

IRISH CLEFTING AND INFORMATION-STRUCTURE

Sebastian Sulger
Universität Konstanz

Proceedings of the LFG09 Conference

Miriam Butt and Tracy Holloway King (Editors)

2009

CSLI Publications

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

Abstract

This paper presents an analysis of Irish clefting couched within Lexical-Functional Grammar. In Irish, cleft sentences are formed using two syntactic permutations. First, a copula is introduced, taking the clefted phrase or word as a predicate. Second, a relative clause is formed containing the remaining material of the original sentence. This basic pattern is valid across a variety of languages. I present different approaches towards copula predication taken within Lexical-Functional Grammar and discuss my analysis of Irish copula constructions. Based on the insights from simple copula predication, I derive my analysis for Irish clefting, claiming that in principle, a parallel syntactic approach for both simple copula clauses and clefting can be assumed. The syntactic analysis of copula clefting has been implemented in a computational LFG grammar using the XLE software.

Next, pragmatic aspects of clefting are discussed. Cleft sentences are not only interesting from a syntactic point of view, but also are used to separate new information from old information. In the modular architecture of Lexical-Functional Grammar, additional levels of representation may be added to allow for extra-syntactic analysis. I use the projection of i(nformation)-structure to map strings in the sentence to discourse functions. Applying this type of analysis to clefting, I arrive at a more complete picture of the form and function of Irish clefts.

1 Introduction

In this paper, I present an analysis of Irish clefting. The analysis extends not only over the syntactic aspects of clefting, but also presents ideas related to the information-structural consequences of clefting. In Modern Irish, as in other languages, the pattern for clefting relies on the use of the copula verb. The basic pattern for clefting in Irish is *copula – XP – relative clause* (Stenson, 1981). I argue that the relative clause in fact resembles a clausal subject, in which the nominal head is understood, though not overtly expressed. The clefting pattern in that case is identical to the general copula sentence pattern; my conclusion is that one syntactic analysis can and should account for both general copula predication and clefting constructions. The difference between the two constructions are to be found in the information-structure of the sentences. I give ideas on how to analyze discourse functions for cleft sentences using a separate projection within LFG architecture called i-structure (King, 1997; Butt and King, 1998; Andréasson, 2007).

Section 2 presents different copula analyses within the LFG literature and discusses why I chose a PREDLINK analysis for Irish copula. In section 3, I provide clefting data from Irish and show the connection between simple copula sentences and cleft sentences. Section 4 discusses consequences of clefting for the

[†]I thank my advisor, Miriam Butt, for many valuable comments on different versions of this paper, and Ingo Mittendorf and Louise Mycock for comments during the LFG09 Conference. Also, I thank the anonymous reviewers of the paper abstract for their constructive criticism.

information-structure of a sentence. The syntactic analysis of Irish copula and clefting has been implemented in a computational LFG grammar of Irish using the XLE grammar writing platform (Crouch et al., 2008).

2 Copula Analyses in LFG

I argue that understanding copula predication is essential to understanding clefting patterns. I therefore first take a look at different copula analyses in the LFG literature. There are three different copula analyses across the literature: a single-tier analysis, an open-complement double-tier analysis (XCOMP) and a closed complement double-tier analysis (PREDLINK) (Attia, 2008).

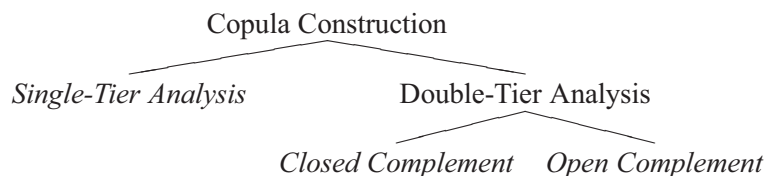


Figure 1: Copula construction analyses in LFG

2.1 Single-Tier Analysis

The single-tier analysis involves the copula predicate (i.e., the adjective in the example below) functioning as a sentential head, meaning that it selects for a subject. Dalrymple et al. (2004) note that this is a preferable analysis for cases in which the copula is optional, such as with Japanese predicative adjectives.

- (1) a. hon wa akai.
 book TOP red
 ‘The book is red.’ (Dalrymple et al., 2004, p. 190)
- b. sono hon wa akai desu.
 this book TOP red is
 ‘This book is red.’ (Dalrymple et al., 2004, p. 191)

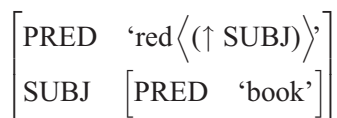


Figure 2: Single-tier analysis in Japanese

A single-tier analysis is problematic for two reasons. First, it has to be assumed and shown that the copula predicate can in fact select for a subject. Second, the copula in Japanese is not optional, but has to surface with nominal predicates. Dalrymple et al. (2004) assume that adjectives are able to select for subjects, while

nouns are not. A different type of analysis would have to be assumed for nominal predicates. This is not ideal since the predication in both cases is the same — the optionality of the copula is no sufficient motivation for assuming two separate analyses of the copula (Attia, 2008).

2.2 Open-Complement Double-Tier Analysis (XCOMP)

In the open-complement double-tier analysis, the copula predicate does not function as a sentential head. The head is the copula itself, when it is present, or a null element, when the copula does not surface. The copula clause predicate selects for a subject which is functionally controlled by the main clause. This means that the subject of the main clause (i.e., the subject of the copula) is unified with the subject of the copula predicate. In LFG, a control relation of this type is constructed using an XCOMP function. The partial XCOMP f-structure alone does not contain a subject value, i.e., is not complete; the subject function is filled with a clause-external subject via functional control through the linking verb (Bresnan, 1982).

It has been argued that this a preferable analysis for cases in which the predicate shows agreement with its subject. Dalrymple et al. (2004) give the example of French adjectives in predicate position, agreeing with the subject of the main clause in the same way as verbs do. The following examples are taken from Dalrymple et al. (2004).

- (2) a. Il est petit.
 he.Masc.Sg COP.3P.Sg.Pres small.Masc.Sg
 ‘He is small.’ (Dalrymple et al., 2004, p. 195)
- b. Elle est petite.
 she.Fem.Sg COP.3P.Sg.Pres small.Fem.Sg
 ‘She is small.’ (Dalrymple et al., 2004, p. 195)

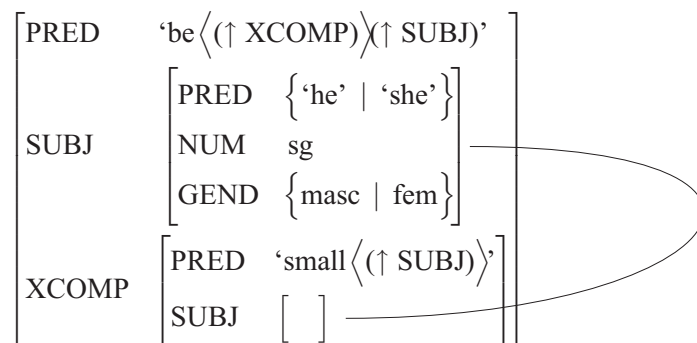


Figure 3: Open double-tier analysis of French copula

Dalrymple et al. (2004) argue that the agreement is a strong indication for a control relation between the subject and the predicate. They also state that if one assumes an XCOMP type of analysis, one can write basic lexical entries for the predicate as in (3).

- (3) *petite* (\uparrow PRED) = 'small < (\uparrow SUBJ) >'
 (\uparrow SUBJ NUM) = c sg
 (\uparrow SUBJ GEND) = c fem

Attia (2008) counters this analysis. He makes the point that in French, the agreement of the copula predicate with the subject is not the same type of agreement verbs show with their subjects. He maintains that the agreement is not enough evidence to assume that the copula predicate subcategorizes for a subject.

Another important argument contra the XCOMP open complement analysis comes from a predicational perspective. The analysis in Figure 3 is not suitable for copula constructions according to Attia (2008) because this is exactly the way normal subject raising verbs (such as *seem*, *appear*) are analyzed in LFG. Assuming an f-structure like Figure 3 would therefore mean that there is no difference between copula constructions and subject raising verbs.

Dalrymple et al. (2004) provide the most compelling evidence against the open-complement XCOMP-type analysis. In cases where the post-copular complement already *has* a subject which is different from the subject of the main clause, the closed complement PREDLINK analysis is the preferred analysis. See the examples in (4).

- (4) a. The good thing is that he did not throw the snowball.
 b. The main goal is (for the student) to succeed in the exam.

If we assume an XCOMP analysis for sentences like these, the result is a clash of PRED values, i.e., because of the control equations, the XCOMP f-structure would contain two subjects which are not unifiable (Attia, 2008; Dalrymple et al., 2004; Butt et al., 1999). See the illformed f-structure in Figure 4 for sentence (4a).

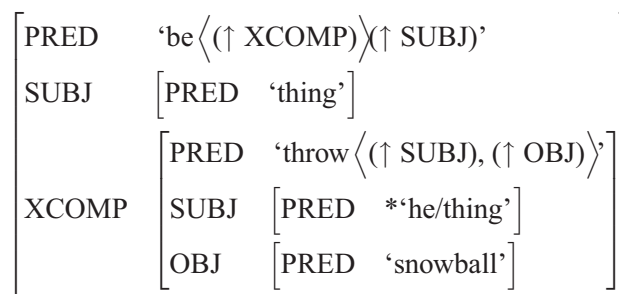


Figure 4: Clashing open double-tier analysis with divergent subjects

2.3 Closed-Complement Double-Tier Analysis (PREDLINK)

Another alternative is the so-called closed complement analysis; “closed” because the PREDLINK function is a closed function and therefore does not allow functional control. The PREDLINK analysis models the fact that a particular property is predicated of the subject in a syntactically reasonable way. The main PRED of the f-structure expresses that a specific property is predicated of the subject. Exactly this is captured by the PREDLINK function (Butt et al., 1999).

The closed complement double-tier analysis is a universal LFG analysis for copula constructions according to both Attia (2008) and Butt et al. (1999). The main advantage of this approach is that since it does *not* necessarily rely on control equations, it does not have any issues with sentences such as in (4). On the other hand, when control equations become mandatory (i.e., when there is obvious agreement), these can also be formulated in a closed complement PREDLINK analysis (Attia, 2008; Butt et al., 1999), although the equations are slightly more troublesome in cases of long-distance agreement. Therefore, the main arguments for an open complement XCOMP type of analysis presented by Dalrymple et al. (2004) are not enough to motivate a pluralist approach to copula constructions in terms of analyses, since all of the advantageous properties of the XCOMP analysis can be reproduced within a PREDLINK closed complement analysis. Attia (2008) maintains that syntactic features such as the presence and absence of the copula form and the presence and absence of agreement on the predicate do not affect the syntactic function of the predication. Only because languages like French show agreement on the predicate, copula predication does not necessarily require diverging syntactic analyses.

There are several advantages to the double-tier closed complement PREDLINK analysis. First, it does not matter what kind of constituent the copula complement is; this analysis seems to be the only one that succeeds in providing valid representations for all constituent types, which can take different semantic roles; see Attia (2008) for an overview. Other approaches seem to have problems with this unified approach. Bresnan (2001), assuming that adjectives can subcategorize for subjects, also assumes that nouns and prepositional phrases can do so. To account for this, she proposes to manipulate the PRED of the noun or preposition by means of lexical rules; see the sentences in (5) and the corresponding rules in (6), cited by Lodrup (2008).

(5) a. The pills made him a monster. (Lodrup, 2008, p. 22)

b. She seems in a bad mood. (Lodrup, 2008, p. 22)

(6) a. ‘monster’ =>
‘be-a-monster<(↑ SUBJ)>’ (Lodrup, 2008, p. 22)

b. ‘in<(↑ OBJ)>’ =>
‘be-in-a-state-of<(↑ SUBJ), (↑ OBJ)>’ (Lodrup, 2008, p. 22)

Both Attia (2008) and Lodrup (2008) find this approach problematic, since it not only results in artificial and complex annotation, but also presupposes that any PP or NP in a given language can in principle subcategorize for a subject. Dalrymple et al. (2004) and also Rosén (1996) in an earlier paper maintain that this type of analysis is certainly not desirable. Within the closed complement analysis, these problems vanish, since there is no XCOMP f-structure, hence we do not need any subject that is functionally controlled. I give the f-structure for (6b) in Figure 5, assuming a closed complement double-tier analysis.

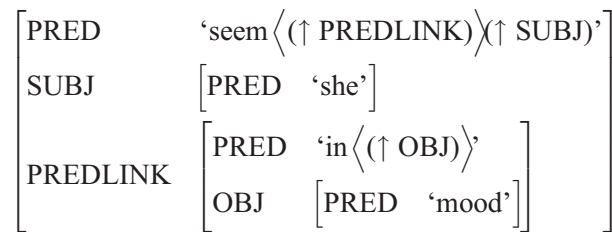


Figure 5: Well-formed closed double-tier analysis of *She seems in a bad mood*.

In sentences where there is no copula, the PREDLINK analysis has intuitive appeal, since, at the f-structure level, it mirrors the juxtaposition of constituents when the copula is missing (Attia, 2008). As many languages contain copula-less sentences (Carnie, 1995), the analysis has cross-linguistic appeal. Attia (2008) further argues that the presence vs. absence of the copula itself is a parameter of variation. Since the copula is generally considered as semantically empty, there is no functional distinction to be made between sentences containing the copula and sentences without the copula. The predication in the absence of the copula is modeled using a `null-be` predicator in the LFG rule notation. See the sentence in (7) and the rule in (8); the resulting f-structure is shown in Figure 6.

- (7) hwa ṭālibun
 he student
 'He is a student.' (Attia, 2008)

- (8) S --> NP: (↑ SUBJ) = ↓;
 ε: (↑ PRED) = 'null-be' < (↑ SUBJ), (↑ PREDLINK) >;
 (↑ TENSE) = pres;
 {NP | AP}: (↑ PREDLINK) = ↓
 (↓ GEND) = (↑ SUBJ GEND)
 (↓ NUM) = (↑ SUBJ NUM)

PRED	‘null-be $\langle (\uparrow \text{SUBJ}), (\uparrow \text{PREDLINK}) \rangle$ ’
SUBJ	[PRED ‘he’]
PREDLINK	[PRED ‘student’]
TENSE	pres

Figure 6: Closed double-tier analysis without surface copula in Arabic

Note that the tense feature is provided by the empty element ε , which captures the insight that the copula in Arabic can only be omitted in present tense. If a copula was present in the sentence, then the tense feature would be provided by the copula itself.

To sum up the discussion about the different approaches towards copula constructions in LFG, I stress that each one of the possible analyses has its advantages and disadvantages. While we have to assume a subcategorization frame for predicate elements in the single-tier and open-complement types of analysis which might be less appropriate for some languages than for others, the PREDLINK approach is more neutral in this respect. I think in this discussion it is important to see LFG in the context of parallel (i.e., cross-linguistic) grammar designing. As long as there are no serious reasons to object to the PREDLINK analysis (e.g., the case of Abkhaz discussed above), I maintain that a universally applicable analysis should be favored to increase the cross-linguistic appeal of LFG.

2.4 Towards an Analysis of Irish Copula in LFG

In this section, I present my own analysis of Irish copula predication in LFG. This analysis provides the basis for the approach to Irish cleft constructions since simple cleft constructions in Irish rely on copula predication.

Consider the examples in (9). They all contain a copula, a predicate and a subject, nothing more and nothing less, as indicated by the bracketing.

- (9) a. Is [le Pól]_{PRED} [an carr.]_{SUBJ}
 COP.Pres with Paul ART.Def.Sg car
 ‘The car belongs to Paul.’ [lit. ‘The car is with Paul.’]
- b. Ba [dhuine deas]_{PRED} [é.]_{SUBJ}
 COP.Past man nice he
 ‘He was a nice man.’
- c. Is [maith li-om]_{PRED} [tae.]_{SUBJ}
 COP.Pres good with-me tea
 ‘I like tea.’

Consider the arguments given by Dalrymple et al. (2004) in favor of a divergent analysis for copula constructions across languages and within a certain lan-

guage. First, agreement is given as an argument for an XCOMP (open complement double-tier) analysis. In Irish, however, the predicate does not show any agreement whatsoever with the subject. Therefore, I maintain that there is no reason concerning agreement to choose XCOMP as a possible copula analysis for Irish.

Second, in cases where the copula is absent, Dalrymple et al. (2004) argue for a special analysis: the single-tier analysis. They present the case of Japanese, where the occurrence of the copula is governed by the category of the predicate. They propose that different analyses have to be assumed depending on the presence or the absence of the copula. Attia (2008) however, as we have seen, maintains that this is merely a case of stylistic variation. In principle, the predication is the same, no matter if the copula is present or not.

In Irish, the factor governing the occurrence of the copula is tense. The copula may be overt or dropped in the present tense, but its occurrence is mandatory when tensed for future or past. Still, the predication in the sentences does not change in principle, whether the copula is present or not. Therefore, I maintain that the presence of the copula is a means of stylistic variation in the present tense, but must be present in clauses with future or past tense. The presence vs. absence of the copula does not lead to semantic differences (see also Ó Siadhail, 1989; Stenson, 1981). Additionally, the present tense copula may not be deleted when the negative form is used (Ó Siadhail, 1989). This can also be modeled via c-structure rule annotations.

I follow Butt et al. (1999) and Attia (2008) in claiming the closed complement double-tier analysis as a universally applicable analysis for copula constructions. The possible variation in the choice of the predicate constituent is immediately reflected by the variation in the PREDLINK f-structure. Different lexical categories can be head of the predicate, hence head of the PREDLINK f-structure. A sample analysis for sentence (9c) is given in Figure 7. In (9c), the head of the predicate is an adjectival construction (i.e., an adjective head *maith* ‘good’ with a prepositional phrase adjunct *liom* ‘with me’).

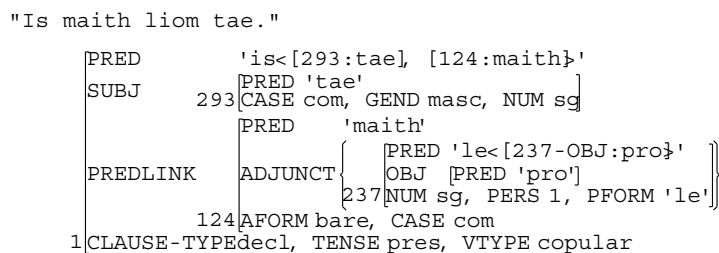


Figure 7: Sample copula analysis using PREDLINK

3 Irish Clefting

This section provides some data from Irish clefting. I compare a special kind of Irish copula construction, called “identification sentence” in the literature (e.g., Stenson, 1981; Ó Siadhail, 1989), to Irish clefts, suggesting that these constructions are very similar. As a consequence, the syntactic analysis of simple Irish clefts becomes straightforward and fits into the closed complement double-tier PREDLINK analysis introduced in the LFG literature perfectly.

3.1 The Irish Data

In this section, I present data illustrating simple Irish clefting. The data shows the general form of clefts and the variability in the choice of the clefted constituent. All of these sentences have the same basic structure. A constituent of the basic sentence appears after the copula as its predicate and a relative clause is formed containing the rest of the sentence, including the main verb. These are the basic syntactic facts given throughout relevant literature (Stenson, 1981; Ó Siadhail, 1989; The Christian Brothers, 1960).

In the following examples, I provide the basic (i.e., non-clefted) sentence and variations of that sentence. That means, for example, (10b) and (10c) are permutations of (10a), and so on. When forming a cleft sentence, the NP can appear in a position before the main verb of the sentence, violating the basic VSO word order of Irish. It is possible to front subject NPs and object NPs. (10c) is another permutation of (10a) whereby the adverbial *inné* ‘yesterday’ occurs in the cleft.

- (10) a. Léigh an múinteoir leabhar inné.
read.Past ART.Def.Sg teacher book yesterday
‘The teacher read a book yesterday.’
- b. Is é an múinteoir a léigh leabhar inné.
COP.Pres AGR.Masc.Sg ART.Def.Sg teacher COMP.Rel read.Past
book yesterday
‘It is the teacher who read a book yesterday.’
- c. Is inné a léigh an múinteoir leabhar
COP.Pres yesterday COMP.Rel read.Past ART.Def.Sg teacher book
‘It is yesterday that the teacher read a book.’

In (11b), the aspectual verb phrase *ag péinteáil cathaoir* ‘painting a chair’ in (11a) occurs in fronted position. The aspectual phrase here expresses progressive action; see Ramchand (1997) for a discussion of Scottish Gaelic aspectual phrases. It also takes on the function of a predicate of the substantive verb¹ in the matrix

¹In Irish, there are two verbs for *to be*: *bí* (often called *substantive verb* in the literature) and

sentence (i.e., it functions as the main verb). In (11c), the object NP *cathaoir* ‘a chair’ is fronted.

- (11) a. Bhí an fear ag péinteáil cathaoir inné.
 be.Past ART.Def.Sg man ASP paint.VN chair yesterday
 ‘The man was painting a chair yesterday.’
 (Ó Siadhail, 1989, p. 236)
- b. Is ag péinteáil cathaoir a bhí an fear
 COP.Pres ASP paint.VN chair COMP.Rel be.Past ART.Def.Sg man
 inné.
 yesterday.
 ? ‘It is painting a chair that the man was (doing) yesterday.’
 (Ó Siadhail, 1989, p. 236)
- c. Is cathaoir a bhí an fear ag péinteáil
 COP.Pres chair COMP.Rel be.Past ART.Def.Sg man ASP paint.VN
 inné.
 yesterday
 ‘It is a chair that the man was painting yesterday.’
 (Ó Siadhail, 1989, p. 236)

(12) and (13) show the possibilities of adjective fronting. In (12b) the predicative adjective *te* ‘hot’ is fronted; it is not possible to front an attributive adjective out of an NP, as is shown by (13b).

- (12) a. Tá sé fuar.
 be.Pres it cold
 ‘It’s cold.’
- b. (Níl sé fuar.) Is te a-tá sé.
 be.Pres.Neg it cold COP.Pres hot COMP.Rel-be.Pres it.
 ‘(It’s not cold.) It is *hot*.’ [lit. ‘It is hot that it is.’]
- (13) a. Léigh sí an leabhar dearg.
 read.Past she ART.Def.Sg book red
 ‘She read a red book.’
- b. * Is dearg a léigh sí an leabhar.
 COP.Pres red COMP.Rel read.Past she ART.Def.Sg book

is (the copula). The copula verb *is* generally takes essential and inherent qualities as predicatives, hence the copula predicates are most commonly noun phrases, such as occupation, nationality, group membership and the like. The substantive verb *bí*, on the other hand, takes as predicatives less inherent qualities, such as temporal specifications, location, (temporal) possession etc. (Ramchand, 1997; Stenson, 1981). This difference in predication is called stage vs. individual level predication in Ramchand (1997).

As is obvious from examples like (11b) and (12b), English is more rigid concerning the choice of the clefted material. While the English equivalent of (11b) is at least questionable, even with the insertion of the verbal residue *doing*, the literal translation presented in (12b) is completely unacceptable.

(13) shows that, as in English and other languages, it is not possible to front an attributive adjective, while (12) shows that it is possible, on the other hand, to front predicative adjectives. The clefting of adjectives often has a marked connotation as a result, as it is the case in (12b). Both Ó Siadhail (1989) and Stenson (1981) note that an exclamatory or question form in these cases is often preferred, as in (14).

- (14) Nach te a-tá sé!
 COP.Pres.NegQ hot COMP.Rel-be.Pres it
 ‘Isn’t it *hot!*’

The structure of clefts seems, at first glance, not to have much in common with normal copula constructions. In the next section, however, I argue along the lines of Stenson (1981) that cleft sentences in fact have much in common with copula constructions — especially with identification sentences.

3.2 Copula Constructions, Questions, Clefting: Some Similarities

Looking at certain types of copula constructions and constituent questions suggests that there are structural similarities between these and cleft sentences. Consider the case of constituent questions, which can be seen as a device for distinguishing presupposed from new (i.e., focused) information. If someone asks the question in (15), then the anticipated referent of the WH-phrase is in the focus, and the remainder of the question is presupposed (see also Stenson, 1981; Sornicola, 1996).

- (15) What cleared the road? (The snowplow cleared the road.)

Interestingly, in Irish, the structure of constituent questions is very similar, if not identical to the structure of simple clefts. Consider (16).

- (16) a. Cé a léigh leabhar?
 who COMP.Rel read.Past book
 ‘Who read a book?’
 b. Is é an múinteoir a léigh
 COP.Pres AGR.Masc.Sg ART.Def.Sg teacher COMP.Rel read.Past
 leabhar.
 book
 ‘It is the teacher who read a book.’

The syntactic similarity between (16a) and (16b) is obvious. While Ó Siadhail (1989) notes that the copula is mostly dropped in the context of clefts, Stenson (1981) mentions:

Indeed, one of the commonest situations for copula deletion is the clefted answer to a Wh-question, suggesting that the opposition between old and new information that the copula expresses has already been established by the question itself. (Stenson, 1981, p. 108)

So, a more appropriate answer to the question in (16a) would be (17), where the copula form is deleted from the surface. The similarity between the question phrasing and the cleft becomes even more obvious then.²

- (17) An múinteoir a léigh leabhar.
 ART.Def.Sg teacher COMP.Rel read.Past book
 ‘It is the teacher who read a book.’

Finally, and most importantly, consider the similarities between copula sentences identifying two NPs with one another and clefts with an NP as the clefted constituent. If we try to distinguish these structures from one another by simply looking at the surface, we might conclude that simple cleft sentences differ from identification sentences only in that they seem to lack one of the NPs which are identified with each other. (18a) is a copula sentence, equating the two NPs as indicated by the bracketing. (18b) is a cleft sentence.

- (18) a. Is é [mo dheartháir] [an fear a
 COP.Pres AGR.Masc.Sg my brother ART.Def.Sg man COMP.Rel
 bhí tinn inné.]
 be.Past sick yesterday
 ‘My brother is the man who was sick yesterday.
- b. Is é [an fear] [a bhí tinn
 COP.Pres AGR.Masc.Sg ART.Def.Sg man COMP.Rel be.Past sick
 inné.]
 yesterday
 ‘It is the man who was sick yesterday.’

Stenson (1981) claims that there is no surface head to the relative clause in cleft sentences such as (18b). Now consider the sentence in (19), which is in fact a pseudo-cleft sentence, using the same type of paraphrase as English pseudo-clefts.

- (19) Is é [mo dheartháir] [an té a
 COP.Pres AGR.Masc.Sg my brother ART.Def.Sg one COMP.Rel
 bhí tinn inné.]
 be.Past sick yesterday
 ‘My brother is the one who was sick yesterday.’

²Note that the agreement marker *é* is dropped together with the copula, which is noted by both Stenson (1981) and Ó Siadhail (1989). Further research relating to the exact role of this marker, and why it is dropped here, might turn out to be interesting.

In pseudo-clefts, the relative clause does have a nominal head; therefore, they resemble run-of-the-mill copula sentences such as (18a) even more. In fact, (19) and (18a) are identical in structure. The nominal head inserted in the pseudo-clefts does not add any semantics, but serves as a syntactic head for the relative clause. Stenson (1981) notes that when the nominal head is not overt, then the head of the relative clause in a cleft sentence like (18b) is understood to be referring to a human referent. In other sentences, the nominal head might not refer to a human referent; then, another provisory head noun (e.g., *rud* ‘thing’ instead of *té* ‘one’) is overt in pseudo-clefts.

The nominal heads in these sentences are understood, even when they do not surface. The conclusion is that cleft sentences such as the one in (18b) are derived from pseudo-clefts such as the one in (19). Since (19) in turn resembles copula sentences like (20a), the structure in (20) is proposed for cleft sentences with an NP in predicate position. Compare this to the structure of the copula sentence in (18a), which is given in (21).

(20) COP [NP]_{predicate} [relative clause]_{subject}

(21) COP [NP]_{predicate} [NP – (relative clause)]_{subject}

In cleft sentences, the subject nominal which is overt in pseudo-clefts does not surface, leaving the relative clause alone as a sentential subject to the matrix clause. The subcategorization frame of the copula is filled by the NP in predicate position which is assigned the PREDLINK function, and the relative clause, analyzed as a sentential subject, which is assigned the SUBJ function. An example of the overall syntactic analysis is given in the f-structure in Figure 9 for the sentence in (22).^{3,4}

(22) Is é an múinteoir a léigh
 COP.Pres AGR.Masc.Sg ART.Def.Sg teacher COMP.Rel read.Past
 leabhar.
 book
 ‘It is the teacher who read a book.’

The main predicate of the whole sentence is supplied by the copula, which in turn has two positions to be filled in its subcategorization frame: subject (SUBJ) and predicate (PREDLINK). The relative clause, the head of which is *léigh* ‘read’, is assigned the topmost SUBJ function, filling one place in the copula’s frame — as indicated by the indices on the f-structures. The NP in predicate position is assigned the PREDLINK function, thus filling the other place in the copula’s frame, while at the same time supplying the SUBJ function within the relative

³(22) is ambiguous between the syntactic reading in Figure 9 and another, admittedly improbable reading where *leabhar* ‘book’ is assigned the SUBJ function and *múinteoir* ‘teacher’ is assigned the OBJ function — both the object and the subject NP of the relative clause may appear in predicate position.

⁴The f-structure in Figure 9, produced by the XLE software, has been simplified for space reasons.

"Is é an múinteoir a léigh leabhar."

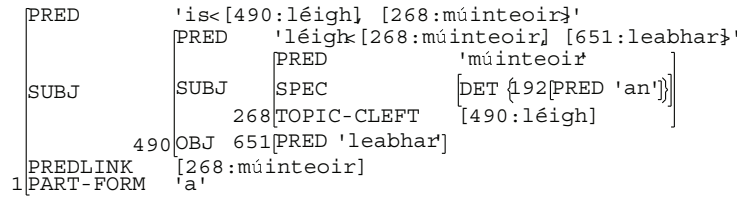


Figure 8: Cleft sentence with NP in predicate position: f-structure

clause — again, indices show how that the NP in predicate position is mapped to different f-structures: the PREDLINK f-structure and the SUBJ f-structure inside the relative clause.

3.3 Variability in Choosing the Clefted Constituent

Note that (21) applies to only those clefts that have an NP in predicate position. It does not explain why other constituents — prepositional phrases, adverbial phrases, adjectives — can also appear in predicate position. I therefore extend Stenson's (1981) view, arguing that not only clefts with NPs in predicate position have the same structure as so-called "identification sentences", but clefts with other constituents in predicate position can be analyzed accordingly. Evidence for this comes from sentences such as in (23). In these examples, the substantive verb *bí* 'be' — the stage level predicate, to go with Ramchand's (1997) terminology — is used to link adjective phrases/prepositional phrases/adverbials and a noun phrase. Note that although word order is different — the subject comes before the predicate, rather than after it — the main syntactic circumstances in these sentences are the same as in the copula examples: the verb is used to link the predicate and the subject. It is the semantics of this predication that is different from the one in copula sentences (Ramchand, 1997; Stenson, 1981). In (24), I give cleft sentences corresponding to the sentences in (23). The fact that the clefts in (24) are formed from sentences with the *substantive* verb follows naturally from the differentiation between stage level and individual level predication — predicates like *tim* 'sick', *faoin mbord* 'under the table' or *amárach* 'tomorrow' cannot be inherent to the subject, but refer to a situation (Ramchand, 1997).

- (23) a. Tá mo dheartháir tinn.
 be.Pres my brother sick
 'My brother is sick.'
- b. Tá an cat faoi-n mbord.
 be.Pres ART.Def.Sg cat under-ART.Def.Sg table
 'The cat is under the table.'

c. Beidh an fhleá amárach.
 be.Fut ART.Def.Sg party tomorrow
 ‘The party will be tomorrow.’

(24) a. Is tinn a-tá mo dheartháir.
 COP.Pres sick COMP.Rel-be.Pres my brother
 ‘My brother is *sick!*’ [lit. ‘It is sick that my brother is!’]

b. Is faoi-n mbord a-tá an cat.
 COP.Pres under-ART.Def.Sg table COMP-Rel-be.Pres ART.Def.Sg cat
 ‘It is under the table where the cat is.’

c. Is amárach a beidh an fhleá.
 COP.Pres tomorrow COMP.Rel be.Fut ART.Def.Sg party
 ‘The party will be *tomorrow!*’ [lit. ‘It is tomorrow that the party will be!’]

The clefting pattern stays the same across the examples: the relative clause containing the main verb (*bí* ‘be’ in these cases) follows the sentence-initial copula and the predicate; see (25). The template for simple stage level predication using the substantive verb *bí* is given in (26).

(25) COP [{AP | PP | ADVP}]_{predicate} [relative clause]_{subject}

(26) *bí* [NP]_{subject} [{AP | PP | ADVP}]_{predicate}

Below I give two sample analyses that show how the PREDLINK analysis can easily be extended to deal with stage level predication in (27) and a clefted version of the same sentence in (28). The corresponding f-structures are given in Figures 9 and 10.

(27) Beidh an fhleá amárach.
 be.Fut ART.Def.Sg party tomorrow
 ‘The party will be tomorrow.’

"Beidh an fhleá amárach."

$$\left[\begin{array}{l} \text{PRED} \quad \text{'bí} \langle [611:\text{fleá}], [355:\text{amárach}] \rangle \\ \text{SUBJ} \quad \left[\begin{array}{l} \text{PRED 'fleá'} \\ \text{SPEC} \left[\text{DET} \langle [110:\text{an}] \rangle \right] \end{array} \right] \\ \text{1} \text{PREDLINK} \text{355} \left[\text{PRED 'amárach'} \right] \end{array} \right]$$

Figure 9: Stage level predication with ADVP as predicate: f-structure

(28) Is amárach a beidh an fhleá.
 COP.Pres tomorrow COMP.Rel be.Fut ART.Def.Sg party
 ‘The party will be *tomorrow!*’ [lit. ‘It is tomorrow that the party will be!’]

"Is amárach a beidh an fhleá."

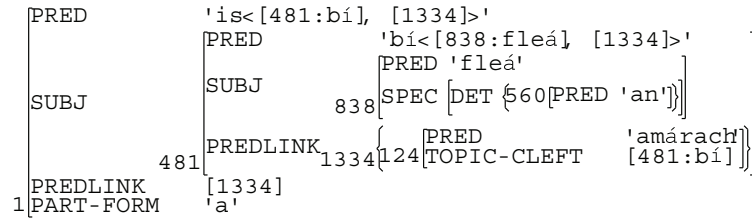


Figure 10: Cleft sentence with ADVP in predicate position: f-structure

Having established the variability in the choice of the constituent appearing in predicate position, the clefting pattern can be modified from (22) and (27) to (29).

(29) COP [XP]_{predicate} [relative clause]_{subject}

I conclude that *syntactically*, copula predication in simple copula sentences such as (20a) on the one hand and stage level predication in sentences such as (27) on the other hand are not different from cleft sentences such as (20b) or (28). A single syntactic analysis — the PREDLINK analysis — has been shown to be able to account for both constructions in a straightforward way. The advantage of such an analysis is that it is kept in parallel for constructions which involve the same type of syntactic processes, demonstrating the universal applicability of the PREDLINK analysis, as mentioned by Attia (2008) and Butt et al. (1999).

4 Information-Structure and Clefting

Much recent work in LFG has focused on integrating discourse functions like topic and focus in the grammar architecture (e.g., King, 1997; Butt and King, 1998; O'Connor, 2004; King and Zaenen, 2004; Andréasson, 2007). Discourse functions can be used to encode extra-syntactic cues within a sentence — cues indicating how the information in the sentence is structured and fits inside a speaker-hearer dialogue. Discourse functions (DF) hence encode Information-Structure (IS), providing an extra level of markup, which can be useful in further semantic processing or other applications that need to access IS, such as anaphora resolution or summarization (e.g., King and Zaenen, 2004).

Traditionally, DFs have been encoded directly in the f-structure (e.g., Bresnan and Mchombo, 1987; Bresnan, 2001). Since, however, there may be mismatches between grammatical functions and DFs, information structure can also be encoded using a separate projection in LFG, called i-structure (e.g., Butt and King, 1998; King and Zaenen, 2004; O'Connor, 2004; Andréasson, 2007) and mapped on top of c-structure. In the remainder of the paper, I present my ideas on how an analysis of cleft sentences can be extended to encode information structure; I use the i-structure projection to do so. The discourse functions I use for the purpose of this

paper, i.e., for describing information-structure within simple Irish clefts, are based on Vallduví (1993a). A simple two-way distinction for structuring information is proposed — *focus* vs. *ground*. Focus is the “informative, newsy, dominant, or contrary-to-expectation” part of the sentence, while ground is the “noninformative, known, or expected” part of the sentence (Vallduví, 1993a). In a cleft sentence, the “newsy” focus part of the information appears in the cleft as the predicate, while the “known” ground part of the information appears in the relative clause — this generalization holds cross-linguistically (Halvorsen, 1977; Declerck, 1988). Figure 11 presents such a mapping for the cleft sentence in (28).

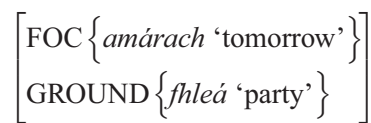


Figure 11: Cleft sentence with focus NP: i-structure

In cases like this, it is easy to assign focus, as there is only one candidate PRED for focus in each sentence — the adverbial phrase *amárach* ‘tomorrow’. In other cases, the mapping between c- and i-structure is not as straightforward, and the focus-ground-division is not as easy. The clefted constituent may contain multiple PREDs and/or contain elements with contrastive stress. In such cases, elements that are part of the clefted constituent, but do not carry contrastive stress are analyzed as being part of the ground (Vallduví, 1993b); see also Halvorsen (1977) and Declerck (1988). (30) is an example.

- (30) Is faoi-n mbord a bhí an cat.
 COP.Pres under-ART.Def.Sg table COMP.Rel be.Past ART.Def.Sg cat
 ‘It is under the table where the cat was.’

From the c-structure of this sentence, three different i-structures can be projected, since it is not clear whether the whole PP *faoi mbord* ‘under the table’ or just the preposition *faoi* ‘under’ or just the preposition’s object *mbord* ‘table’ should be assigned the DF FOC. See Figures 12-14.

$$\left[\begin{array}{l} \text{FOC} \{ \textit{faoin mbord} \text{ 'under the table'} \} \\ \text{GROUND} \{ \textit{cat} \text{ 'cat'} \} \end{array} \right]$$

Figure 12: Cleft sentence with focus PP: i-structure

$$\left[\begin{array}{l} \text{FOC} \{ \textit{faoin} \text{ 'under'} \} \\ \text{GROUND} \{ \textit{cat} \text{ 'cat'}, \textit{mbord} \text{ 'table'} \} \end{array} \right]$$

Figure 13: Cleft sentence with focus PP: i-structure

$$\left[\begin{array}{l} \text{FOC} \{ \textit{mbord} \text{ 'table'} \} \\ \text{GROUND} \{ \textit{cat} \text{ 'cat'}, \textit{faoin} \text{ 'under'} \} \end{array} \right]$$

Figure 14: Cleft sentence with focus PP: i-structure

Here, syntax clearly reaches its limits, and prosodic phonology takes over. Clefting, as a *syntactic* focusing device, is only able to select the whole PP as the clefted constituent. It is not possible to have *just* the preposition or the preposition's object in this position. So if a speaker intends to focus *just* these, they have to employ contrastive focus, which is a prosodic device. Such prosodic cues are, however, not normally present in written text — which is why a grammar operating on written text can and should merely produce multiple i-structures in cases like (30). Note in this respect the work by Bögel et al. (2009), who have experimented with parsing prosodic information using a computational LFG grammar. Given the appropriate markup in the text, such a grammar could be used to disambiguate between structures such as Figures 12-14.

5 Conclusion

The paper has presented a syntactic analysis of Irish clefting. It started out by giving alternatives for analyzing copula constructions in LFG and showed why the PREDLINK analysis is a favorable one for Irish. The paper discussed data from Irish clefting and the similarities between simple copula predication and stage level predication using the substantive verb *bí* as well as clefting. It was argued that in fact one syntactic analysis can be used for both types of constructions. Section 4 extended the analysis to include Information Structure by encoding discourse functions in a separate projection of LFG, i-structure.

The similarities between run-of-the-mill copula sentences and clefts strongly suggest that a single analysis should be chosen, since the syntactic predication

is identical across constructions. With the approach presented in this paper, one does not have to rely on different syntactic analysis — the PREDLINK analysis models the parallelism. Another advantage of the proposed analysis is that it can be extended to other languages in a straightforward way, since several languages use copula predication in cleft sentences.

References

- Andréasson, Maia. 2007. The Architecture of I-Structure. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG07 Conference*, CSLI Publications.
- Attia, Mohammed. 2008. A Unified Analysis of Copula Constructions in LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG08 Conference*, CSLI Publications.
- Bögel, Tina, Butt, Miriam, Kaplan, Ronald M., King, Tracy Holloway and Maxwell III, John T. 2009. A treatment of clitics via prosodic phonology in LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG09 Conference*.
- Bresnan, Joan. 1982. Control and Complementation. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, MIT Press.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Blackwell Publishing.
- Bresnan, Joan and Mchombo, Sam A. 1987. Topic, pronoun, and agreement in Chicheŵa. In William Bright (ed.), *Language*, volume 63, Waverly Press.
- Butt, Miriam and King, Tracy Holloway. 1998. Interfacing Phonology with LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG98 Conference*, Stanford: CSLI Publications.
- Butt, Miriam, King, Tracy Holloway, Niño, María-Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. CSLI Publications.
- Carnie, Andrew. 1995. *Non-verbal predication and head movement*. Ph.D.thesis, MIT.
- Crouch, Dick, Dalrymple, Mary, Kaplan, Ronald M., King, Tracy Holloway, Maxwell III, John T. and Newman, Paula. 2008. *XLE Documentation*. Palo Alto Research Center.
- Dalrymple, Mary, Dyvik, Helge and King, Tracy Holloway. 2004. Copular Complements: Closed or Open? In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG04 Conference*.

- Declerck, Renaat. 1988. *Studies on Copular Sentences, Clefts and Pseudo-Clefts*. Leuven University Press.
- Halvorsen, Per-Kristian Giaever. 1977. *Syntax and Semantics of Cleft Constructions*. Ph. D.thesis, University of Texas at Austin.
- King, Tracy Holloway. 1997. Focus Domains and Information-Structure. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG97 Conference*, Stanford: CSLI Publications.
- King, Tracy Holloway and Zaenen, Annie. 2004. F-structures, Information Structure, and Discourse Structure. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG04 Conference*.
- Lodrup, Helge. 2008. Functional Structure. In Robert D. Borsley and Kersti Börjars (eds.), *Nontransformational Syntax: A Guide to Current Models*, Blackwell Publishing, to appear.
- O'Connor, Robert. 2004. *Information Structure in Lexical-Functional Grammar: The Discourse-Prosody Correspondence in English and Serbo-Croatian*. Ph. D.thesis, University of Manchester.
- Ó Siadhail, Mícheál. 1989. *Modern Irish - Grammatical Structure and Dialectal Variation*. Cambridge University Press.
- Ramchand, Gillian Catriona. 1997. *Aspect and Predication*. Oxford University Press.
- Rosén, Victoria. 1996. The LFG Architecture and “Verbless” Syntactic Constructions. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG96 Conference*.
- Sornicola, Rosanna. 1996. Topic, Focus, and Word Order. In Keith Brown and Jim Miller (eds.), *Concise Encyclopedia of Syntactic Theories*, Elsevier.
- Stenson, Nancy. 1981. *Studies in Irish syntax*. Gunter Narr Verlag.
- The Christian Brothers. 1960. *Graiméar Gaeilge na mBráithre Críostaí (The Christian Brothers' Irish Grammar)*. M. H. Mac an Ghoill agus a mhac Tta.
- Vallduví, Enric. 1993a. Information packaging: A survey. Technical Report, Centre for Cognitive Science & Human Communication Research Centre.
- Vallduví, Enric. 1993b. *The Informational Component*. Ph. D.thesis, University of Pennsylvania at Philadelphia.