that adjectives in English do not take SUBJs, and hence that examples like (1b) take closed PREDLINKs rather than open XCOMPs. Corresponding arguments would lead to the same conclusions for PPs, as in (1c).

- (1) b. The books are flimsy.
 c. The books are on the table.

However, as we will see in section 3, there are control constructions involving English copular verbs which seem to require an open complement analysis. This once again suggests that copular constructions even within a single language may not have a uniform analysis.

2 Subjects in Post-Copular Constructions

The closed complement PREDLINK analysis is mandated when the post-copular element already has a subject, as with post-copular *that*-clauses, certain gerunds, and some modal uses of the copula with null pronominal subjects.³ These are repeated in (11).

³Some modal uses of the copula do not pose a problem in this respect because the subject of the copula is the same as the subject of the post-copular verb. This is similar to the behavior of more prototypical modals such as *would* and *should*. An example is shown

- (11) a. The problem is that they appear.
 b. The problem is their appearing.
 c. The problem is (for them) to leave before 6.

In all of these examples, the post-copular element contains a verb which has a subject distinct from the subject of the copula. In (11a) it is *they*, in (11b) it is the possessive *their*, and in (11c) it is either an arbitrary pronoun or the *them* in *for them*.

The XCOMP analysis of these constructions results in a clash of PRED values. The control equation (\uparrow SUBJ)=(\uparrow XCOMP SUBJ) equates the subject of the copula with the subject of the post-copular constituent. However, since these are distinct, the result is a failure in unification of the PREDs of the XCOMP's subject and an ungrammatical structure, as shown in (12) for (11a).

(12) Open Complement

$$\left[\begin{array}{l} \text{PRED} \quad \text{'be<XCOMP>SUBJ'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad \text{'problem'} \right] \\ \text{XCOMP} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'appear<SUBJ>'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad * \text{'they/problem'} \right] \end{array} \right] \end{array} \right]$$

This problem is avoided if the copula takes a closed complement, as shown in (13).

(13) Closed Complement

$$\left[\begin{array}{l} \text{PRED} \quad \text{'be<PREDLINK>SUBJ'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad \text{'problem'} \right] \\ \text{PREDLINK} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'appear<SUBJ>'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad \text{'they'} \right] \end{array} \right] \end{array} \right]$$

Under the closed PREDLINK analysis, there is no control equation to unify the two subjects and the result is a grammatical analysis. Thus, the PREDLINK analysis is the only one possible for constructions where the copular complement has its own subject, distinct from the matrix subject.

3 Control Relations

In sharp contrast to the examples in section 2 in which the subjects differed and hence required the closed PREDLINK analysis, there are constructions which are better treated under an XCOMP analysis. The XCOMP analysis is excellent for capturing certain control relations, such as those in (14).

- (14) a. It is likely/bound/certain to rain.
 b. They are eager/foolish/loathe to leave.

in (i).

- (i) They are to leave at six.

Whether these modal uses of the copula should be treated similarly to the more canonical uses is not explored here.

In these examples, the subject of the copula is also the subject of the verb embedded in the post-copular constituent.

If these adjectives have subjects, then the chain of control from the matrix subject through the adjective to its verbal complement is standardly described by the control equation (\uparrow SUBJ)=(\uparrow XCOMP SUBJ). The f-structure for (14a) is shown in (15).

$$(15) \left[\begin{array}{l} \text{PRED} \quad \text{'be<XCOMP>SUBJ'} \\ \text{SUBJ} \quad \left[\begin{array}{l} \text{PRON-FORM} \quad \text{it} \\ \text{I} \end{array} \right] \\ \text{XCOMP} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'likely<XCOMP>SUBJ'} \\ \text{SUBJ} \quad [] \\ \text{XCOMP} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'rain<>SUBJ'} \\ \text{SUBJ} \quad [] \end{array} \right] \end{array} \right] \end{array} \right]$$

In contrast, if these constructions had a closed PREDLINK analysis, as in (16), then the control equation for adjectives like *likely* would have to be as in (17).

$$(16) \left[\begin{array}{l} \text{PRED} \quad \text{'be<PREDLINK>SUBJ'} \\ \text{SUBJ} \quad \left[\begin{array}{l} \text{PRON-FORM} \quad \text{it} \\ \text{I} \end{array} \right] \\ \text{PREDLINK} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'likely<COMP>'} \\ \text{COMP} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'rain<>SUBJ'} \\ \text{SUBJ} \quad [] \end{array} \right] \end{array} \right] \end{array} \right]$$

$$(17) \text{ likely} \quad (\uparrow \text{PRED})=\text{'likely<COMP>'} \\ (\uparrow \text{COMP SUBJ})=((\text{PREDLINK } \uparrow) \text{ SUBJ})$$

Although it is possible to write equations to enforce the desired subject control, these equations are quite complicated and do not follow general patterns of control equations found crosslinguistically.

The contrast between the copular constructions in section 2 and the ones in this section, both from English, suggests once again that within languages, there is variation as to the grammatical function of the post-copular element: some copular complements are open, and some are closed.

4 Agreement

The XCOMP analysis also works well for languages like French and Norwegian, in which the postcopular complement shows agreement with the subject of the copula. A French example is shown in (18).

- (18) a. Elle est petite.
 she.F.SG is small.F.SG
 'She is small.' (French)
- b. Il est petit.
 he.M.SG is small.M.SG
 'He is small.' (French)

In the XCOMP analysis, the adjective simply agrees with its own SUBJ, in the same way as verbs do. Consider the structure in (19), which provides the open complement (XCOMP) analysis for the sentence in (18a).

$$(19) \left[\begin{array}{l} \text{PRED} \quad \text{'be<XCOMP>SUBJ'} \\ \text{SUBJ} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'she'} \\ \text{NUM} \quad \text{sg} \\ \text{GEND} \quad \text{fem} \end{array} \right] 1 \\ \text{XCOMP} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'small<SUBJ>'} \\ \text{SUBJ} \quad [] 1 \end{array} \right] \end{array} \right]$$

Given a structure of this type, the adjective can have a basic lexical entry as in (20).

$$(20) \text{ petite} \quad (\uparrow \text{PRED}) = \text{'small<SUBJ>'} \\ (\uparrow \text{SUBJ NUM}) = \text{c sg} \\ (\uparrow \text{SUBJ GEND}) = \text{c fem}$$

In contrast, consider the structure in (21), which provides the closed complement (PREDLINK) analysis for the same sentence:

$$(21) \left[\begin{array}{l} \text{PRED} \quad \text{'be<SUBJ,PREDLINK>'} \\ \text{SUBJ} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'she'} \\ \text{NUM} \quad \text{sg} \\ \text{GEND} \quad \text{fem} \end{array} \right] \\ \text{PREDLINK} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'small'} \end{array} \right] \end{array} \right]$$

If we assume that the adjective specifies the agreement features of the subject of the sentence, the adjective must have a lexical entry such as (22) under the closed-complement analysis, since the agreement specifications are not associated with a local argument of the adjective.

$$(22) \text{ petite} \quad (\uparrow \text{PRED}) = \text{'small'} \\ ((\text{PREDLINK } \uparrow) \text{SUBJ NUM}) = \text{c sg} \\ ((\text{PREDLINK } \uparrow) \text{SUBJ GEND}) = \text{c fem}$$

Although it is possible to formulate equations so that agreement works in these copular constructions with a PREDLINK analysis, the XCOMP analysis allows for a much simpler analysis and one which is similar to that of other cases of subject-predicate agreement, such as subject-verb agreement.

In other languages, however, some considerations may weaken the status of agreement as an argument for assuming an XCOMP analysis. In languages like Norwegian, for example, there is no subject-verb agreement, so that subject-adjective agreement must be treated differently from subject-verb agreement in any case. Another issue is that predicative adjective agreement may be governed by semantic rather than syntactic features. Specifically, and unlike the case of agreement between attributive adjective and head noun, the form of the predicative adjective may be governed by properties of the intended referent rather than the grammatical properties of the subject, as in (23).

$$(23) \text{ a. Ekteparet er syke} \\ \text{the-married-couple.N.SG is ill.PL} \\ \text{'The couple are ill.' (Norwegian)}$$

- b. Postbudet er syk
the-mailman.N.SG is ill.M.SG
'The mailman is ill.' (Norwegian)
- c. Bil er dyrt
car.M.SG is expensive.N.SG
'Having/buying/using/... a car is expensive.' (Norwegian)

This suggests that the agreement in predicative adjectives is not a strict grammatical rule, and hence less compelling as an argument for assuming that the predicative adjective takes its own SUBJ.

5 Discussion

We have laid out several issues in the syntactic analysis of copular constructions, and in particular the role of the postcopular complement. Much more work needs to be done to determine the range of copular constructions found crosslinguistically and how these should be analyzed in the overall LFG syntactic framework. This initial examination of copular constructions suggests several conclusions that we hope can guide future analyses.

First, the XCOMP analysis is appropriate for some copular constructions but not for others, even within the same language. That is, languages may have both closed and open postcopular constituents whose occurrence may be governed by different factors such as the c-structure category of the constituent.

Second, the same construction can be open in some languages and closed in others. For example, postcopular adjectives appear to be open categories in some languages and closed ones in others; this may hold of other categories as well.

Third, we can determine whether a particular copular construction is open or closed by examining phenomena like agreement and the obligatory presence of the copula. We hope that future work will identify other syntactic tests which can be used to determine the status of a postcopular constituent both within and across languages.

Assuming that postcopular complements can be open, requiring a subject argument for completeness, raises a number of questions that we have not addressed here. One of the most pressing is: to what extent does this status carry over to other environments? In particular, the subject requirement of words when they function predicatively may be different than when they have a non-predicative role. For example, if the adjectives in (14) and (18), partially repeated here as (24), are open in postcopular use, should it follow that they also open in attributive position, (25)?

(24) a. They are eager/foolish/loathe to leave.

- b. Elle est petite.
she.F.SG is small.F.SG
'She is small.' (French)

(25) a. the foolish boy

- b. la petite fille
the.F.SG little.F.SG girl
'the little girl' (French)

This does not seem implausible for adjectives, especially in languages such as French with adjectival agreement, but is less so for PPs and particularly for NPs. That is, it seems unlikely that every NP in a given

language, regardless of the syntactic construction in which it appears, requires a subject. We thus speculate that open postcopular complements, which occur in a predicative environment, need not always be open in other, non-predicative environments. For example, postcopular noun phrases may be open in a language, but closed in other environments.

Finally, it is our hope that examining these constructions can illuminate the difference between a grammatical function and a semantic argument, and that untangling the two notions can lead to a better understanding of their interaction in LFG.

Acknowledgments

This paper is based in presentations and discussions conducted over many years within the ParGram project (www.parc.com/istl/groups/nltp/pargram/). We are grateful to the participants of ParGram, in particular Hiroshi Masuichi, and to the participants of LFG04 for helpful discussion.

Bibliography

- Andrews, A. (1982). "The Representation of Case in Modern Icelandic". In Joan Bresnan (ed.) *The Mental Representation of Grammatical Relations*, pp. 427-503. The MIT Press, Cambridge, MA.
- Butt, M., King, T. H., Niño, M., and Segond, F. (1999). *A Grammar Writer's Cookbook*. CSLI Publications, Stanford University.
- Chvany, C. (1975). *On the Syntax of BE-sentences in Russian*. Slavica Publications.
- Falk, Y. (2004). "The Hebrew Present-tense Copula as a Mixes Category". In M. Butt and T. H. King (eds.) *Proceedings of the LFG04 Conference*. CSLI Publications, Stanford University. (This volume.)
- Ionin, T., and R. Matushansky. (2002). "Elephants, times, and predicate case in the Russian copular". Presented at *Existence*, Nancy, France. Handout available on-line.
- Kaplan, R. and J. Maxwell. (1996). *LFG Grammar Writer's Workbench*. Technical report, PARC. Available from: <ftp://ftp.parc.xerox.com/pub/lfg/lfgmanual.ps>.
- Keenan, Edward L. (1976). "Towards a universal definition of subject". In C. N. Li (editor), *Subject and Topic*. New York: Academic Press.
- Rosén, V. (1996). "The LFG Architecture and 'Verbless' Syntactic Constructions". In M. Butt and T. H. King (eds.) *Proceedings of the LFG96 Conference*. CSLI Publications, Stanford University.
- Shlonsky, U. (1997). *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. Oxford University Press.