

# An LFG Account of Discontinuous Nominal Expressions

Liselotte Snijders

Waseda University

Tokyo, Japan

liselottesnijders@gmail.com

## Abstract

This paper presents an overview of an LFG treatment of discontinuous nominal expressions involving modification, making the claim that cross-linguistically different types of discontinuity (i.e. in Warlpiri and English) should be captured by the same overall analysis, despite being licensed in different ways. LFG's separation of grammatical functions from phrase structural positions intuitively accounts for discontinuous expressions, and its use of glue semantics ensures that discontinuous and contiguous expressions receive the same semantic analysis.

## 1 Introduction

Discontinuity of nominal expressions, a phenomenon in which two or more parts of a semantic nominal unit are non-adjacent in phrase structure, is prevalent in languages traditionally classified as “non-configurational” (Hale, 1983), e.g. the Australian languages Warlpiri, Wambaya, Jaminjung (Simpson, 1991; Nordlinger, 1998; Schultze-Berndt and Simard, 2012), Latin (Devine and Stephens, 2000; Spevak, 2010), Ancient Greek (Devine and Stephens, 2006), and are also attested in a number of Slavic languages, e.g. Russian (Sekerina, 1997; Sekerina, 1999) and Polish (Siewierska, 1984). An example of nominal discontinuity from Warlpiri is shown in (1) (Simpson, 1991, p. 282):<sup>1</sup>

<sup>1</sup>This type of Warlpiri example has another interpretation, which can be translated as ‘The child<sub>i</sub> is chasing it and it<sub>i</sub> is small’ (Simpson, 1991). Based on Simpson’s work this appears to be secondary predication rather than discontinuity, therefore I only take the interpretation in (1) into account.

- (1) *Kurdu-ngku ka wajilipi-nyi*  
child-ERG PRES chase-NONPAST  
*wita-ngku*  
small-ERG  
‘The small child is chasing it.’

In (1) a head noun is separated from a modifier, but both parts map to the same grammatical function (subject). The two parts of the discontinuous expression share the same case-marking. A similar type of discontinuity involving modification is attested in English, in the cases of relative clause extraposition in (2a) and NP-PP split in (2b) (Kirkwood, 1977, p. 55):<sup>2</sup>

- (2) a. **The man** entered **who I met yesterday**.  
b. **A number of stories** soon appeared  
**about Watergate**.

A similar type of discontinuity is in fact also attested in Warlpiri (Hale, 1976, p. 78):<sup>3</sup>

- (3) *Ngajulu-rlu rna yankirri pantu-rnu*  
I-ERG AUX emu.ABS spear-PAST  
*kuja-lpa ngapa nga-rnu.*  
COMP-AUX water.ABS drink-PAST  
‘I speared the emu that was drinking water.’

<sup>2</sup>Another type of discontinuity in English involving modification is partial fronting, e.g. *About Japan, the woman wrote many books*; additional examples are discussed in Section 6.

<sup>3</sup>Hale (1976) refers to this type of example as ‘adjoined relative clause’: it can also precede the sentence as a whole (something like a hanging topic). It can also have a temporal reading: ‘I speared the emu while it was drinking water’.

Discontinuity of nominal expressions, whether of the kind in which two words are marked with the same case (as in (1)), or of the kind in which a modifier of an argument is postposed to follow the clause (as in (2) and (3)) presents a challenge for syntactic theory, as it requires the two or more parts of syntactic information to be united in the semantics.

In this paper I illustrate how discontinuous nominal expressions can be accounted for within the Lexical-Functional Grammar (LFG) framework based on previous work. With its constraint-based nature and parallel architecture, LFG provides a straightforward way of handling discontinuity by allowing two or more separate parts of phrase structural information to map to the same functional structure, and thereby to the same semantic structure. The focus of this paper is on nominal discontinuity involving modification specifically (i.e. a head and a modifier being separated), to limit the scope of discussion, but Section 8 briefly addresses a different type of discontinuity in comparison. I make the claim that discontinuous nominal expressions (involving modification) in typologically different languages (i.e. in Warlpiri and English) are instances of the same phenomenon and therefore require the same analysis, despite being licensed by somewhat different phrase structure rules. I propose a definition that captures both types of discontinuity, and illustrate that LFG is capable of accounting for the different types in a straightforward fashion.

Overall I aim to illustrate LFG’s potential in contributing to a potential implementation of discontinuity in NLP systems, because of its straightforward account of discontinuity, not requiring any special mechanisms. Discontinuous data is more challenging for approaches which parse sentences based on linear ordering, such as dependency grammar and approaches relying on surface phrase structure configuration. LFG is computationally implementable and has been implemented in the XLE system (Crouch et al., 2011). This paper thereby contributes to potential enhancements of NLP tasks with regards to discontinuity.

## 2 Lexical-Functional Grammar

Lexical-Functional Grammar (LFG) (Kaplan and Bresnan, 1982) is a constraint-based syntactic

framework which posits a parallel architecture, separating information about grammatical functions from phrase structural configuration. For this reason it is well-suited to account for languages with relatively free ordering of grammatical functions (e.g. Warlpiri). LFG posits two syntactic levels: constituent structure (c-structure) and functional structure (f-structure). C-structure encodes information about linear precedence, dominance relations and constituency, and is represented as a phrase structure tree. F-structure hosts information about the grammatical functions of the predicate of a sentence (including adjuncts), along with a range of morphosyntactic information such as case, number, tense and aspect. It is represented as an attribute-value matrix. C-structure nodes are locally annotated with information about grammatical functions and/or lexical information. Each c-structure node is associated with a particular f-structure, and a local annotation on the c-structure node ensures the mapping of the c-structure node to this f-structure via the  $\phi$  function. Specifically, the local annotation on a particular c-structure node specifies the relation of the f-structure associated with this node to the f-structure associated with its mother node. An example of this mapping for the simple sentence *John walked* is shown in Figure 1.

In the annotations on c-structure,  $\uparrow$  points to the f-structure of the mother node and  $\downarrow$  points to the f-structure of the current node. The annotation  $\uparrow = \downarrow$  thus expresses that the f-structure of the mother node is the same as the f-structure of its daughter node, and the annotation  $(\uparrow \text{SUBJ}) = \downarrow$  expresses that the daughter node maps to the subject of its mother’s f-structure. In English, this subject annotation is structurally associated with the NP node that is in the specifier of IP position (see Section 5). The grammatical functions SUBJ and OBJ (as well as a number of other functions assumed in LFG, such as  $\text{OBJ}_\theta$  and  $\text{OBL}_\theta$ ) are unique, e.g. a verb can only have one subject which it subcategorizes for. This subcategorization is marked in the lexical annotation on *walked* in Figure 1,  $(\uparrow \text{PRED}) = \text{‘walk<SUBJ>’}$ , which states that the predicate (PRED) value of this word is ‘walk’ and takes one argument, SUBJ. Adjuncts (ADJ) are not unique, but map to a set, by means of the annotation  $\downarrow \in (\uparrow \text{ADJ})$ , to be discussed in more detail in the next section.

### 3 Discontinuity in LFG

#### 3.1 Previous Work

Discontinuous nominal expressions have been analyzed in LFG by Simpson (1991) (Warlpiri), Kuhn (1999; 2001) (German), Cavar and Seiss (2011) (New-Shtokavian) and are discussed by Snijders (2012; 2015).<sup>4</sup> Simpson (1991) and Kuhn (1999; 2001) share a similar overall analysis in assuming that two parts of a discontinuous nominal expression map to the same f-structure. Here I focus on Simpson’s (1991) analysis. Simpson’s analysis for example (1), adapted to fit with her more recently proposed c-structure for Warlpiri (Simpson, 2007), is shown in Figure 2 (leaving out lexical annotations).<sup>5</sup>

Crucially, the two parts of the discontinuous expressions both map to the f-structure of the subject of the predicate. The annotation  $\downarrow \in (\uparrow \text{ADJ})$  on the N node states that this node ( $\downarrow$ ) maps to the set of adjuncts of the NP node ( $\uparrow$ ). Note that the object, despite being absent in c-structure, is present in f-structure as the verb requires an object: its PRED value is ‘PRO’.

The overall f-structure for the contiguous example is the same, as shown by the annotations on the c-structure of the contiguous version of example (1), shown in Figure 3. Unlike in English, in the Warlpiri c-structures the subject annotation comes from the case-marking, not from the structural position.

The example with the discontinuous expression and the contiguous one thus have the same f-structure, which is mapped to the same semantic structure, as will be discussed in Section 7.

#### 3.2 English Extraposition

Previous work in LFG does not discuss discontinuous nominal expressions in English, nor does it provide an account of how discontinuity in differ-

<sup>4</sup>There is other work on long-distance dependencies in LFG (e.g. Kaplan and Zaenen (1989), Clément et al. (2002) among others), but this literature discusses cases of arguments appearing outside the clause that they are part of (e.g. *wh*-fronting out of embedded clauses). This is a different type of long-distance dependency than the discontinuous nominal expressions discussed here, as the latter case involves two parts of the same grammatical function being separated.

<sup>5</sup>It is generally assumed that adjective-like elements in Warlpiri are of the N category (Hale, 1983; Simpson, 1991; Hale et al., 1995).

ent languages is licensed in different ways. One type of English discontinuity, extraposition, will be addressed here, and I show that it requires the same treatment as the discontinuity with two case-marked nominals as found in Warlpiri, providing a cross-linguistic definition of discontinuous nominal expressions in Section 4. I show that any cross-linguistic differences are due to differences in c-structure rules (phrase structure rules).

Discontinuity involving extraposition and discontinuity involving two words with the same case-marking have in common the fact that two parts of the same grammatical functions are not adjacent in phrase structure. In the instance of relative clause extraposition in English, I propose to represent the extraposed clause by means of an adjoined CP clause.<sup>6</sup> The structure I propose for (2a) is shown in Figure 4 with partial annotations. In the case of (2b) we would instead have an adjoined PP.

I note a few differences between the type of discontinuity with two case-marked nominals (Figure 2) and the English extraposition one (Figure 4). Linguistically, a difference is the type of categories that may be separated from each other, i.e. in Warlpiri two nominals, while in English an NP and a CP or PP. A second point of linguistic variation is the position that the two or more parts of the discontinuous expression may appear in. In English this is restricted, shown by the unacceptability of *\*A number of stories soon about Watergate appeared*. In the case of Warlpiri discontinuity in which the subparts have the same case-marking, the placement of the subparts is much freer, reflecting Warlpiri’s property of free placement of grammatical functions.<sup>7</sup>

Another difference between the discontinuous expression in Warlpiri in Figure 2 and the English one in Figure 4 is the type of annotation on the modifier:

<sup>6</sup>The CP does not form a constituent with the VP, as preposing of both is ruled out: *\*Entered who I saw yesterday, the man*. For this reason adjunction to IP is appropriate.

<sup>7</sup>However, like word order, discontinuity is not random, but is triggered by information structure, as discussed by De Jong (1986) (Latin), Cavar and Seiss (2011) (New-Shtokavian) and Schultze-Berndt and Simard (2012) (Jaminjung). A full discussion of the information structure of discontinuous expressions is beyond the scope of this paper. Also, recall that Warlpiri does appear to have a type of extraposition as shown in (3), which seems more restricted in its placement than the type of discontinuity involving case-marked nominals.

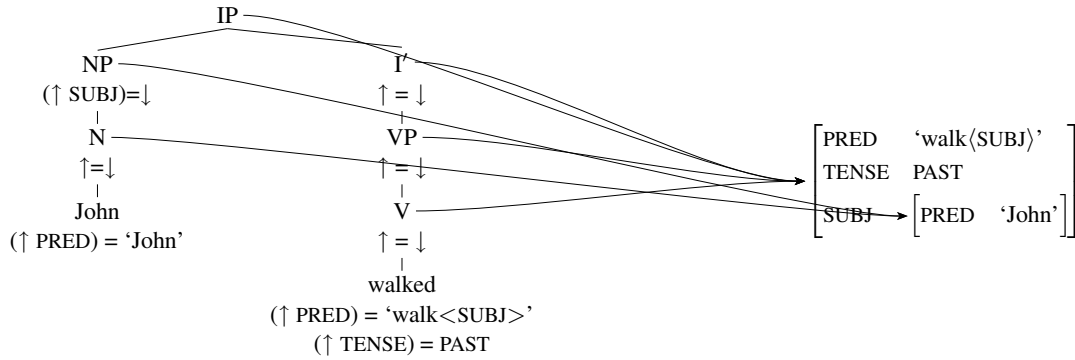


Figure 1: An illustration of c- to f-structure mapping.

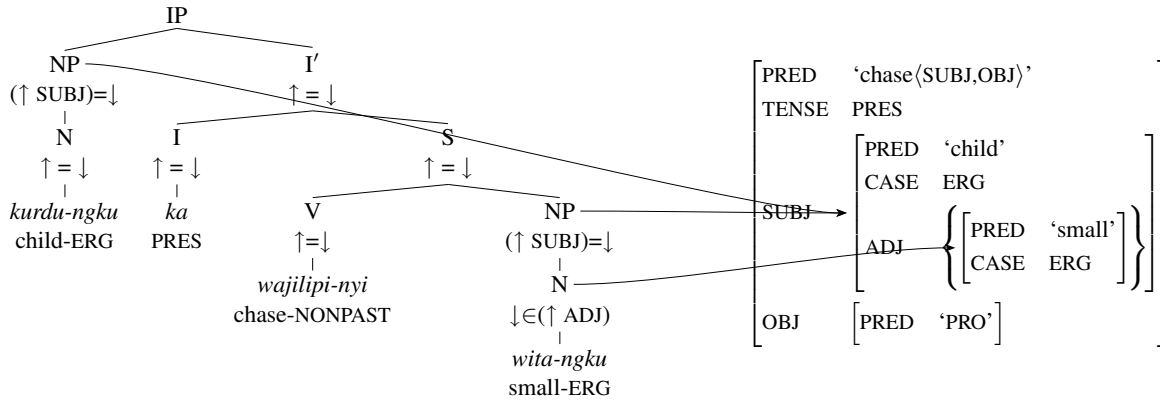


Figure 2: C- to f-structure mapping of a discontinuous expression.

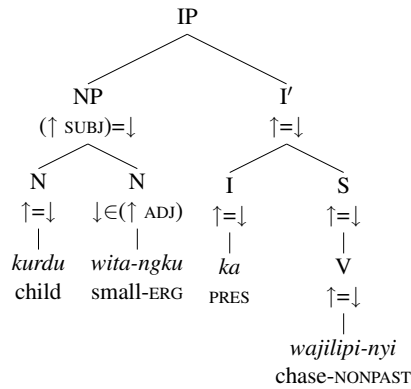


Figure 3: C-structure for a contiguous example, mapping to the f-structure in Figure 2.

in the Warlpiri case both head and modifier map to the same f-structure directly (subject of the predicate by means of the annotation  $(\uparrow \text{SUBJ}) = \downarrow$ ). In the English case, the modifier maps to the f-structure of the adjunct of the SUBJ, whereas the head maps to the overall f-structure of the SUBJ. The annotations are thus somewhat different, but the end result for both examples is the same: both head and modifier are contained within the f-structure of the subject. This

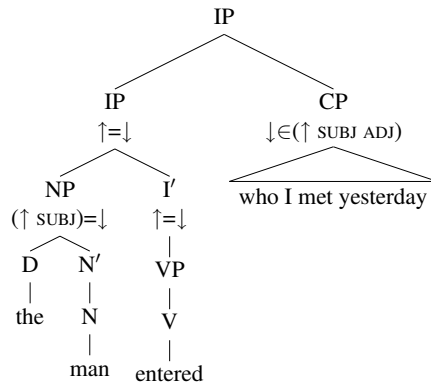
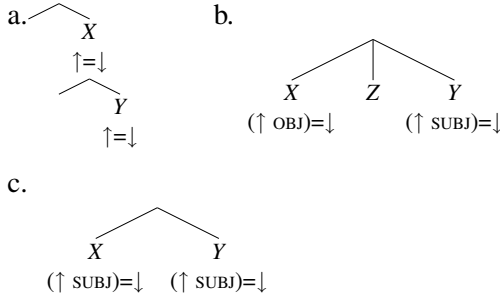


Figure 4: C-structure for English relative clause extraposition.

is an important observation, as we would have the same f-structure mapping even with different annotations. For example, in the Warlpiri c-structure in Figure 2, the annotations on the two NPs are both  $(\uparrow \text{SUBJ}) = \downarrow$ , both mapping to the same f-structure, but we could also imagine a set of annotations where the adjunct has the annotation  $\downarrow \in (\uparrow \text{SUBJ ADJ})$ . A definition of discontinuity needs to abstract away from this variation in annotation, which is reflected in the definition proposed in the following section.



**Figure 5:** Structures not fulfilling the conditions of definition (4).

#### 4 Definition of Discontinuity in LFG

In order to capture nominal discontinuity in a more formal way, I propose the following definition:<sup>8</sup>

(4) **Nominal discontinuous expressions:**

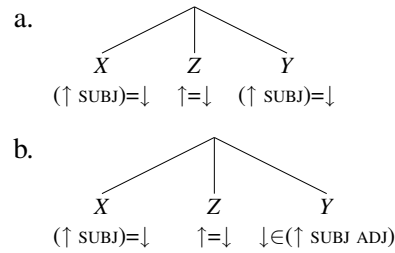
Given two c-structure constituents  $X$  and  $Y$ ,  $X \neq Y$ ,  $\{X, Y\}$  form a discontinuous nominal expression iff:

- i. Neither  $X$  nor  $Y$  dominate the other; and
- ii.  $X$  and  $Y$  map to the f-structure or sub-f-structure of the same grammatical function; and
- iii. The yield of  $X$  is not string adjacent to the yield of  $Y$ ; and
- iv. The constituent(s) that intervene(s) between  $X$  and  $Y$  do(es) not map to the f-structure or any sub-f-structure of the grammatical function that  $X$  and  $Y$  map to.

The key here is that the two parts of a discontinuous expression both map to the f-structure of the same grammatical function, or to an f-structure that is contained within the f-structure of this grammatical function. Consider the partial structures shown in Figure 5, all of which do not fulfill all conditions of definition (4).

Condition (i) rules out structure (a) in Figure 5 as being discontinuous;  $X$  and  $Y$  here map to the same f-structure, but  $X$  dominates  $Y$ . Condition (ii) rules

<sup>8</sup>One reviewer points out that these conditions (especially (i), (iii) and (iv)) do not confine to nominals. However, this paper restricts itself to nominal expressions (involving modification); extending the definition to other kinds of discontinuity is briefly addressed in Section 8.



**Figure 6:** Instances of discontinuous nominal expressions.

out the structure in Figure 5 (b) as an instance of discontinuity. In this structure the yield of  $X$  is not string adjacent to the yield of  $Y$  (in other words, the edges of  $X$  and  $Y$  do not coincide), but  $X$  and  $Y$  do not map to the f-structure (or sub-f-structure) of the same grammatical function. Condition (iii) ensures that there is an intervening element, and rules out the structure in Figure 5 (c) as an instance of discontinuity.

The structure in Figure 6 (a) fulfills all conditions as listed in (4), including condition (iv). Finally, according to condition (ii) in (4), the structure in Figure 6 (b) is also a case of discontinuity.

#### 5 Constraining Discontinuity

The definition of nominal discontinuity in (4) covers both the type of nominal discontinuity attested with two case-marked nominals (one of which modifies the other) and the type with an extraposed modifier clause. Their analysis is very similar in terms of their c- to f-structure mapping, but the way in which the two different types are licensed is somewhat different. The case-marked nominal type of discontinuity is made possible by the assumption in LFG that in languages like Warlpiri grammatical functions are assigned lexically and not by phrase structure configuration (Dalrymple, 2001; Bresnan et al., 2016). Free assignment of grammatical functions enables the existence of discontinuous expressions as in (1).<sup>9</sup> Cases of English extraposition are more constrained as grammatical functions are generally assumed to

<sup>9</sup>Again, this does not mean that discontinuity is unconstrained: it appears constrained by information structure. Moreover, there are languages which like Warlpiri have free assignment of grammatical functions, but which lack discontinuous nominal expressions in which two words with the same case-marking are separated. We can say that Warlpiri has *head optionality*, allowing for a modifier to appear without its head nominal.

be assigned configurationally.

In LFG c-structure is licensed and constrained by c-structure rules (phrase structure rules), assumed to be static constraints on c-structure. In English, a rule ensuring that the subject appears pre-verbally and VP-externally is as follows:<sup>10</sup>

$$(5) \quad \text{IP} \rightarrow \begin{array}{cc} \text{NP} & \text{I}' \\ (\uparrow \text{SUBJ}) = \downarrow & \uparrow = \downarrow \end{array}$$

The annotation on the NP ensures that if the NP is present, it obligatorily hosts the subject.<sup>11</sup> The rule in (5) partly licenses the c-structure in Figure 1. In Warlpiri, with no obligatory annotations for any grammatical function, the c-structure rules are less constrained in this dimension. An example is the IP rule partly licensing the c-structures in Figures 2 and 3, where GF = ‘grammatical function’:<sup>12</sup>

$$(6) \quad \text{GF} \equiv \{ \text{SUBJ} \mid \text{OBJ} \} \\ \text{IP} \rightarrow \left\{ \begin{array}{cc} \text{NP} & \mid \text{V} \\ (\uparrow \text{GF}) = \downarrow & \uparrow = \downarrow \end{array} \right\} \text{I}' \\ \uparrow = \downarrow \quad \uparrow = \downarrow \quad \uparrow = \downarrow$$

The node preceding I' ranges over NP and V, as either an NP or a V may appear preceding the AUX constituent (which is in I position), as long as it bears a focus function. The annotations on NP and V are different. Relevant for the current discussion is the annotation on the NP: it is unspecified for its grammatical function (in this case, it ranges over SUBJ and OBJ, but this set can be extended depending on the attested data). Assuming that all NPs in Warlpiri's c-structure rules have the same unspecified annotation ( $\uparrow \text{GF} = \downarrow$ ), it is possible for an annotation for the

<sup>10</sup>The subject can also appear in Spec,CP position, for example when it is a *wh*-word.

<sup>11</sup>I note ‘if the NP is present’, as LFG adheres to the principle of *Economy of Expression*, which states that all phrase structure nodes are optional, unless required by independent principles (Bresnan et al., 2016). An example of an independent principle is satisfaction of subcategorization requirements, e.g. if the subject is expressed elsewhere in the c-structure (e.g. in Spec,CP) then the NP node in (5) is absent.

<sup>12</sup>The reason for assuming an IP in Warlpiri (following Austin and Bresnan (1996), Simpson (2007)) is the set positions of two types of constituents. The first is the verb-like constituent referred to as AUX (‘auxiliary’) in the Warlpiri literature, like *ka* (glossed ‘PRES’) in (1), assumed to appear in I position. The second is the constituent immediately preceding the AUX, which Simpson (2007) assumes to always have a focus discourse function (similarly, she assumes that Spec,CP always hosts a topic function).

same GF to appear on two NPs, enabling discontinuity (under the assumption that in Warlpiri syntactic heads are optional). The actual annotation on the NPs is determined lexically, by case-marking.

In English the situation is somewhat different. Discontinuity of the kind shown in (2) and Figure 4 is licensed by the rule in (5) and a rule like the one in (7) for adjunction of CP or PP to IP:

$$(7) \quad \text{XP} \equiv \{ \text{CP} \mid \text{PP} \} \\ \text{IP} \rightarrow \begin{array}{cc} \text{IP} & \text{XP} \\ \uparrow = \downarrow & \downarrow \in (\uparrow \text{GF ADJ}) \end{array}$$

The annotation on the XP,  $\downarrow \in (\uparrow \text{GF ADJ})$ , maps to the adjunct set of a GF function. Note that this GF is not restricted to SUBJ, as we can have examples like *Mary mentioned the claim yesterday that John is intelligent*. Despite configurational assignment of GFs in English, there is some nonspecificity in annotation here. It appears that Warlpiri adjoined relative clauses (as in (3)) can be licensed by a similar rule with a similar annotation on an extraposed CP (but with the difference that the clause can appear on either side of the main IP). A more in-depth investigation of the data is needed to posit a specific rule like this, but a generalized rule of extraposition seems plausible and would be promising for a uniform approach to this type of discontinuity. Discontinuity involving case-marked nominals and discontinuity involving extraposition have somewhat different underlying mechanisms. We can say that different restrictions on annotations on c-structure rules in the different constructions (and languages) lead to very similar outcomes, namely that two parts of c-structure which are non-adjacent can map to the f-structure or sub-f-structure of the same grammatical function.

## 6 More complex cases

There are other types of more complex cases of nominal discontinuity involving modification in English, namely extraposition with embedding (in (8a)) (Müller, 2016, p. 443), extraction out of complex NPs (in (8b)) and secondary predication (in (8c)):<sup>13</sup>

- (8) a. **Many proofs of the theorem appeared that I wanted to think about.**

<sup>13</sup>I thank the anonymous reviewer for suggesting to include these examples in the paper.

- b. **Who** did they take **pictures of**?
- c. She watched **him naked**.

Extrapolation with embedding ((8a)) is a more complex version of example (2). The extraposed clause can be assumed to either modify the head *many proofs* or the modifier *of the theorem*. In the first case, rule (7) applies, with  $GF = \text{SUBJ}$ . For the latter case we need an extension of rule (7) to ensure that the relative clause can map to the adjunct set of the adjunct of the head, which can be achieved by an additional possible annotation on the XP in (7) of the form  $\downarrow \in (\text{ADJ} \in \text{ADJ GF} \uparrow)$ .

As for the type of discontinuity involving extraction from complex NPs ((8b)), this is captured by the definition of discontinuity in (4) if we assume that the first part (*who*) maps to the adjunct set of the object, and that *pictures of* maps to the object. The preposition *of* can by itself map to the adjunct set of the object as well, meaning that both *who* and *of* are part of the adjunct set.<sup>14</sup> Therefore this type of example does not contradict the generalizations proposed.

The example of secondary predication in (8c) is not covered by the definition of nominal discontinuity in (4), as there is no intervening material between *him* and *naked*. This construction appears very similar to relative clause extrapolation in the sense that a modifier follows the sentence as a whole. Under the current approach this is not assumed to be a case of discontinuity, even if the two adjacent words *him* and *naked* do not form a syntactic constituent. I leave this issue open for discussion.

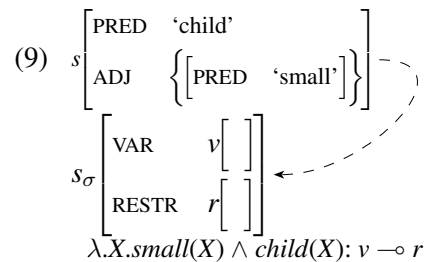
An approach to discontinuity following the definition in (4) thus covers most cases, but in an implementation of this approach, one might need to consider specific constructions individually to achieve accurate results.

## 7 Mapping to Semantics

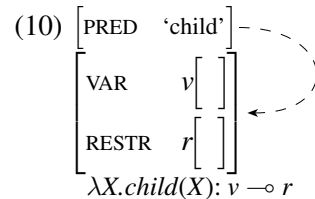
For completeness, I discuss how discontinuous nominal expressions involving modification can be analyzed semantically. The c- to f-structure mapping in LFG via the  $\phi$  function ensures that a minimal

<sup>14</sup>However, this mapping does not ensure that *who* and *of* end up as part of the same f-structure in this set; this issue will be addressed in Section 8.

pair of sentences with or without a discontinuous expression have the same f-structure, as shown above. Semantics in LFG is represented on the level of s-structure. Following Dalrymple and Nikolaeva (2011, p. 90), I assume that f-structure is mapped directly to s-structure via the  $\sigma$  function. The direct mapping ensures that sentences with the same f-structure will receive the same semantics. A discontinuous expression and a contiguous expression will therefore have the same semantics, as also pointed out by Dalrymple (2001). This is achieved by glue semantics (Dalrymple et al., 1993; Dalrymple, 1999; Dalrymple, 2001), the linguistic theory of semantic composition commonly used in LFG, which relies on linear logic. Glue semantics associates meaning constructors, instructions on how to combine meanings to form the meaning of the sentences, with lexical items (or in some cases with phrase structural positions). Semantic composition is therefore largely separate from c-structure constituency, which is especially beneficial for the purpose of accounting for discontinuous nominal expressions, as we want the same semantic analysis for two different c-structural configurations. For example, in the Warlpiri example in Figure 2, the two subparts of the discontinuous expression each contribute their own meaning constructors. Before looking at these, consider the f- to s-structure mapping for the SUBJ (the discontinuous expression) of example (1):



The f- to s-structure mapping of the head by itself is very similar:



In (9), the subject's f-structure is labelled  $s$ , and its s-structure is labelled  $s_{\sigma}$ . The s-structure  $s_{\sigma}$  has two

*kurdu*  
 $(\uparrow \text{ PRED}) = \text{'child'}$   
 $\lambda X.\text{child}(X) : (\uparrow_{\sigma} \text{ VAR}) \multimap (\uparrow_{\sigma} \text{ RESTR})$

*wita*  
 $(\uparrow \text{ PRED}) = \text{'small'}$   
 $\lambda P.\lambda X.\text{small}(X) \wedge P(X)$   
 $[(\text{ADJ} \in \uparrow)_{\sigma} \text{ VAR}] \multimap ((\text{ADJ} \in \uparrow)_{\sigma} \text{ RESTR})] \multimap$   
 $[(\text{ADJ} \in \uparrow)_{\sigma} \text{ VAR}] \multimap ((\text{ADJ} \in \uparrow)_{\sigma} \text{ RESTR})]$

**Figure 7:** Lexical entries for *kurdu* and *wita*.

attributes: VAR which has as its value an s-structure labelled  $v$  and RESTR which has as its value an s-structure labelled  $r$ . The attribute VAR represents a variable of type  $e$  and RESTR is of type  $t$  and represents a restriction the variable of type  $e$ . For (9) the restriction would state that the variable must range over individuals that are both children and small. For (10) the only restriction is that the variable must range over individuals that are children. The notation  $\multimap$ , the *linear implication* symbol of linear logic, signifies that if there is an attribute VAR ( $v$ ) in the s-structure  $(\uparrow_{\sigma})$  then there is also an attribute RESTR ( $r$ ) in that same s-structure.

The s-structure in (9) of the subject of example (1) comes from the lexical entries of the two words of the discontinuous expression, with the one entry restricting the other. These lexical entries are shown in Figure 7, leaving out case marking. The lexical entry of the head, *kurdu* ('child'), states that *kurdu* provides a value for the PRED attribute in the f-structure, namely *child*. The second part of the lexical entry of *kurdu* makes a statement about the mapping to s-structure (signaled by the use of  $\uparrow_{\sigma}$  mapping to s-structure). It states the restriction on the variable VAR: it must range over individuals that are children. The lexical specification of the modifier *wita* is somewhat more complicated. Here  $(\text{ADJ} \in \uparrow)$  refers to the f-structure of which  $\uparrow$  (the adjunct) is a member (the set of adjuncts, or modifiers),  $(\text{ADJ} \in \uparrow)_{\sigma}$  refers to the s-structure corresponding to that f-structure and  $((\text{ADJ} \in \uparrow)_{\sigma} \text{ VAR})$  refers to the value of the VAR attribute of that s-structure. Likewise,  $((\text{ADJ} \in \uparrow)_{\sigma} \text{ RESTR})$  refers to the value of the RESTR attribute of the s-structure  $(\text{ADJ} \in \uparrow)_{\sigma}$ . Referring to these s-structures with the labels  $v$  and  $r$ , as shown in the examples in (9) and (10), the meaning constructor premises for the two individual words of the discontinuous expression in example (1) are

**[child]**  $\lambda X.\text{child}(X) : v \multimap r$   
**[small]**  $\lambda P.\lambda X.\text{small}(X) \wedge P(X) : [v \multimap r] \multimap [v \multimap r]$

**Figure 8:** Meaning constructors for *kurdu* and *wita*.

$$\frac{\lambda X.\text{child}(X) : v \multimap r \quad \lambda P.\lambda X.\text{small}(X) \wedge P(X) : [v \multimap r] \multimap [v \multimap r]}{\lambda X.\text{small}(X) \wedge \text{child}(X) : v \multimap r}$$

**Figure 9:** Deduction of the meaning of the discontinuous expression.

shown in Figure 8, with the meaning constructors in bold and brackets (with the glue semantics side on the right).

From the meaning constructors for the two individual words, one can deduce the meaning of the overall expression as shown in Figure 9. The meaning constructor for the modifier consumes the contribution of the noun ( $v \multimap r$ ), and thereby provides a new meaning, also associated with  $v \multimap r$ . Without providing a full overview of glue semantics and its use in LFG, this section has shown that by associating meaning constructors directly with lexical items and not with phrase structural positions, one can have the same semantic derivation for both contiguous and discontinuous nominal expressions.

## 8 Remaining Issues

There are a few remaining issues left to be resolved with regards to this approach to discontinuity. First, one outstanding technical issue that was brought up by Snijders (2012) is the problematic account of discontinuous adjuncts. Discontinuous adjuncts are found for example in Latin, (Bolkestein, 2001, p. 255). In c-structure both parts of a discontinuous adjunct will be marked with the annotation  $\downarrow \in (\uparrow \text{ ADJ})$ , to ensure that both parts map to the adjunct set of the predicate. However, unlike SUBJ or OBJ, ADJ is not a unique grammatical function. This is apparent from the set notation. Any NP annotated with the adjunct annotation will map to the adjunct set, and in principle each NP will form its own f-structure in this set. There is no clear way to distinguish between the case in which two nominals (with the same case, number, gender, assuming that this a constraining factor for discontinuity) form two separate f-structure adjuncts (separate elements of the adjunct set) or are part of the same adjunct f-structure. The only constraining factor in this is 'PRED clash': grammatical functions (including ADJ)



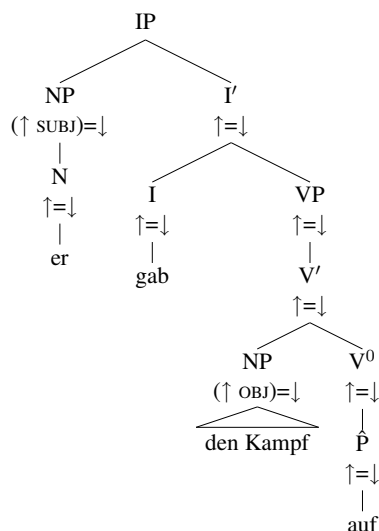


Figure 10: C-structure for example (12a).

may only have one PRED value:

$$(11) \left[ \begin{array}{l} \text{PRED} \quad \text{'VERB(SUBJ)'} \\ \text{SUBJ} \quad \left[ \text{PRED} \quad \text{'SUBJ'} \right] \\ \text{ADJ} \quad \left\{ \left[ \text{PRED} \quad \text{'ADJ1'} \right], \left[ \text{PRED} \quad \text{'ADJ2'} \right] \right\} \end{array} \right]$$

PRED clash ensures that two adjuncts which both contribute a PRED value do not unify in f-structure. Nonetheless, there is no straightforward way to distinguish between the two cases just described, and to ensure that two parts of a discontinuous adjunct map to the same f-structure (or sub-f-structure).

A second issue is that the definition in (4) covers discontinuous nominal expressions, but it does not encompass other types of discontinuity, such as the one created by ‘particle verbs’ (Ackerman, 1983; Piñón, 1992; Lüdeling, 2001; Booij, 2002; Toivonen, 2003; Forst et al., 2010).<sup>15</sup> Consider an example from German (Forst et al., 2010, p. 229):

- (12) *Er gab den Kampf auf.*  
 he gave the.ACC fight up  
 ‘He gave up the fight.’

I follow Toivonen’s (2003) analysis of verbal particles as non-projecting words, marked as  $\hat{P}$ . Example (12) then has the c-structure as shown in Figure

<sup>15</sup>Also, discontinuous nominal expressions not involving modification but rather with a separated determiner and a noun, as found for example in Latin (Devine and Stephens, 2006, p. 524), have not been discussed. The semantic mapping for this will be different than the mapping for modification described in Section 7.

10. In this c-structure the two parts of the phrasal verb both map to the same f-structure, namely the overarching f-structure of the sentence (shown by  $\uparrow = \downarrow$ ). Aiming for a general definition of discontinuity, this makes it seem appropriate to change condition (ii) in the definition of discontinuity in (4) to read ‘X and Y map to the same f-structure or its sub-f-structure’, with no mention of a grammatical function. However, this rephrasing makes inaccurate predictions, because if we refer to the highest f-structure level (the one of the whole sentence) and its sub-f-structures, we refer to all of the f-structures contained in the sentence.<sup>16</sup> Making reference to all f-structures in the definition of discontinuity makes it impossible to make any point about discontinuity. One solution is to change the phrasing of condition (ii) to ‘X and Y map to the same f-structure’, but under this condition the English structure in Figure 4 would not be an instance of a discontinuous expression, at least not with the annotations as shown. However, we want the definition to cover both extraposition and discontinuity like example (1), as ultimately both are somewhat different instances of the same mechanism.

## 9 Conclusion

This paper has illustrated an LFG approach to discontinuous nominal expressions involving modification, i.e. by letting two (or more) c-structural constituents map to the same (sub-)f-structure of a specific grammatical function. In this I follow Simpson (1991), Kuhn (1999; 2001), while making the explicit claim that different types of discontinuity (e.g. two constituents with the same case-marking or the case of extraposed XPs) should be captured by the same overall analysis, despite being licensed in different ways. Crucially, discontinuous expressions and contiguous expressions receive the same mapping to semantics, enabled by glue semantics’ association of meaning constructors with lexical items, not phrase structure. This paper has thereby aimed to illustrate LFG’s potential in contributing to a potential implementation of discontinuity in NLP systems.

<sup>16</sup>One reviewer suggests referring to immediate sub-f-structures only, but the immediate sub-f-structures of the sentence’s overarching f-structure include those of the arguments of the predicate, and thereby of the sentence as a whole.

## Acknowledgments

I gratefully acknowledge the JSPS for funding this work. I also thank Jelke Bloem, Mary Dalrymple, Ryo Otaguro, Marjolein Poortvliet and the anonymous reviewers for their useful suggestions.

## References

- Farrell Ackerman. 1983. *Miscreant Morphemes: Phrasal Predicates in Ugric*. Ph.D. thesis, UC Berkeley.
- Peter Austin and Joan Bresnan. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language and Linguistic Theory*, 14(2):215–268.
- A. Machteld Bolkestein. 2001. Random Scrambling? Constraints on Discontinuity in Latin Noun Phrases. In C. Moussy, editor, *De lingua Latina novae quaestiones*, pages 245–258. Peeters, Louvain.
- Geert Booij. 2002. Separable Complex Verbs in Dutch: A Case of Periphrastic Word Formation. In Nicole Dehé, Ray Jackendoff, Andrew McIntyre, and Silke Urban, editors, *Verb-Particle Explorations*, pages 21–42. Mouton de Gruyter, Berlin.
- Joan Bresnan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler. 2016. *Lexical-Functional Syntax*. Wiley Blackwell, Chichester, West Sussex.
- Damir Cavar and Melanie Seiss. 2011. Clitic Placement, Syntactic Discontinuity and Information Structure. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG11 Conference*, pages 131–151. CSLI Publications.
- Lionel Clément, Kim Gerdes, and Sylvain Kahane. 2002. An LFG-Type Grammar for German Based on the Typological Model. In M. Butt and T.H. King, editors, *Proceedings of the LFG02 Conference*, pages 116–129. CSLI Publications.
- Richard Crouch, Mary Dalrymple, Ronald M. Kaplan, Tracy King, III Maxwell, John T., and Paula Newman. 2011. XLE Documentation. Palo Alto Research Center (PARC), Palo Alto, CA.
- Mary Dalrymple and Irina Nikolaeva. 2011. *Objects and Information Structure*. Cambridge Studies in Linguistics. Cambridge University Press, Cambridge.
- Mary Dalrymple, John Lamping, and Vijay Saraswat. 1993. LFG semantics via constraints. In *Proceedings of the 6th Meeting of the EACL*, pages 97–105, Utrecht. European Association for Computational Linguistics.
- Mary Dalrymple, editor. 1999. *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. MIT Press, Cambridge, MA.
- Mary Dalrymple. 2001. *Lexical-Functional Grammar*, volume 34 of *Syntax and Semantics*. Academic Press, San Diego; London.
- Andrew M. Devine and Laurence D. Stephens. 2000. *Discontinuous syntax: hyperbaton in Greek*. Oxford University Press, Oxford/New York.
- Andrew M. Devine and Laurence D. Stephens. 2006. *Latin Word Order*. Oxford University Press, Oxford/New York.
- Martin Forst, Tracy Holloway King, and Tibor Laczkó. 2010. Particle Verbs in Computational LFGs: Issues from English, German and Hungarian. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG10 Conference*, pages 228–248. CSLI Publications.
- Kenneth L. Hale, Mary Laughren, and Jane Simpson. 1995. Warlpiri syntax. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, editors, *Syntax. Ein internationales Handbuch zeitgenössischer Forschung: An International Handbook of Contemporary Research*, pages 1430–51. Walter de Gruyter, Berlin New York.
- Kenneth L. Hale. 1976. The adjoined relative clause in Australia. In R.M.W. Dixon, editor, *Grammatical categories in Australian languages*, pages 78–105. Australian Institute of Aboriginal Studies, Canberra.
- Kenneth L. Hale. 1983. Warlpiri and the Grammar of Non-Configurational Languages. *Natural Language and Linguistic Theory*, 1.1:5–47.
- Jan R. De Jong. 1986. Hyperbaton en informatiestructuur. *Lampas*, pages 323–331.
- Ronald M. Kaplan and Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 173–281. MIT Press, Cambridge, MA.
- Ronald M. Kaplan and Annie Zaenen. 1989. Long-Distance Dependencies, Constituent Structure and Functional Uncertainty. In Mark Baltin and Anthony Kroch, editors, *Alternative Conceptions of Phrase Structure*, pages 17–42. University of Chicago Press, Chicago, IL. Reprinted in M. Dalrymple, R.M. Kaplan, J.T. Maxwell III and A. Zaenen, editors, *Formal Issues in Lexical-Functional Grammar*, pages 137–165, CSLI Publications, Stanford University, 1995.
- H.W. Kirkwood. 1977. Discontinuous Noun Phrases in Existential Sentences in English and German. *Journal of Linguistics*, 13(1):53–66.
- Jonas Kuhn. 1999. The syntax and semantics of split NPs in LFG. In Francis Corblin, Carmen Dobrovie-Sorin, and Jean-Marie Marandin, editors, *Selected papers from the Colloque de Syntaxe et Sémantique à Paris (CSSP 1997)*, pages 145–166. Thesus, The Hague.

- Jonas Kuhn. 2001. Resource Sensitivity in the Syntax-Semantics interface and the German Split NP Construction. In T. Kiss and D. Meurers, editors, *Constraint-Based Approaches to Germanic Syntax*. CSLI Publications, Stanford, CA.
- Anke Lüdeling. 2001. *On Particle Verbs and Similar Constructions in German*. CSLI Publications, Stanford, CA.
- Stefan Müller. 2016. *Grammatical Theory: From transformational grammar to constraint-based approaches*. Textbooks in Language Sciences 1. Language Science Press, Berlin.
- Rachel Nordlinger. 1998. *Constructive Case: Evidence from Australian Languages*. CSLI Publications, Stanford, CA. Revised version of 1997 Stanford University dissertation.
- Chris Piñón. 1992. The Preverb Problem in German and Hungarian. In *Proceedings of BLS*, pages 395–408.
- Eva Schultze-Berndt and Candide Simard. 2012. Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure. *Linguistics*, 50:1015–1058.
- Irina Sekerina. 1997. *The Syntax and Processing of Split Scrambling Constituents in Russian*. Ph.D. thesis, SUNY Graduate School and University Center.
- Irina Sekerina. 1999. The Scrambling Complexity Hypothesis and Processing of Split Scrambling Constituents in Russian. *Journal of Slavic Linguistics*, 7(2):265–304.
- Anna Siewierska. 1984. Phrasal Discontinuity in Polish. *Australian Journal of Linguistics*, 4:57–71.
- Jane Simpson. 1991. *Warlpiri Morpho-Syntax: a Lexicalist Approach*, volume 23 of *Studies in Natural Language and Linguistic Theory*. Kluwer Academic Publishers, Dordrecht.
- Jane Simpson. 2007. Expressing Pragmatic Constraints on Word Order. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling, and Chris Manning, editors, *Architectures, Rules, and Preferences: Variations on Themes by Joan W. Bresnan*, pages 403–427. CSLI Publications, Stanford, CA.
- Liselotte Snijders. 2012. Issues Concerning Constraints on Discontinuous NPs in Latin. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG12 Conference*, pages 565–581. CSLI Publications.
- Liselotte Snijders. 2015. *The Nature of Configurationality in LFG*. Ph.D. thesis, University of Oxford.
- Olga Spevak. 2010. *Constituent Order in Classical Latin Prose*. Studies in Language Companion Series. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Ida Toivonen. 2003. *Non-Projecting Words: A Case Study of Swedish Particles*. Studies in Natural Language and Linguistic Theory. Kluwer Academic Publishers, Dordrecht.