### **NUMBER MARKING: AN LFG OVERVIEW**

# Mary Dalrymple University of Oxford

Proceedings of the LFG12 Conference

Miriam Butt and Tracy Holloway King (Editors)

2012

**CSLI** Publications

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

#### **Abstract**

This paper presents a general overview of plurality and number marking and its treatment in LFG. It was originally prepared as an introduction to the Workshop on Number and Plurals at LFG12.

Broadly speaking, number marking is morphological marking which indicates the number of verbal dependents or events involved in a situation. Number marking can appear on pronouns or nouns, indicating the number of members in the group referred to, or as agreement marking on determiners, adjectives, verbs, prepositions, and other categories. Verbal number, or pluractionality, indicates the number of events in a complex event description. Here we present a general overview of number marking patterns and their analysis in LFG. For a thorough and in-depth treatment of these issues from a crosslinguistic perspective, see Corbett (2000) and Kibort and Corbett (2008).

#### 1 Number systems

English and many European languages make a two-way distinction in number, contrasting singular and plural:

(1) a. the boy (singular: one boy)

b. the boys (plural: more than one boy)

More complex number systems are common in Austronesian languages. Subject pronouns in Boumaa Fijian (Austronesian, Oceanic; Dixon 1988) distinguish four numbers: singular, dual (two participants), paucal (a small number of participants), and plural (a larger number of participants).

(2)		singular	dual	paucal	plural
	first person inclusive	_	etaru	tou	eta
	first person exclusive	au	'eirau	'eitou	'eimami
	second person	О	omudrau	omudou	omunuu
	third person	e	erau	eratou	era

It is possible for number distinctions to vary across the pronominal paradigm, or to vary according to the type of noun being marked. For example, personal pronouns in Biak (Austronesian, South Halmahera-West New Guinea; Mofu 2009) make a four-way distinction in the third person, but a three-way distinction in first and second person.<sup>1</sup>

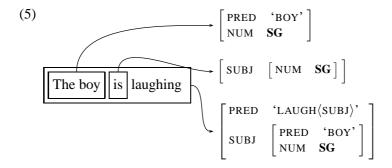
<sup>&</sup>lt;sup>1</sup>Biak has an additional animate/inanimate distinction in the plural only, violating Greenberg's Universal 45, which states that if a language makes gender distinctions in the plural, it also makes some gender distinctions in the singular; for more discussion, see Steinhauer (1985) and Mofu (2009).

(3)					plural	
		singular	dual	paucal	animate	inanimate
•	first person inclusive	_	ku	_	ko	_
	first person exclusive	aya	nu	_	inko	_
	second person	au	mu	_	mko	_
	third person	i	su	sko	si	na

#### 2 Number specification and number agreement

Number can be cospecified by the noun and the verb, as in the English examples in (4), where a singular subject requires a singular verb, and a plural subject requires a plural verb:

In LFG treatments of verb agreement, the features of the dependent must match the features of the agreeing verb. Here the subject *the boy* is singular, and the number specified by the finite verb *is* must be compatible with the number of the subject:<sup>2</sup>

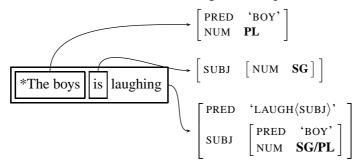


As usual in agreement relations, number specifications must match:

The examples in (6) are ruled out because of a clash between the subject's number and the number required by the verb:

 $<sup>^2</sup>$ To avoid clutter, we omit most f-structure features other than number, including definiteness marking, person, gender, tense, aspect, and other grammatical features.

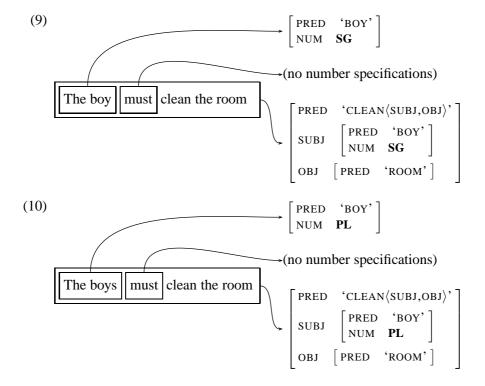
(7) Ill-formed f-structure with clashing number specification:



In English and many other languages, number is not always overtly specified. Many English verbs, including modals like *must*, do not impose person or number constraints on their subjects. In such cases, number may be specified only by the noun:

(8) The boy/boys must clean the room.





## 3 Indeterminacy, ambiguity, or underspecification?

For a small number of English nouns, singular and plural forms are not distinguished. When such nouns appear as subjects, the verb may provide the only indication of number:

(11) a. The *sheep/fish/deer is* moving quickly.

? s

b. The sheep/fish/deer are moving quickly.

' F

There are in principle several possibilities for the treatment of the number value of these nouns, according to standard LFG analyses of indeterminacy, ambiguity, and underspecification, as follows:

- Ambiguity: there are two homophonous nouns, sheep<sub>SG</sub>: [NUM SG], sheep<sub>PL</sub>: [NUM PL]
- Underspecification: No NUM value is specified for *sheep*; it can be specified by another component of the sentence, such as the determiner or verb, as either sg or PL.
- Indeterminacy: *sheep* is both singular and plural (in a sense to be defined in the following).

We begin by exploring and dismissing a treatment of nouns like *sheep* as having indeterminate number. A hallmark of indeterminacy is the ability to simultaneously satisfy conflicting requirements on the same feature (Dalrymple and Kaplan, 2000; Dalrymple, King and Sadler, 2006). The case feature for the Polish noun *kogo* 'who' is indeterminate, as shown by Dyła (1984), since it can simultaneously satisfy an ACC requirement and a GEN requirement:

```
(12) Kogo Janek lubi a Jerzy nienawidzi (Polish) who Janek likes and Jerzy hates
? OBJ CASE = ACC OBJ CASE = GEN
'Who does Janek like and Jerzy hate?' (Dyła, 1984)
```

This is also true for the case feature in German; as shown by Groos and van Reimsdijk (1979), was 'what' can simultaneously satisfy an ACC requirement and a NOM requirement:

```
(13) Ich habe gegessen was übrig war (German)

I have eaten what was left

OBJ CASE=ACC NOM/ACC SUBJ CASE=NOM

'I ate what was left.' (Groos and van Reimsdijk, 1979)
```

Formal analyses of indeterminacy have been proposed by Dalrymple and Kaplan (2000) and Dalrymple, King and Sadler (2006); though there are important differences between these analyses, they share the property that indeterminate features

have complex values which can simultaneously satisfy multiple conflicting constraints.

- (14) Specification of NOM/ACC case indeterminacy:
  - a. was according to Dalrymple and Kaplan (2000): CASE {NOM, ACC}
  - b. was according to Dalrymple, King and Sadler (2006):  $\begin{bmatrix} CASE & \begin{bmatrix} NOM & + \\ ACC & + \end{bmatrix} \end{bmatrix}$

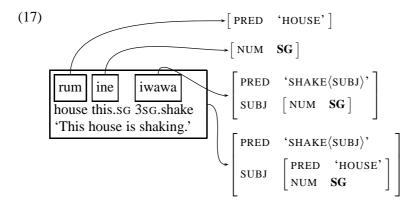
However, such analyses are inappropriate for the number feature, since the same form cannot satisfy singular and plural requirements at the same time:

Hence, the correct treatment for nouns like *sheep* must involve ambiguity or underspecification, not indeterminacy.

Although it is difficult to provide clear evidence as to whether English nouns like *sheep* are ambiguous or underspecified, in other languages it is clear that underspecification is the right treatment. Biak nouns are never marked for number, though number distinctions are obligatory elsewhere in the language; in the examples in (16), the demonstrative determiners and verbs show number agreement (Mofu, 2009).

- (16) a. rum ine iwawa house this.sg 3sg.shake 'This house is shaking.'
  - b. rum suine suwawa house this.DUAL 3DUAL.shake 'These two houses are shaking.'
  - c. rum skoine skowawa house this.PAUCAL 3PAUCAL.shake 'These (several) houses are shaking.'
  - d. rum nane nawawa house this.PL.INANIM 3PL.INANIM.shake 'These houses are shaking.'

Occam's razor precludes an analysis of all nouns in Biak as four ways ambiguous, with no morphological evidence for the ambiguity: Biak nouns are underspecified for NUM. As shown in (17), number specifications imposed by the determiner and the verb constrain the same feature, and these features must be compatible.



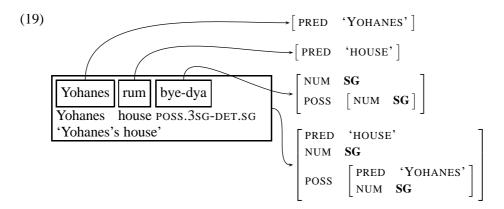
#### 4 Double indexing

In the examples we have seen so far, the verb specifies the required features of one of its dependents, its subject. In more complex cases, the same word can specify number constraints for more than one dependent. The Biak possessive construction illustrates this pattern (Mofu, 2009): the possessive determiner specifies the person and number of the possessor, and additionally specifies the number of the possessee head noun. An overt possessor may appear before the head noun, as in (18e).

- (18) a. roma bye-di son POSS.3SG-DET.SG 'his son'
  - b. roma bye-suya son POSS.3SG-DET.DUAL 'his two sons'
  - c. roma aye-di son POSS.1SG-DET.SG 'my son'
  - d. roma aye-skoi son POSS.1SG-DET.PAUCAL 'my (several) sons'
  - e. Yohanes rum bye-dya Yohanes house POSS.3SG-DET.SG 'Yohanes's house'

Such patterns are sometimes referred to as "double indexing".<sup>3</sup> The LFG analysis is straightforward; the doubly-agreeing word specifies the number of the possessor as well as the possessee:

<sup>&</sup>lt;sup>3</sup>This use of the term "double indexing" is different from its use in formal semantics, which refers to the representation of different kinds of referential dependencies by different kinds of indices that can appear on the same phrase (e.g. Heim 1993).



Palmer (2012, this volume) provides more discussion of double indexing in Oceanic languages.

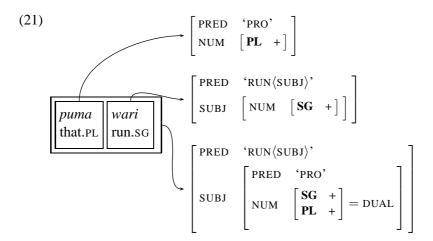
#### 5 Representing the number feature

In much LFG work, the number feature is assumed to have atomic values such as SG, DUAL, and PL, as in the examples above. More recent work on *constructed number* assumes that the value of the NUM feature is not atomic, but a complex value whose form may be constrained in different ways by different parts of the sentence. Sadler (2011) provides an analysis of the constructed dual in Hopi which assumes a complex value for the NUM feature.

Sadler (2011) presents the following data from Hopi, taken from Corbett (2000, 169), and notes that "dual number is expressed constructively through the combination of a plural pronoun and a verb showing sG agreement", as shown in (20c):

- (20) a. Pam wari that.sg run.perfect.sg 'S/he ran.'
  - b. Puma yùutu that.PL run.PERFECT.PL 'They ran.'
  - c. Puma wari that.PL run.PERFECT.SG 'They (two) ran.'

According to Sadler's analysis, number may be partially specified by different components of the sentence: singular verbs in Hopi contribute the specification [sg +] for the NUM feature, and plural pronouns contribute [PL +]. Dual number is defined as [sg +, PL +]:



A similar analysis for Marori, involving a nonatomic value for the NUM feature, is discussed below and by Arka (2012, this volume).

This analysis of constructed number may appear similar to Dalrymple, King and Sadler's treatment of feature indeterminacy, discussed above: in both analyses, a complex f-structure appears as the value of a feature (the value of the feature CASE for Dalrymple, King and Sadler 2006, NUM for Sadler 2011), with the possibility for more than one component of the complex structure to have the value '+' ([NOM +, ACC +] for Dalrymple, King and Sadler, [SG +, PL +] for Sadler). Importantly, however, the two analyses are in fact very different. For indeterminate features like CASE, a complex value such as [NOM +, ACC +] allows an indeterminate form to simultaneously fulfil **conflicting** case requirements. In contrast, Sadler's analysis does not entail that dual nouns are in any sense simultaneously SG and PL; instead, complete patterns of feature values holistically represent the kinds of values that are expressed as atoms in other analyses (SG, DUAL, PL). This shows that structures that are similar in appearance can be used to express very different linguistic intuitions.

#### 6 Syncretism and number

Recent LFG work has explored patterns of syncretism and underspecification in the morphology of number. We review this work here as an illustration of the general issue of the tradeoff in complexity between grammatical constraints and grammatical structures: often, similar phenomena can be analysed either in terms of relatively simple constraints on complex structures, or by complex statements of constraints on simple structures.

#### 6.1 Murrinh-Patha

Nordlinger (2011) presents an analysis of the morphology of number agreement on verbs in Murrinh-Patha (Australian), which we will use as an example of complex

constraints on simple structures. Nordlinger's analysis has been considerably simplified for the purposes of this discussion; for the complete analysis, see Nordlinger (2011) and, for more discussion of number marking in Murrinh-Patha, Nordlinger (2012, this volume).

A partial paradigm for the Murrinh-Patha verb meaning 'see' is shown in (22):

#### (22) a. bam-ngkardu

3sg-see

'He/she saw him/her.'

#### b. bam-ngintha-ngkardu

3sg-fem.dual-see

'They two (female non-siblings) saw him/her.'

#### c. pubamka-ngkardu

3DUAL-see

'They two (siblings) saw him/her.'

#### d. pubamka-ngkardu-ngime

3DUAL-see-FEM.PAUCAL

'They (paucal female nonsiblings) saw him/her.'

#### e. pubam-ngkardu

3<sub>PL</sub>-see

'They (paucal siblings/plural) saw him/her.'

A notable feature of this analysis is the reuse of forms in different and seemingly incompatible parts of the paradigm. As Nordlinger (2011) notes, an analysis involving accidental coincidence of form is unsatisfying; the patterns shown here are systematic.

(23)	bam	singular		
	bam + ngintha/nintha	dual non-sibling		
	pubamka	dual sibling		
	pubamka + ngime/neme	paucal non-sibling		
	pubam	paucal sibling		
	pubam	plural		

In her analysis of these forms, Nordlinger (2011) proposes to use the standard atomic values SG, DUAL, PAUCAL, and PL for the NUM feature. Crucially, the constraints associated with each form involve disjunction and the use of constraining equations to control the contribution of the morpheme combinations, as follows:

Consider, for example, the form bam-ngkardu:

(25) a. bam-ngkardu
3sG-see
'He/she saw him/her.'
b. bam: {(↑ subj num) = sG
| (↑ subj num) = c dual}

The constraints associated with *bam* can be paraphrased as follows: *bam* contributes the value SG for the number feature, or it appears in a context in which the value DUAL for the number feature is provided by another form. In (26) there is no other form to provide the value DUAL, so SG is correctly chosen:

(26) 
$$\begin{bmatrix} PRED & `SEE \langle SUBJ, OBJ \rangle ` \\ SUBJ & [NUM & \textbf{SG}] \end{bmatrix}$$

If the form *ngintha* is present, it contributes DUAL number, which is incompatible with a SG value:

(27) a. bam-ngintha-ngkardu
3sG-FEM.DUAL-see
'They two (female non-siblings) saw him/her.'

Since the sG specification for *bam* cannot be satisfied, the DUAL constraining equation for *bam* must be satisfied. The DUAL value contributed by *ngintha* satisfies the constraining equation, and the result is as in (28):

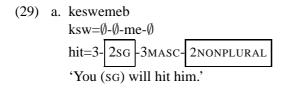
(28) 
$$\begin{bmatrix} PRED & SEE(SUBJ,OBJ) \\ SUBJ & SUBJ \\ SIB & - \end{bmatrix}$$

Nordlinger's analysis uses disjunctive constraints over atomic values such as SG, DUAL and PL to achieve simple and familiar f-structures for Murrinh-Patha verbs. For further discussion of Murrinh-Patha verb morphology, see Nordlinger (2012, this volume).

#### 6.2 Marori

Arka (2011) provides a discussion of verb morphology in Marori (isolate, Trans New Guinea) which is similar to Sadler's (2011) in using complex values for the NUM feature; the use of complex values allows a very simple statement of the contribution of different forms in the Marori verbal paradigm. Arka's analysis distinguishes singular, dual, paucal, and plural agreement through a verb-internal constructive strategy. Like the preceding analysis, the discussion of Marori presented in the following has been considerably simplified; see Arka (2011) for the full analysis, and Arka (2012, this volume) for more discussion of number marking in Marori.

Like Hopi, the Marori dual is formed as a combination of nonsingular and nonplural. Subject agreement in example (29a) is singular (singular and nonplural), (29b) is dual (nonsingular and nonplural), and (29c) is plural (nonsingular and plural):



# b. kesneme ksw=0-n-me-0 hit=3-2NONSING-3MASC-2NONPLURAL

'You (2) will hit him.

# c. kesnemem ksw=\( \Phi - n - me - \Phi \) hit=3-\( \frac{2 \text{NONSING}}{2 \text{NONSING}} - 3 \text{MASC-} \( \frac{2 \text{PL}}{2 \text{PL}} \) 'You (more than 2) will hit him.'

Constructed dual can also be expressed by a nonsingular subject with a nonplural verb:

Singular subjects take nonplural agreement:

A nonsingular pronominal form in combination with a plural verb means 'three or more' (nonsingular, nondual):

Like Sadler, Arka (2011) assumes that the value of the NUM feature is nonatomic; unlike Sadler, dual is treated as [SG -, PL -] rather than [SG +, PL +]. Arka's feature treatment for Marori is as in (33):

(33) Singular: 
$$\begin{bmatrix} SG & + \\ PL & - \end{bmatrix}$$
Dual: 
$$\begin{bmatrix} SG & - \\ PL & - \end{bmatrix}$$
Plural: 
$$\begin{bmatrix} SG & - \\ PL & + \end{bmatrix}$$

Crucial to Arka's analysis is that nonatomic features give rise to natural classes of features: singular and dual are [PL -], dual and plural are [SG -]. The form glossed 'singular' contributes [SG +], the nonsingular form is [SG -], plural is [PL +], and nonplural is [PL -]. This allows the following analyses:

(34) 'You (sg) will hit him': keswemeb ksw= $\emptyset$ - $\emptyset$ -me- $\emptyset$  hit=3-2sg-3masc-2nonplural [subj [sg +]] [subj [pl -]] = singular

(35) 'You (2) will hit him':

kesneme
ksw=∅-n-me-∅
hit=3-2NONSING-3MASC-2NONPLURAL

[SUBJ [SG -]] [SUBJ [PL -]] = dual

Rather than using disjunctive constraints on atomic features, Arka's analysis makes use of simple underspecified constraints on complex features: complex **structures** and simple **constraints**. It is not clear whether such a strictly compositional approach to constructed number is possible for all languages: Nordlinger (2012, this volume) presents data from Murrinh-Patha that are difficult to analyse on an approach involving complex values and underspecification. Arka (2012, this volume) provides additional discussion of number marking and verbal number in Marori.

#### 7 Inverse number

We conclude with a brief description of two phenomena for which there are no standard or established analyses within LFG, but which must be addressed in a complete analysis of the morphology, syntax, and semantics of number.

In languages with **inverse number**, different classes of nouns have different 'unmarked' or default values for the number feature, and inverse morphological marking changes the number value to a 'marked' value; see Corbett (2000) for more discussion. A notorious case of inverse number is found in Kiowa (Kiowa-Tanoan; Watkins 1984; Harbour 2008), and shown in (36), excerpted from Harbour (2008, Chapter 2). Nouns in Kiowa fall into a number of classes, three of which are illustrated here. In the class represented by the noun 'young man', the unsuffixed noun is singular or dual, and nouns with the inverse number suffix  $d\acute{o}$  are plural. In the class represented by 'tomato', the unsuffixed noun is dual, and the suffixed form is singular or plural; in the class represented by 'tree', the unsuffixed noun is dual or plural, and the suffixed noun is singular. Such patterns pose interesting questions for the morphology-syntax-semantics interface.

More discussion of inverse number in Oceanic is provided by Palmer (2012, this volume).

#### 8 Verbal number

Verbal number marking indicates that an event took place a number of times or that an event had a number of (usually theme) participants. Veselinova (2006) illustrates the phenomenon with examples from Mupun (Chadic; Frajzyngier 1993; Veselinova 2006). The Mupun verb meaning 'kill' has two forms: one for a single event, and one for multiple killing events:

```
(37) a. tù: 'kill, singular action'b. tù-é: 'kill, plural action'
```

The form in (37a) is used for one killing event, and the form in (37b) is used for several events. Since a rat can be killed only once, in (37b) several rats must have been involved; plural marking on the noun meaning 'rat' is possible but not necessary.

Suppletive forms are also found:

```
(39) a. cīt: 'beat, singular action'b. nás: 'beat, plural action'
```

The singular-action form cannot be used with a plural object, as shown in example (40a). Example (40b) shows that it is possible to use the plural-action form with a singular object to describe multiple beating events involving the same individual:

Arka (2012, this volume) discusses verbal number in Marori, which involves multiple exponence of the kind discussed in Section 6.

#### 9 Conclusion

LFG provides a solid theoretical basis for work exploring the syntax and semantics of number, but a good deal of work remains to be done. Besides a complete treatment of inverse number and verbal number, described in the final two sections of this overview paper, unexplored or underexplored areas include the following areas:

- the determination and specification of number for coordinate structures, including structures with singular number, such as "my friend and colleague", as well as structures with plural number, like "George and Fred" and "the dog and cat"
- the analysis of nouns which seem to exhibit both singular and plural properties, such as British English "company", which require singular determiners but can appear with plural verbs ("this company are well managed"): see Hristov (2012) for interesting and illuminating discussion of these examples, and proposals for their analysis from an LFG perspective
- the formal representation of the number feature, particularly for languages whose number systems include dual, trial, or paucal in addition to singular and plural
- related to the general issue of the representation of the number feature, the
  question of whether the number feature should be treated as privative, with
  one of the values of the feature represented as the absence of a value for the
  feature

The papers presented in the Workshop provide a firm basis for the exploration of these and other issues.

#### Acknowledgments

Thanks to I Wayan Arka for proposing and organising the Workshop on Number at LFG12, to the workshop participants I Wayan Arka, Rachel Nordlinger, and Bill Palmer, and to the audience at LFG12 for stimulating discussion of these issues. And as always, thanks to Miriam Butt and Tracy King for helpful comments on an earlier draft of this paper, and for their usual exemplary efficiency in preparing the conference proceedings for publication.

#### References

Arka, I Wayan. 2012. Verbal Number, Argument Number and Plural Events in Marori. In *On-line Proceedings of the LFG2012 Conference*.

- Arka, Wayan. 2011. Constructive Number Systems in Marori and Beyond. In Miriam Butt and Tracy Holloway King (eds.), *On-line Proceedings of the LFG2011 Conference*.
- Corbett, Greville G. 2000. *Number*. Cambridge, UK: Cambridge University Press. Dalrymple, Mary and Kaplan, Ronald M. 2000. Feature Indeterminacy and Feature Resolution. *Language* 76(4), 759–798.
- Dalrymple, Mary, King, Tracy Holloway and Sadler, Louisa. 2006. Indeterminacy by Underspecification. In Miriam Butt and Tracy Holloway King (eds.), *On-line Proceedings of the LFG2006 Conference*.
- Dixon, Robert M. W. 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Dyła, Stefan. 1984. Across-The-Board Dependencies and Case in Polish. *Linguistic Inquiry* 15(4), 701–705.
- Frajzyngier, Zygmunt. 1993. *A Grammar of Mupun*. Berlin: Dietrich Reimer Verlag.
- Groos, Anneke and van Reimsdijk, Henk. 1979. Matching Effects in Free Relatives: A Parameter of Core Grammar. In Adriana Belletti, Luciana Brandi and Luigi Rizzi (eds.), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, pages 171–216, Pisa: Scuola Normale Superiore di Pisa.
- Harbour, Daniel. 2008. *Morphosemantic Number: From Kiowa Noun Classes to UG Number Features*. Studies in Natural Language and Linguistic Theory, Springer.
- Heim, Irene. 1993. Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach. Technical Report, University of Tübingen, reprinted in Sauerland and Percus (1998, 205–246).
- Hristov, Bozhil. 2012. *Agreement, Case Assignment, and Nominal Coordination*. D.Phil. dissertation, University of Oxford.
- Kibort, Anna and Corbett, Greville G. 2008. Number. In *Grammatical Features*, University of Surrey.
- Mofu, Suriel. 2009. Biak Morphosyntax. D.Phil. dissertation, Oxford University.
- Nordlinger, Rachel. 2011. LFG and Language Documentation. Abstract in *On-line Proceedings of the LFG2011 Conference*, ed. Miriam Butt and Tracy Holloway King.
- Nordlinger, Rachel. 2012. Number Marking in the Daly River Languages. In *Online Proceedings of the LFG2012 Conference*.
- Palmer, Bill. 2012. Nominal Number in Meso-Melanesian. In *On-line Proceedings* of the LFG2012 Conference.
- Sadler, Louisa. 2011. Indeterminacy, Complex Features and Underspecification. *Morphology* 21, 379–417.
- Sauerland, Uli and Percus, Orin (eds.). 1998. *The Interpretive Tract*. MITWPL, Department of Linguistics and Philosophy, MIT.
- Steinhauer, Hein. 1985. Number in Biak: Counterevidence to Two Alleged Language Universals. *Bijdragen tot de Taal-, Land- en Volkenkunde* 141, 462–485.

Veselinova, Ljuba N. 2006. Verbal Number and Suppletion. In Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie (eds.), *The World Atlas of Language Structures*, Oxford: Oxford University Press.

Watkins, Laurel. 1984. A Grammar of Kiowa. Lincoln: University of Nebraska Press.