COMPLEX PREDICATES IN ARRERNTE

Mark Dras[1]   François Lareau[1]
Benjamin Börschinger[1,2]   Robert Dale[1]
Yasaman Motazedi[1]   Owen Rambow[3]
Myfany Turpin[4]   Morgan Ulinski[3]

[1]Macquarie University   [2]Universität Heidelberg
[3]Columbia University   [4]University of Queensland

Abstract

Using the example of Murrinh-Patha, Seiss (2011) illustrates how Australian Aboriginal languages can shed light on the morphology-syntax interface: one aspect of their polysynthetic nature is that information often encoded in phrases and clauses in other languages is instead found in a single morphological word. In this paper, we look at another instance, the Australian Aboriginal language Arrernte, and in particular at complex predicates within the language, to examine the implications for the morphology-syntax interface. Following from this, we show how a glue semantics-based approach can be applied to Arrernte complex predicates, in a way that fits neatly with the use of glue semantics to model lexical functions in LFG in a multilingual natural language generation environment.

1   Introduction

Using the example of Murrinh-Patha, Seiss (2011) illustrates how Australian Aboriginal languages can shed light on the morphology-syntax interface: one aspect of their polysynthetic nature is that information often encoded in phrases and clauses in other languages is instead found in a single morphological word, one manifestation of the morphology-competes-with-syntax idea discussed in Bresnan (2001). In this paper, we look at another instance, the Australian Aboriginal language Arrernte, and in particular at complex predicates within the language, to examine the implications for the morphology-syntax interface.

The context for this work is a data-to-text multilingual natural language generation (MNLG) system, where one of the languages to be generated is Eastern/Central Arrernte. For the language realisation component, the grammar is developed in XLE and the morphology in XFST. Some aspects of the language can be handled quite straightforwardly using standard LFG mechanisms. Complex predicates, however, require more consideration.

Complex predicates in LFG have most often been handled using some kind of predicate composition, sometimes at the level of argument structure elaborated via Jackendoff's Lexical Conceptual Structures (LCS) (Jackendoff, 1990), and in some cases through the use of the restriction operation (Butt, 1993; Alsina, 1997; Andrews and Manning, 1999; Butt and King, 2006; Nordlinger, 2010, exemplify some approaches). An alternative proposed in an appendix of Andrews and Manning (1999) and elaborated in Andrews (2007) is to use glue semantics, which is outlined there with a sketch of Romance causatives.

Following from our consideration of the morphology-syntax interface, we show how a glue semantics-based approach can be applied to Arrernte complex predicates, in a way that fits neatly with the use of glue semantics to model lexical collocations in LFG in an MNLG environment (Lareau et al., 2011).

2 Arrernte: The Language

2.1 Structure of the Grammar

Eastern/Central Arrernte is a language of the Arandic group of the Pama-Nyungan family of Australian languages. It is one of the larger Australian languages, with perhaps 1500 to 2000 speakers, who mostly reside around Alice Springs in Australia's Northern Territory. It is also a language of regular communication, which children still speak as a first language. In addition to the early work of Strehlow (1944), there are two major written descriptions of aspects of Arrernte grammar, PhD theses by Wilkins (1989) and Henderson (1998); in addition, there is a large dictionary (Henderson and Dobson, 1994).

Henderson (1998, Sec 1.4) gives an overview of the structure of the language:

> E/C Arrernte is suffixing and agglutinative and is rich in compounding morphology of various types. The principal parts of speech are nominals, verbs, adverbs and particles/clitics. ... The core syntactic cases are Ergative (ERG), Nominative (NOM), Accusative (ACC) and Dative (DAT). In nominals other than pronouns Nominative and Accusative are syncretised, while in pronouns Ergative and Nominative are syncretised except for the first person singular which distinguishes all four cases. Other cases include Locative/Instrumental (LOC/INST), Ablative (ABL), Allative (ALL), Possessive (POSS) and Proprietive (PROP).[1] Pronouns distinguish singular, dual and plural .... All verbs must bear one of a set of suffixes referred to as the obligatory morphology. These indicate tense, mood and clausal status including switch reference marked dependent clauses.

There are also non-obligatory suffixes that precede obligatory suffixes; these may include aspect and subject number. There is a special kind of verbal category among the non-obligatory suffixes called Associated Motion, also found in other Australian languages, which indicates that "a verb-stem action happens against the background of a motion event with a specific orientation in space" (Wilkins, 2006). The (somewhat simplified) morphological structure of verbs is schematically depicted in Figure 1.

There is also reduplication, which when applied to verbs does not have a straightforward relationship to the verb structure of Figure 1. There are many varieties of reduplication: one that we refer to later in the paper is the Attenuative. The Attenuative form of some verb X is often glossed as 'start to X', and is

---

[1] Authors' note: Other abbreviations we use in examples are ASSOC MOTION = Associated Motion, ATTEN = Attenuative reduplicant, DO.COMING = type of Associated Motion marker, EMPH = emphasis particle, FOC = focus particle, INTENS = intensifier, ITER = iterative, IV = intransitive verbaliser, NOMLSR = nominaliser, NUM = number, PRES = present tense, PRIOR.MOTION = type of Associated Motion marker, PST = past tense, PV = preverb, QUICK:DO&GO.BACK = type of Associated Motion marker, RECIP = reciprocal, REFL = reflexive, SR = switch reference, SS = same subject, TV = transitive verbaliser, V = verb.

V root (TV) Base + IV (TV) PV ... Root Non-V + TV	REFL / RECIP	number	ASSOC MOTION	aspect	ITER	tense (SR) mood (SR) SR NOMLSR

Figure 1: Simplified Eastern/Central Arrernte Verb Structure, adapted from Henderson (1998, p.276).

formed by taking the first syllable of the verb stem, adding *elpe* or *erlpe*, and then repeating the (inflected) verb; an example, giving the reduplicated form of the verb *ampangkeme*[2] in (1a), is in (1b), with the reduplicant and the reduplicated part of the base underlined.

(1)   a. *ampangk-eme*
           groan-PRES

           (he) is groaning

       b. *ampelpe-ampangk-eme*
           ATTEN-groan-PRES

           (he) is starting to groan (Henderson, 1998, (4.60a))

Henderson (1998, Section 5.3.1) contends that "a large proportion of verb components, and therefore verb forms, involve more than one stem or morphological word": these are referred to as complex, as opposed to simple, verbs. These complex verbs include lexical compounds consisting of preverbs (defined below) in combination with (following) inflecting verbs. Complex verbs and the phenomenon of Associated Motion will be central to this paper; we discuss them in more detail in Section 4 on complex predicates.

In contrast to the rigid morphology, Arrernte clausal syntax is relatively free, with no obvious constraints on the order of phrases, although it may be the case, as in some other Australian languages (Simpson, 2007, for example), that it is pragmatically constrained. Word order within the Noun Phrase, however, is much more restricted, and case is marked exclusively on the final element of the NP.

2.2   An LFG Analysis

Following Nordlinger and Bresnan (2011), we capture Arrernte's free word order at the sentence level by assuming a flat exocentric c-structure rooted in S. We do not model in the grammar the potential pragmatic factors that control linearization; instead, for the actual system we have a separate reranking post-process to handle this. There is no explicit copula in the present tense; like Nordlinger and Sadler (2007), we allow all nominals to act predicatively. The head of a sentence can be

---

[2]In this paper we follow the standard practice of referring to verbs by their 'dictionary form', the stem combined with the present tense ending *-eme*. Also, in glossing we use the morpheme boundaries of Henderson (1998) and Henderson (2002). In full sentences, we follow the usual orthographic convention of starting with an uppercase letter.

a verb or a nominal, but only a finite verb can carry tense; with other heads, the auxiliary *aneme* (lit., 'sit') can be added to carry such information if desired. Grammatical functions are specified with dependent-marking (Bresnan, 2001, p111), and NPs with a semantic case such as LOC function as modifiers.

NPs, in contrast, have relatively fixed internal word order, with the possible exception of some 'floating' of demonstratives and counting terms; we model this with the separation of immediate dominance and linear precedence constraints (Falk, 2001). NP case (ergative/absolutive, as well as the numerous other cases such as locative, ablative, etc.) is handled via 'particles' in syntax, which always appear at the end of the whole NP. We model this by adding a projection level over the core NP.

For the most part, verb morphology is handled in the morphological component; this includes some relatively complicated cases of prefixing reduplication including the Attenuative mentioned in Section 2.1, which we handle in XFST using compile-replace rules (Beesley and Karttunen, 2003; Bögel et al., 2007). In between these clear-cut cases of aspects handled by the grammar versus those handled by the morphology, however, there is the grey area of complex predicates.

3   Definitions of Complex Predicates

According to an influential definition by Butt (1993), a complex predicate has to satisfy three conditions: (a) the argument structure is complex (two or more semantic heads contribute arguments); (b) the grammatical functional structure is that of a simple predicate — it is flat, and there is only a single predicate (paraphrased by Nordlinger (2010) as "monoclausal"); and (c) the phrase structure may be either simple or complex — it does not necessarily determine the status of the complex predicate.

While this definition is widely accepted within the LFG community, it ought to be noted that there also are more general definitions, e.g. Amberber et al. (2010) in their book on complex predicates, where they acknowledge that there is no agreed set of criteria for defining a complex predicate; Butt (1993) also gives the same caveat. Further, there are interesting cases that are still naturally analyzed within LFG as compex predicates although they do not meet Butt's criteria, such as the Associated Motion construction in Wambaya as described by Nordlinger (2010). We review this briefly here, along with work by Wilson (1999) on the Australian language Wagiman, both for what they have to say about the definitions of complex predicates, and for the characteristics that are similar to the complex predicates that we discuss in this paper.

Wambaya is a non-Pama-Nyungan language, and therefore relatively distant from Arrernte. Its word order is free, but there is an auxiliary obligatorily in second position (Nordlinger, 1998b). An Associated Motion marker is optionally attached to the auxiliary; there must also be a main verb. Nordlinger (2010) notes: "When combined with a motion verb, the Associated Motion marker adds the direction of

the motion event, since motion verbs in Wambaya are direction-oriented. When the main verb is a non-motion verb, however, the Associated Motion affix encodes a sequential event 'go/come and VERB'." Following Broadwell (2000) on Choctaw, Nordlinger (2010) characterises these as single events and dual events respectively.

She asserts that the Associated Motion construction is monoclausal, as it contains only a single main verb and a single subject, and the Associated Motion marker cannot constitute a clausal predicate on its own. The satisfaction of condition (a) above, however, is less clear, in terms of the extent to which the Associated Motion marker can be considered a semantic predicate in the absence of syntactic predication tests. Nordlinger (2010) argues that it does on the grounds that it adds a motion predication, as well as sensitivity to the semantics of its lexical verbs.

Wagiman is also a non-Pama-Nyungan language, aspects of which are described by Wilson (1999) and reanalysed in Andrews and Manning (1999). It contains coverbs analogous to the preverbs of Arrernte (see Section 4), which are analysed by Wilson (1999) as complex predicates. They differ from other instances of complex predicates in that both components can occur independently and act as full predicates, in contrast to earlier examples of complex predicates where one component was essentially a light verb (e.g. the Romance causatives of Alsina (1997) or the Urdu permissive of Butt (1993)).

In this paper we take the definition of what counts as a complex predicate from Henderson (2002). Our analysis based on this definition is broadly in the spirit of Butt (1993); we will draw attention to the situations where it is not.

4   Complex Predicates in Arrernte

4.1   Three Types of Complex Predicate

Henderson (2002), our source of the definition of Arrernte complex predicates, more generally investigates the problematic nature of a word in Arrernte: in some contexts an element of the language will appear to be, say, a derivational morpheme, and in others a separate word. Henderson (1998) claims that "a large proportion of verb components, and therefore verb forms, involve more than one stem or morphological word, [and that this notion of complex predicates] provides a coherent account of a number of phenomena"; he then gives phonological, morphological (e.g. the possibility of reduplication) and grammatical criteria for determining wordhood in Arrernte. These broadly agree, although not always. The following are two of the phonological criteria from Henderson (2002).

Prosodically conditioned allomorphy   The forms of the Reciprocal, Dual and Plural verb suffixes depend on the number of syllables between the beginning of the phonological word and later verb suffixes. For Dual and Reciprocal, the morphemes *err* and *irr* are used if the stem has an odd or even number of syllables, respectively; for Plural these morphemes are *errirr* and *irrer*, with an additional alternative *ewarr* that may apply to stems of more than one syllable.

Stress   Each word bears a primary stress on the first syllable beginning with a consonant. In (relatively uncommon) words of four or more syllables, there may be stress on alternating syllables after the primary stress.

Henderson (2002) notes that there is no simple definition of a word in Arrernte in terms of grammatical criteria. Nominal morphology is limited to compounding. Verbs take suffixes as described in Section 2.1; the order of these morphemes is largely fixed, as in Figure 1. There are a number of other factors that can be taken as indicators of word status, however, including the two following:

Intervening material   It is possible for some non-verbal morphemes to intervene at specific points within the verb. For example, in (2), the particle *akwele* 'supposedly' (which can appear on its own, outside of a verb or an NP) appears inside the verb, between the stem *arrerne* and its suffixes.

(2)   *arrerne akwele lh-eme*
       place   SUPPO REFL-PRES
       supposedly sit down (Henderson, 2002, (9))

Reduplication   In addition, the manner and location of reduplication, which applies to verbs, gives an indication of the boundaries of a verbform.

Given the definition of complex predicate for Arrernte based on the criteria of Henderson (2002), we look at three particular types: the intransitive verbaliser (IV), lexical compounds, and Associated Motion. In the following subsections, we then consider, in light of the problematic nature of the notion of word, whether each of these three should be handled in the syntax or the morphology of our overall LFG grammar; and if in the syntax, what kind of verbs — e.g. full verbs, light verbs or auxiliaries — are involved.

Intransitive Verbaliser   The IV *irreme*, in (3), is a highly productive element of the language that follows and combines with a base that can be a nominal, adverb, NP, or sometimes a clause: in (3a), it combines with the nominal *mwerre* 'good'. Its basic sense is inchoative, although it can function as a copula; the derived form of base + IV functions as an intransitive verb. The IV has often been treated as a derivational morpheme, as in Figure 1.

(3)   a.   *Utyene tiwelhe-me   mwerre-irr-eme-le*
           sore   fall.off-PRES good-IV-PRES-SS
           The scab falls off and the sore gets better.
           (Henderson and Dobson, 1994, entry for *tiweme*)

      b.   ***Alakenhe** re   ampe akweke mpwe ulk-etyenh-ele*
           thus   3sg.NOM child small   urine excrete-FUT-SS
           irr-entye.akngerre
           IV-NOMLSR
           Little kids behave that way when they need to have a leak.
           (Henderson, 2002, (20))

Lexical compounds   Lexical compounds as in (4) consist of a preverb followed by and combined with an inflecting verb: in (4a), *lthere* is the preverb, and *iweme* the inflecting verb. In some ways, then, it is similar to the IV, although it is much less productive and more lexically idiosyncratic; the fixed order and lack of productivity also distinguish them from adverb-verb combinations. Also as with the IV, lexical compounds have often been treated as the result of derivational morphology, as in Figure 1.

(4)   a.   *Arelhe-le   ampe lthere iw-eme*
      woman-ERG child pinch$_1$ pinch$_2$-PRES
      The woman is pinching the child. (Henderson, 1998, (5.28), modified)

   b.   *Angeme the   pelhe-iw-eke*
      fly   1sg.ERG spit$_1$-spit$_2$-PST
      I spat the fly out. (Henderson and Dobson, 1994, entry for *pelhe-iweme*)

   c.   *Ampe yanhe-le-ame   apmere irnterre anthurre*
      child that-ERG-EMPH place   INTENS INTENS
      akerre-iw-eme
      scatter$_1$-scatter$_2$-PRES
      That child is scattering things all over the camp.
      (Henderson and Dobson, 1994, entry for *akerre-iweme*)

   d.   ***Ikerrke*** *anthurre akwele re*   ***iw-elh-eke***
      stick$_1$ INTENS SUPPO 3sg.NOM stick$_2$-REFL-PST
      He supposedly got himself really stuck. (Henderson, 1998, (5.5))

Associated Motion   In Associated Motion constructions as in (5), the Associated Motion marker occurs between the verb stem and the obligatory morphology: in (5a), the Associated Motion marker is *artn.alp*,[3] which adds the meaning of quickly going and returning while performing some other action.

(5)   a.   *Artwe angk-artn.alp-eke*
      man   speak-QUICK:DO&GO.BACK-PST
      The man quickly spoke and then went back.
      (Wilkins, 2006, (15c), modified)

   b.   *Ar-ety-arle   akwele alh-err-eme*
      see-PRIOR.MOTION-FOC SUPPO GO-DUAL-PRES
      Two supposedly go and then see. (Henderson, 2002, (29))

   c.   *Artwe angk-inty-eke*
      man   speak-DO.COMING-PST
      The man spoke while coming this way. (Wilkins, 2006, (15a), modified)

---

[3]The period in the middle is conventionally used in a gloss of a separable Associated Motion marker, discussed below.

Associated Motion can occur with almost all verbs; the exceptions are what Wilkins (2006) characterises as "deictic" motion verbs (e.g. *alheme* 'go', *alpeme* 'go (and come) back'). This incompatibility accords with Associated Motion having much the same semantics as the deictic motion verbs. Apart from this, the semantic contribution of the Associated Motion marker is broadly the same for motion and non-motion verbs. The Associated Motion marker in Arrernte, then, is not sensitive to the verb to which it is linked in the manner of Wambaya, but the grounds of semantic predication for complex predicate status still hold. Wilkins (1989) sees the Associated Motion marker as a morpheme for which there is a specific slot in the verb stem.

4.2 Syntax or Morphology?

The default position, then, might be to handle all three types in the morphology, as suggested by Figure 1. However, in light of Henderson (2002), we note the following points and then make proposals about where to handle each type.

For the IV construction, Attenuative reduplication — with its two possible positions as given in (6) — in conjunction with the other criteria mentioned above, indicates that the combined form is not a single simple verb. If the verb were a simple one, and the IV consequently an unequivocal derivational morpheme, only (6a) would be valid; (6b) indicates that *irreme* has at least a quasi-independent status.

(6) a. *mwelpe-mwerre-irr-eme*
         ATTEN-good-IV-PRES
         start to get better (Henderson, 1998, (4.66))

     b. *mwerre-irrerlpe-irr-eme*
         good-ATTEN-IV-PRES
         start to get better (Henderson, 1998, (4.66))

The same argument can be made for lexical compounds, to which the Attenuative applies in a similar fashion. Each inflecting verb used in a lexical compound is in all cases homophonous with a free verb, whose meaning is sometimes obviously related but sometimes not. In (4a) and (4c), the inflecting verb is homophonous with *iweme* 'throw (away)'. Henderson (1998) notes that there is a "continuum of semantic compositionality" ranging from cases where there is almost no sense of the free verb, as in (4d), to ones with a more transparent sense, as in (4c).

For associated motion, Henderson (2002), based on a range of further data than Wilkins (1989), notes that the construction e.g. in (5a) could instead be glossed as an Associated Motion particle *artn* and the full verb of motion *alpeme* 'to go and come back'; that there are phonological grounds for considering them separate words; and that some intervening material is possible. (5b) contains the Associated Motion marker *ty.alh* (potentially an Associated Motion particle *ty* and the full verb

of motion *alheme* 'go'), which is separated by the focus particle *arle* and *akwele* 'supposedly' (and is consequently written as two words in the example).

We note that all of these constructions in fact permit intervening material. In addition to (5b) just described, in (4d), the lexical compound *ikerrke-iweme* (in bold) also has *akwele* intervening, as well as the intensifier *anthurre* and the pronoun *re*; in (3b), between the base *alakenhe* 'thus' and the IV (also in bold) there is a dependent clause. This brings into focus the question of whether all of these should be considered separate words for an LFG analysis, and so, for an XLE implementation, perhaps more naturally handled in the grammar rather than the morphology.

In coming to a view about this, it is useful to consider the separability cline of Henderson (2002), reproduced here. This cline groups into classes the kinds of intervening material permitted in complex predicates:

1. certain particles and clitics: e.g. *anthurre* Intensifier, *akwele* 'supposedly', *arle* FOC;

2. *akwete* 'still';

3. third person singular pronoun functioning non-referentially as an emphatic;

4. other pronominal NPs;

5. simple non-pronominal NPs, most likely being a single nominal;

6. other adverbs, complex NPs;

7. dependent clauses.

In terms of applicability to the various kinds of complex verbs, the cline ranges from most to least widespread: that is, intensifiers such as *anthurre* in class 1 are applicable to the widest range of complex verbs, while the dependent clauses of class 7 are the most restricted. Furthermore, as can be seen from the cline above, the size of the units of intervening material broadly increases from class 1 to class 7. In our examples, (5b) illustrates intervening material from classes 1 and 2; (4d) from classes 1 and 3; and (3b) from class 7, the most extensive attested type of intervening material. Henderson (2002) notes that the cline is roughly implicational in a number of ways, in particular that if a type of complex verb allows intervening material of class n, it also allows intervening material of classes $1 \ldots n-1$.

As illustrated by (3b) for the IV construction, the potentially unbounded amount of intervening material from all separability classes between the base and the IV suggests handling these within the grammar as separate words.

Verbs with Associated Motion, by contrast, are attested as permitting intervening material only from classes 1 and 2, like the class 1 element *akwele* 'supposedly' in (5b); these are finite and quite small in extent. In addition, not all Associated Motion morphemes can be decomposed into smaller components that correspond to some motion verb: (5c) has the morpheme *intye*, typically glossed 'do X while coming this way', which has no obvious free verb counterpart.[4] The nature of the

---

[4]Wilkins (1989, p277) states that *intye* "is itself likely to have originated from a former motion

Associated Motion construction is then similar to those of Wambaya, as described in Nordlinger (2010) (and briefly in Section 3), although there is not the same freedom of movement of the component parts. An appropriate place to handle Associated Motion would then be in the morphology, with a slot inside the verb structure as in Figure 1, and with additional internal slots for morphemes of classes 1 and 2 of the separability cline.

Lexical compounds fit somewhere in the middle. There are attested utterances with intervening material of classes 1–5, but none with the extent of intervening material of the IV construction. The situation here then is more ambiguous between syntax and morphology than the other two cases. We propose to handle this in the syntax, given its similarity to the IV construction; there are also further reasons, which will become clear in Section 5.1, where we discuss our glue semantics-based approach and the incorporation into our representation of the notion of lexical functions from Meaning-Text Theory (MTT).

4.3 Full Verb, Light Verb or Auxiliary?

We now consider the two constructions to be handled in the grammar, the IV construction and lexical compounds. Our position is that neither the IV nor the inflecting verb of lexical compounds should be considered a full verb, which would imply that the base or preverb respectively would consequently be some kind of argument. Attenuative reduplication applies only to verbs, so the base or preverb, which can validly be reduplicated as in (6a), cannot be a plain nominal (or adverb, etc).

In addition, case marking supports this. If *mwerre* 'good' in (3a) were the object of a full verb *irreme*, then the derived compound form would be transitive, and the subject *utyene* 'sore' would be marked with the ergative marker *-le*; and this is not the case. And for lexical compounds, case marking (the ergative *-le* on the NP *ampe yanhe* 'that child') in (4c) indicates that the verb is transitive.

Regarding the status of the IV and inflecting verbs as light verbs versus auxiliaries, we draw on the helpful synthesis of Seiss (2009), which aims to bring together various definitions that have been used in the field for auxiliaries, serial verbs and light verbs, in order to work towards a common cross-linguistic usage.

Our lexical compounds consisting of preverb + inflecting verb are in fact quite similar to the inflecting verb + coverb that is part of the case study of the Australian language Ngan'gityemerri that Seiss (2009) uses to illustrate definitions of inflecting elements as auxiliaries, serial verbs or light verbs. Given this parallel, and drawing on the characterisation of light verbs in Butt (2010), the inflecting verb would be a light verb: a key characteristic is that "light verbs exhibit subtle lexical semantic differences in terms of combinatorial possibilities", and lexical compounds as we have already noted, have a high degree of lexical idiosyncrasy in terms of inflecting verbs.

---

verb meaning 'come' ", and that this has been argued for related languages by Koch (1984). However, this has no bearing on a synchronic analysis as in this paper.

The IV is less straightforward. It is not lexically idiosyncratic in the manner of the lexical compound's inflecting verb. However, it would be odd to characterise it as an auxiliary. Seiss (2009) notes that auxiliaries typically do not contribute semantic information about the type of event, whereas light verbs can: the IV by its nature typically adds an inchoative meaning to the whole complex verb. In addition, light verbs and not auxiliaries can change the valency of a construction. In the IV construction, the base is fundamentally a nominal (or adverb, etc), which can perhaps be considered to be acting as a verb in this context, based on the Attenuative as discussed above. It is unclear what the valency of the base by itself would be, but the IV definitely enforces intransitivity as its fundamental function.

We then treat the IV and the lexical compound's inflecting verb as light verbs. The discussion following links this to our existing treatment of lexical resources that have much in common with light verbs, and hence to our handling of complex predicates at the level of f-structure. In light of the complexities of the Arrernte system of spatial and movement relations (Wilkins, 2006), we leave open exactly what the semantics for Associated Motion should be, and consequently what the resulting complex predicate would look like; we thus only discuss the IV and lexical compounds in the remainder of the paper.

5 A Representation of Arrernte Complex Predicates

5.1 Meaning-Text Theory's Lexical Functions

It has been the experience of large-scale MNLG systems (Wanner et al., 2010, for example) that as much of the system as possible should be language-independent, a position advocated from the early days of the field by Bateman et al. (1991) and Cahill et al. (2000) among others. Among the mechanisms for enhancing language independence are the so-called lexical functions from MTT (Mel'cuk, 1996; Kahane and Polguère, 2001), which embody recurrent patterns of collocations. These abstract away from language-dependent collocations, such as the English *outright lie* versus French *mensonge éhonté* 'shameless lie', as well as language-internal collocational variation, such as *heavy rain*, *strong wind* or *intense bombardment* which all refer to the intensification of some phenomenon. This particular semantic notion of intensification or strength is represented by `Magn(L)`; another lexical function of interest is $Oper_1(L)$, where a semantically (mostly) empty verb serves as syntactic support to link a predicative noun to its most prominent semantic argument, for example $Oper_1$(TALK)=GIVE, $Oper_1$(ATTENTION)=PAY. Lexical functions provide an efficient mechanism for describing a wide range of collocations. In Lareau et al. (2011) we showed how these can be incorporated into LFG using glue semantics; see also the companion paper in this volume.

Among the types of collocations described in MTT are support verbs; one such is $Oper_1(L)$ above. While there is no universally agreed definition of support verbs, we follow Fillmore et al. (2003) who, in discussing the nature of "semantically transparent" lexical elements as part of the FrameNet formalisation, charac-
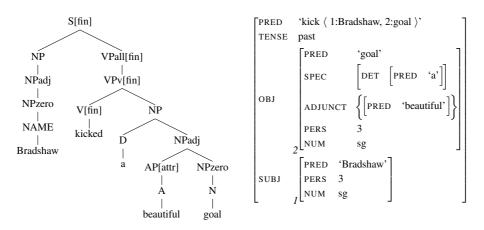
```
                S[fin]
               /      \
            NP          VPall[fin]
            |              |
          NPadj          VPv[fin]
            |            /       \
         NPzero      V[fin]       NP
            |          |         /   \
          NAME       kicked    D      NPadj
            |                   |     /    \
        Bradshaw               a  AP[attr]  NPzero
                                    |         |
                                    A         N
                                    |         |
                                beautiful    goal
```

$$
\begin{bmatrix}
\text{PRED} & \text{‘kick}\,\langle\,\text{1:Bradshaw, 2:goal}\,\rangle\text{’} \\
\text{TENSE} & \text{past} \\
\text{OBJ} &
\begin{bmatrix}
\text{PRED} & \text{‘goal’} \\
\text{SPEC} & \begin{bmatrix} \text{DET} & \begin{bmatrix} \text{PRED} & \text{‘a’} \end{bmatrix} \end{bmatrix} \\
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{‘beautiful’} \end{bmatrix} \right\} \\
\text{PERS} & 3 \\
\text{NUM} & \text{sg}
\end{bmatrix}_2 \\
\text{SUBJ} &
\begin{bmatrix}
\text{PRED} & \text{‘Bradshaw’} \\
\text{PERS} & 3 \\
\text{NUM} & \text{sg}
\end{bmatrix}_1
\end{bmatrix}
$$

Figure 2: c-structure (left) and f-structure (right) for *Bradshaw kicked a beautiful goal*

terise support verbs as "broader than the traditional notion of light verb". We apply our mechanism for this broader class of support verbs to the light verbs of Section 4.3. We thus review briefly here our mechanism for dealing with lexical functions, which then leads in to our glue-semantics-based approach to handling Arrernte complex predicates.

The domain of our MNLG system is Australian Football League (AFL) football. Consider sentence (7), with c-structure and f-structure as in Figure 2.[5]

(7)   Bradshaw kicked a beautiful goal.

In MNLG, the system starts with some representation of the input, and generates text from that, generally passing through a number of stages; in an LFG context, this first-stage input (after selection of the relevant content) might be first-order predicate logic, or perhaps more expressive representations such as intensional logic or Discourse Representation Theory (DRT) (Kamp and Reyle, 1993), with this semantic representation then mapping to f-structure and then c-structure, and at the end producing the generated text. If the meaning representation for generating sentence (7) were as in (8a), a more literal meaning representation where there is a bijection between words and terms in the meaning representation, the mapping to f-structure would be quite straightforward; if, however, the meaning representation were as in (8b), where the semantically empty element is omitted and the attribute generalised, the mapping would need some more powerful mechanism. This second meaning representation is in fact the relevant one for our system: it abstracts away from the collocationally determined use of *beautiful* to describe a good goal (which in the language of football commentary is largely interchangeable with *magnificent*, *superb*, . . .) and *kick* to describe the scoring of the goal. A

---

[5]The c-structure here broadly follows the PARC Starter English Grammar: `http://www2.parc.com/isl/groups/nltt/xle/doc/PargramStarterGrammar/starternotes.html`.

Bradshaw	N	$(\uparrow\text{PRED})=$'Bradshaw'
		Bradshaw : $\uparrow_\sigma$
goal	N	$(\uparrow\text{PRED})=$'goal'
		goal : $\uparrow_\sigma$
beautiful	A	$(\uparrow\text{PRED})=$'beautiful'
		$\lambda X.\text{beautiful}(X) : (\text{ADJ} \in \uparrow)_\sigma \multimap (\text{ADJ} \in \uparrow)_\sigma$
kicked	V	$(\uparrow\text{PRED})=$'kick$\langle(\uparrow\text{SUBJ}),(\uparrow\text{OBJ})\rangle$'
		$(\uparrow\text{TENSE})=$past
		$\lambda X.\lambda Y.\text{kick}(X,Y) : (\uparrow\text{SUBJ})_\sigma \multimap [(\uparrow\text{OBJ})_\sigma \multimap \uparrow_\sigma]$
a	D	$(\uparrow\text{PRED})=$'a'
		$\lambda X.X : (\text{DET} \uparrow)_\sigma \multimap (\text{DET} \uparrow)_\sigma$

Figure 3: Lexical entries with meaning constructors for mapping between the f-structure of Figure 2 and the literal meaning of (8a).

goal, worth six points, can only be scored in AFL by kicking; touching with any other body part results in a 'behind', worth one point. *Kick* therefore is semantically empty.

(8) a. kick(bradshaw, beautiful(goal))

   b. good(goal(bradshaw))

To describe the mapping between our desired semantics in (8b) and our f-structure in Figure 2, we use glue semantics as described in Dalrymple (2001). Briefly, in a glue semantics approach a lexical entry contains a meaning constructor made up of two parts: the lefthand (meaning) side represents the meaning, and the righthand (glue) side represents a logical formula over semantic structures corresponding to those meanings. We first give in Figure 3 the lexical entries that would be required for mapping the literal semantics of (8a) to the f-structure. The entry for *kick*, for example, is just the standard one for a transitive verb.[6]

To handle the mapping between the semantics of (8b) and the f-structure, we would add the entries of Figure 4. Here GOAL, by contrast, is a unary predicate: $\lambda X.\text{goal}(X)$, i.e. 'X goals', so to speak. However, in the construction under consideration here, its semantic predicativity is not echoed in syntax: there is no verb *to goal* in standard English, which is why a support verb is needed in the first place. *Kick* is this support verb, and so adds nothing to the final semantic form. *Kick* is only a support verb, however, in the context of *goal*, which is enforced by the constraining equation. *Beautiful* is similar to before, but has the more generic semantics of positive appreciation: $\lambda X.\text{good}(X)$.

Many of these more complex syntax-semantics mappings are in fact fairly regular for MTT lexical functions, and can be captured using templates, an XLE mechanism that can be used to implement LFG's Lexical Rules. For example, for the

---

[6]There are various ways of handling the determiner *a*, e.g. as a quantifier. We give only a simple treatment for illustrative purposes, where *a* contributes nothing to the semantics.

goal	N	($\uparrow$PRED)='goal'
		$\lambda X.\mathrm{goal}(X) : ((\mathrm{OBJ}\uparrow)\ \mathrm{SUBJ})_\sigma \multimap \uparrow_\sigma$
kicked	V	($\uparrow$PRED)='kick$\langle(\uparrow\mathrm{SUBJ}),(\uparrow\mathrm{OBJ})\rangle$'
		($\uparrow$OBJ PRED)$=_c$ 'goal'
		($\uparrow$TENSE)=past
		$\lambda X.X : (\uparrow\mathrm{OBJ})_\sigma \multimap \uparrow_\sigma$
beautiful	A	($\uparrow$PRED)='beautiful'
		$((\mathrm{ADJ} \in \uparrow)\ \mathrm{PRED})=_c$ 'goal'
		$\lambda X.\mathrm{good}(X) : (\mathrm{ADJ} \in \uparrow)_\sigma \multimap (\mathrm{ADJ} \in \uparrow)_\sigma$

Figure 4: Additional lexical entries with meaning constructors for mapping between the f-structure of Figure 2 and the desired meaning of (8b).

lexical function $\mathtt{Oper_1(L)}$, which represents the use of support verbs in contexts such as that of *kick* in our examples, the following template in (9) could be defined.

(9) @OPER1(L)=($\uparrow$PRED)='%stem$\langle(\uparrow\mathrm{SUBJ}),(\uparrow\mathrm{OBJ})\rangle$'
            ($\uparrow$OBJ PRED)$=_c$ 'L'
            $\lambda X.X : (\uparrow\mathrm{OBJ})_\sigma \multimap \uparrow_\sigma$

5.2   Handling Complex Predicates via Glue

As may already be apparent from the analysis of the IV and lexical compound complex predicates as containing light verbs, we can use this exact same mechanism to handle them in the grammar component: we can take the elements of the f-structure corresponding to complex predicates and use glue semantics to combine them together. This is quite different from the more common approach in LFG, which has generally used the LCS of Jackendoff (1990). In an early version this approach, exemplified by Butt (1995), complex predicates are formed at a-structure by combining an LCS containing a 'transparent event' position (for example, an LCS corresponding to a light verb) with a fully specified LCS representing its argument. Andrews and Manning (1999) demonstrate some problems with this approach, including that it does not handle the combination of two full predicates, such as in Wagiman. They propose instead an approach to complex predicates using restriction projections; their approach can also use the mechanism of LCS, but they also put forward (in Appendix A) glue semantics as an alternative mechanism for the complex predicate combination. Andrews (2007) notes, albeit with a different formulation of glue semantics in that paper, that in some ways glue semantics is mimicking the effect of the alternative LCS approach at a-structure. In fact, it does away with a-structure, enforcing in the appropriate way the combination of complex predicate components at s-structure. Our approach in this section has some similarities to those of Andrews and Manning (1999) and Andrews (2007), but is implemented using the lexical functions we have adopted from MTT, as described in Section 5.1. We illustrate it using the IV construction and lexical compounds.

For the IV, the appropriate lexical function is $\texttt{IncepOper}_1\texttt{(L)}$, similar to $\texttt{Oper}_1\texttt{(L)}$ but referring to a support-like verb indicating the start of something (e.g. *contract a disease*). We define the template INCEPOPER1 at the top of Figure 5. To illustrate its use, we take a slightly simpler version of sentence (3a) above, in (10). *irreme* would instantiate this template; *mwerre* 'good' is a nominal that is verbalised by *irreme*. For the generalised f-structure and lexical items with semantics (Figure 5 centre left), we obtain via the glue semantics of the lexical entries the desired overall semantics for the complex predicate (Figure 5, bottom and centre right, respectively).[7]

(10) *Artwe mwerre-irr-eme*
      man   good-IV-PRES
      The man is starting to get better.

Note that in contrast with the LCS approach to complex predicates, *mwerre* 'good' appears as a separate, embedded f-structure, not yet combined with the IV, as this happens at s-structure via glue semantics. We noted in Section 4.3 that the base *mwerre* would not be an object; in fact it is not obvious what the most appropriate category might be. For this example we use PREDLINK (Butt et al., 1999): it is used standardly for copula or copula-like constructions (e.g. by Sulger (2009) for a closed complement double-tier analysis of Irish copulas). We also note that the lexical function $\texttt{IncepOper1}$ is contributing the inchoative aspect of the semantics, which is somewhat different from the earlier LCS approaches.

For lexical compounds, we consider examples (4a) and (4b). One question is whether there are two full predicates here that are able to operate independently, as in Wagiman (Section 3), or whether a support verb analysis would be suitable. While for (4a) there is no attested independent use of the nominal *lthere*, in (4b) *pelhe* is an attested independent nominal meaning 'spit', and as noted earlier there is the full verb *iweme* 'throw' which in this context is perhaps indicating the motion of the spittle. However, although *pelhe* as a preverb can occur with the Attenuative (suggesting some characteristics of a verb) it cannot occur with any other verbal morphology, and it is not attested as acting as a full predicate independently, so for the construction as a whole we adopt the support verb analysis. Consequently, we use the lexical function $\texttt{Oper}_{12}$ to represent a plain support verb that takes as its subject the first semantic argument of the base of the collocation (for (4a), *arelhe*), the second argument as its direct object (*ampe*), and the base itself (*lthere*); the structure is parallel for (4b).

The question then is what grammatical function is appropriate for the base. In the case of (4b), it may again be PREDLINK: this is a resultative construction which is quite similar to a construction of the Urdu/Hindi reference dependency bank for complex predicates (Ahmed et al., 2012), there termed the resultative complex predicate with noun. There, PREDLINK represents what is predicated of a certain

---

[7]We note that there is nothing tying us to this particular semantics. Neo-Davidsonian semantics would work equally well, or the DRT or intensional semantics mentioned earlier.

@INCEPOPER1(L)=

    ($\uparrow$PRED)='IncepOper1$\langle$($\uparrow$SUBJ),($\uparrow$PREDLINK)$\rangle$'

    ($\uparrow$PREDLINK PRED)=$_c$'L'

    $\lambda X.\text{start}(X)$: ($\uparrow$PREDLINK)$_\sigma$ $\multimap$ $\uparrow_\sigma$

$$\begin{bmatrix} \text{PRED} & \text{'IncepOper}_1\langle \textit{2:}\text{artwe, } \textit{3:}\text{mwerre}\rangle\text{'} \\ \text{TENSE} & \text{pres} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'artwe'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{sg} \end{bmatrix}_2 \\ \text{PREDLINK} & \begin{bmatrix} \text{PRED} & \text{'mwerre'} \end{bmatrix}_3 \end{bmatrix}_1 \qquad \text{start(well(man))}$$

[artwe]   man : $\uparrow_\sigma$

[mwerre] $\lambda X.\text{well}(X)$: ((PREDLINK$\uparrow$) SUBJ)$_\sigma$ $\multimap$ $\uparrow_\sigma$

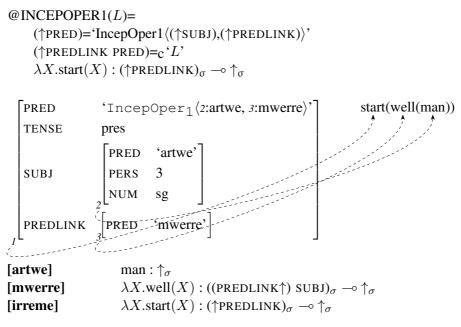**[irreme]** $\lambda X.\text{start}(X)$: ($\uparrow$PREDLINK)$_\sigma$ $\multimap$ $\uparrow_\sigma$

Figure 5: Analysis for (10): lexical function template for INCEPOPER1 (top left), f-structure for IV (centre left), semantic representation for IV (centre right), lexical entries with glue semantics (lower left)

entity (in their case and ours, the syntactic object; for us, what is predicated is that it has been spat out). Ahmed et al. (2012) comment that the resultative nature of the construction is not overt in the f-structure, and must be inferred from the existence of PREDLINK. This is somewhat different from the canonical use of PREDLINK, and for us does not seem applicable to (4a).

    An alternative could be COMP, assuming as does Lødrup (2012) that nominal COMPs are possible. This analysis is perhaps more controversial, and we do not have space to present a full justification of it here, but we note that the essential point is that what we are looking for is a fairly neutral grammatical function whose role is just to give the base a place in the f-structure. This is indeed how Lødrup (2012) describes nominal COMP: "The intuition behind the COMP function could be verbalized this way: COMP differs from the other complement functions by not having their properties; it is a complement that just 'is there', and does not take part in grammatical processes."[8] Our analysis is then as in Figure 6. COMP would

---

[8]Rachel Nordlinger (personal communication) has considered nominal COMP to be potentially applicable to Australian Aboriginal languages such as Wambaya, for instance in (11).

(11)   *Ngawu ngu   bungmanya mirra*
        1sg.NOM 1sg.S-FUT old.woman be

        I will live to be an old woman. (Nordlinger, 1998a)

She notes that NPs such as *bungmanya*, often called (subject) complements in the literature, are a kind of nominal which relates in some way to another nominal without it clearly being a modifica-

$$\begin{bmatrix} \text{PRED} & \text{`Oper}_{12}\langle\textit{2}\text{:arelhe, }\textit{3}\text{:ampe, }\textit{4}\text{:lthere}\rangle\text{'} \\ \text{TENSE} & \text{pres} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`arelhe'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{sg} \end{bmatrix}_{\textit{2}} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`ampe'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{sg} \end{bmatrix}_{\textit{3}} \\ \text{COMP} & \begin{bmatrix} \text{PRED} & \text{`lthere'} \end{bmatrix}_{\textit{4}} \end{bmatrix}_{\textit{1}}$$

pinch(woman, child)

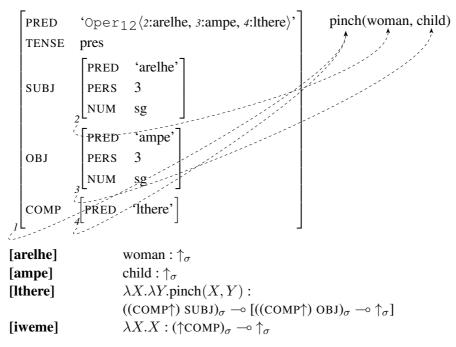|	
[arelhe]	woman : $\uparrow_\sigma$
[ampe]	child : $\uparrow_\sigma$
[lthere]	$\lambda X.\lambda Y.\text{pinch}(X, Y)$:
	$((\text{COMP}\uparrow)\ \text{SUBJ})_\sigma \multimap [((\text{COMP}\uparrow)\ \text{OBJ})_\sigma \multimap \uparrow_\sigma]$
[iweme]	$\lambda X.X$: $(\uparrow\text{COMP})_\sigma \multimap \uparrow_\sigma$

Figure 6: Analysis for (4a): f-structure for lexical compound (upper left), semantic representation (upper right), lexical entries with glue semantics (lower left)

also be applicable to the IV construction from earlier.

From these examples it can be seen that the fundamental indicator of complex predication in our approach is the occurrence of lexical functions as the PRED value, in conjunction with the result at s-structure: we thus also arguably diverge from the LFG-specific definition of complex predicate by Butt (1993), in that in our representation they do not constitute a single predicate at f-structure.[9]

6   Conclusion

In this paper we have looked at three kinds of complex predicate in the Australian Aboriginal language Eastern/Central Arrernte: the intransitive verbaliser construction, lexical compounds, and Associated Motion. Based on the criteria of Henderson (2002) for characterising a word in Arrernte, we have argued that the first two types of complex predicate are more naturally handled in an LFG grammar, and the last in the morphology. We have then shown how a mechanism for incorporating the lexical functions of Meaning-Text Theory into LFG via glue semantics, developed as part of a multilingual natural language generation system, extends

---

tional structure.

[9]Note that whereas our semantic structure in Figure 5 happened to have two separate elements at s-structure corresponding to the two parts of the complex predicate, because of choices made in the semantic representation, in Figure 6 the two parts of the complex predicate combine to give a single element at s-structure, as in more traditional approaches to complex predicate combination in LFG.

naturally to form a mechanism for formation of complex predicates of at least the first two types in the semantics. In terms of future work, a full treatment of Associated Motion remains to be carried out. A potential direction here is the approach detailed in recently published monograph of Mani and Pustejovsky (2012), which describes itself as "analyz[ing] the semantics of motion expressions in terms of the formalisms of qualitative spatial reasoning".

References

Ahmed, Tafseer, Butt, Miriam, Hautli, Annette and Sulger, Sebastian. 2012. A Reference Dependency Bank for Analyzing Complex Predicates. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey.

Alsina, Alex. 1997. A Theory of Complex Predicates: Evidence from Causatives in Bantu and Romance. In Alex Alsina, Joan Bresnan and Peter Sells (eds.), *Complex Predicates*, pages 203–246, Stanford, CA, US: CSLI Publications.

Amberber, Mengistu, Baker, Brett and Harvey, Mark (eds.). 2010. *Complex Predicates*. Cambridge, UK: Cambridge University Press.

Andrews, Avery. 2007. Projections and Glue for Clause-Union Complex Predicates. In *Proceedings of the LFG07 Conference*, pages 44–65.

Andrews, Avery and Manning, Christopher. 1999. *Complex Predicates and Information Spreading in LFG*. Stanford, CA, US: CSLI Publications.

Bateman, John, Matthiessen, Christian, Nanri, Keizo and Zeng, Licheng. 1991. The re-use of linguistic resources across languages in multilingual generation components. In *Proceedings of the 1991 International Joint Conference on Artificial Intelligence*, volume 2, pages 966–971, Sydney.

Beesley, Kenneth R. and Karttunen, Lauri. 2003. *Finite State Morphology*. Stanford, CA, US: CSLI.

Bögel, Tina, Butt, Miriam, Hautli, Annette and Sulger, Sebastian. 2007. Developing a Finite-State Morphological Analyzer for Urdu and Hindi. In T. Hanneforth and K. M. Würzner (eds.), *Finite-State Methods and Natural Language Processing, Revised Papers of the Sixth International Workshop on Finite-State Methods and Natural Language Processing*, pages 86–96, Potsdam, The Netherlands: Potsdam University.

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

Broadwell, George Aaron. 2000. Choctaw directionals and the syntax of complex predication. In Miriam Butt and Tracy King (eds.), *Argument Realization*, pages 111–134, Stanford, CA, US: CSLI Publications.

Butt, Miriam. 1993. *The Structure of Complex Predicates in Urdu*. Ph. D.thesis, Stanford University.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Stanford, CA, US: CSLI Publications.

Butt, Miriam. 2010. The Light Verb Jungle: Still Hacking Away. In Mengistu Am-

berber, Brett Baker and Mark Harvey (eds.), *Complex Predicates*, pages 48–78, Cambridge, UK: Cambridge University Press.

Butt, Miriam and King, Tracy Holloway. 2006. Restriction for Morphological Valency Alternations: The Urdu Causative. In Miriam Butt, Mary Dalrymple and Tracy Holloway King (eds.), *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan*, pages 235–258, Stanford, CA, US: CSLI Publications.

Butt, Miriam, King, Tracy Holloway, Niño, Maria-Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. Stanford, CA, US: CSLI Publications.

Cahill, Lynn, Doran, Christie, Evans, Roger, Kibble, Rodger, Mellish, Chris, Paiva, Daniel, Reape, Mike, Scott, Donia and Tipper, Neil. 2000. Enabling resource sharing in language generation: an abstract reference architecture. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC)*, Athens, Greece.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*, volume 42 of *Syntax and Semantics Series*. New York, NY, US: Academic Press.

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

Fillmore, Charles, Johnson, Christopher and Petruck, Miriam. 2003. Background to Framenet. *International Journal of Lexicography* 16(3), 235–250.

Henderson, John. 1998. *Topics in Eastern and Central Arrernte grammar*. Ph. D.thesis, University of Western Australia.

Henderson, John. 2002. The Word in Eastern/Central Arrernte. In R. M. W. Dixon and Alexandra Aikhenvald (eds.), *Word: A Cross-Linguistic Typology*, pages 100–124, Cambridge, UK: Cambridge University Press.

Henderson, John and Dobson, Veronica. 1994. *Eastern and Central Arrernte to English Dictionary*. Alice Springs, Australia: IAD Press.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA, US: MIT Press.

Kahane, Sylvain and Polguère, Alain. 2001. Formal foundation of lexical functions. In *Proceedings of ACL 2001*, Toulouse, France.

Kamp, Hans and Reyle, Uwe. 1993. *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, volume 42 of *Studies in Linguistics and Philosophy*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Koch, Harold. 1984. The Category of 'Associated Motion' in Kaytej. *Language in Central Australia* 1, 23–34.

Lareau, Francois, Dras, Mark, Börschinger, Benjamin and Dale, Robert. 2011. Collocations in Multilingual Natural Language Generation: Lexical Functions meet Lexical Functional Grammar. In *Proceedings of the Ninth Australasian Language Technology Workshop (ALTA'11)*, pages 95–104, Canberra, Australia.

Lødrup, Helge. 2012. Is there a 'third object'? Searching a nominal counterpart to COMP. In *Proceedings of the LFG12 Conference*.

Mani, Inderjeet and Pustejovsky, James. 2012. *Interpreting Motion: Grounded Rep-*

resentations for Spatial Language. Oxford, UK: Oxford University Press.

Mel'cuk, Igor. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In Leo Wanner (ed.), *Lexical Functions in Lexicography and Natural Language Processing*, pages 37–102, Amsterdam, The Netherlands: Benjamins.

Nordlinger, Rachel. 1998a. *A Grammar of Wambaya*. Canberra, Australia: Pacific Linguistics.

Nordlinger, Rachel. 1998b. *Constructive Case: Evidence from Australian Languages*. Stanford, CA, US: CSLI Publications.

Nordlinger, Rachel. 2010. Complex predicates in Wambaya: detaching predicate composition from syntactic structure. In Mengistu Amberer, Brett Baker and Mark Harvey (eds.), *Complex Predicates*, pages 237–258, Cambridge, UK: Cambridge University Press.

Nordlinger, Rachel and Bresnan, Joan. 2011. Lexical-Functional Grammar: interactions between morphology and syntax. In Robert Borsley and Kersti Börjars (eds.), *Non-Transformational Syntax: Formal and Explicit Models of Grammar*, West Sussex, UK: Wiley-Blackwell.

Nordlinger, Rachel and Sadler, Louisa. 2007. Verbless clauses: revealing the structure within. In J. Grimshaw, J. Maling, C. Manning, J. Simpson and A. Zaenen (eds.), *Architectures, Rules and Preferences: A Festschrift for Joan Bresnan*, pages 139–160, Stanford, CA, USA: CSLI Publications.

Seiss, Melanie. 2009. On the Difference Between Auxiliaries, Serial Verbs and Light Verbs. In *Proceedings of the LFG09 Conference*, pages 501–519.

Seiss, Melanie. 2011. Implementing the Morphology-Syntax Interface: Challenges from Murrinh-Patha Verbs. In *Proceedings of the LFG11 Conference*, pages 482–502.

Simpson, Jane. 2007. Expressing pragmatic constraints on word order in Warlpiri. In Annie Zaenen, Jane Simpson, Chris Manning and Jane Grimshaw (eds.), *Architectures, Rules, and Preferences: A Festschrift for Joan Bresnan*, Stanford, CA, US: CSLI Publications.

Strehlow, T. G. H. 1944. *Aranda phonetics and grammar*. Sydney, Australia: Oceania Monographs.

Sulger, Sebastian. 2009. Irish clefting and information-structure. In *Proceedings of the LFG09 Conference*, pages 562–582.

Wanner, Leo, Bohnet, Bernd, Bouayad-Agha, Nadjet, Lareau, François and Nicklaß, Daniel. 2010. MARQUIS: Generation of user-tailored multilingual air quality bulletins. *Applied Artificial Intelligence* 24(10), 914–952.

Wilkins, David. 1989. *Mparntwe Arrernte (Aranda): studies in the structure and semantics of grammar*. Ph. D.thesis, Australian National University.

Wilkins, David. 2006. Towards an Arrernte grammar of space. In Stephen Levinson and David Wilkins (eds.), *Grammars of space: explorations in cognitive diversity*, pages 24–62, Cambridge, UK: Cambdridge University Press.

Wilson, Stephen. 1999. *Coverbs and complex predicates in Wagiman*. Stanford, CA, US: CSLI Publications.