

CHALLENGES AT THE  
SYNTAX-SEMANTICS-PRAGMATICS  
INTERFACE

*A Role and Reference Grammar Perspective*

EDITED BY  
Robert D. Van Valin Jr.

# Challenges at the Syntax-Semantics- Pragmatics Interface



# Challenges at the Syntax-Semantics- Pragmatics Interface:

*A Role and Reference Grammar  
Perspective*

Edited by

Robert D. Van Valin, Jr.

**Cambridge  
Scholars  
Publishing**



**Challenges at the Syntax-Semantics-Pragmatics Interface:  
A Role and Reference Grammar Perspective**

Edited by **Robert D. Van Valin, Jr.**

This book first published 2021

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2021 by **Robert D. Van Valin, Jr. and contributors**

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-6747-8

ISBN (13): 978-1-5275-6747-4

# TABLE OF CONTENTS

Contributors .....	vii
Editor's Introduction .....	xi
<b>Part 1. Lexical Representation and Argument Structure</b>	
Activities, Accomplishments and Causation .....	3
<i>Rainer Osswald</i>	
Simplex and Complex Predicates in Persian: An RRG Analysis .....	31
<i>Jens Fleischhauer</i>	
Linking Syntax and Semantics in Comparatives of Taiwan Sign Language .....	63
<i>Chien-hung Lin &amp; Jung-hsing Chang</i>	
<b>Part 2. Argument Marking</b>	
Dative Case and Three-Place Predicates in Japanese .....	91
<i>Hideki Kishimoto</i>	
Purposive Case and Semantic Arguments in Australian Western Desert Dialects .....	113
<i>Conor Pyle</i>	
A Neo-Jakobsonian Account of Default Cases: Instrumental vs. Dative .....	135
<i>Wataru Nakamura</i>	
A Cross-Linguistic Survey of the Instrument-Subject Alternation .....	169
<i>Koen Van Hooste</i>	
Two-Theme Constructions and Preposition Assignment in Spanish .....	189
<i>Sergio Ibáñez Cerda</i>	

### Part 3. Syntactic Structure

An Analysis of Non-Iconic Word Order in the Bamunka Reference Phrase .....	213
<i>Ciara Anderson</i>	
Cosubordination .....	241
<i>Robert D. Van Valin, Jr.</i>	

### Part 4. Information Structure

An RRG Account of Aspects of the Information Structure-Syntax Interface in Tagalog .....	257
<i>Anja Latrouite &amp; Robert D. Van Valin, Jr.</i>	
Why Eve Shouldn't Eat the Snake: An Intelligent Answer from Corpus-driven Information Structure and Reference Tracking in Biblical Hebrew .....	285
<i>Nicolai Winther-Nielsen</i>	
Language Index .....	309
Subject Index .....	311

## CONTRIBUTORS

Ciara Anderson  
Centre for Language and Communication Studies  
Room 4091, Arts Building  
Trinity College  
Dublin 2,  
Ireland  
ciarabrowne86@gmail.com

Jung-hsing Chang  
Institute of Linguistics  
National Chung Cheng University  
168 University Road, Min-Hsiung  
Chia-Yi, Taiwan 62102, R.O.C.  
lngjhc@gmail.com

Jens Fleischhauer  
Department of General Linguistics  
Heinrich Heine University Düsseldorf  
Universitätsstr. 1  
40225 Düsseldorf  
Germany  
fleischhauer@phil.uni-duesseldorf.de

Sergio Ibáñez Cerdá  
Instituto de Investigaciones Filológicas  
Universidad Nacional Autónoma de México (UNAM)  
Tenorios 298, Edificio Ciprés "A", No. 202,  
Colonia Hacienda Coapa, Delegación Tlalpan, México, D.F., México.  
C.P. 14300  
cecser@prodigy.net.mx



Hideki Kishimoto  
Department of Linguistics  
Graduate School of Humanities  
Kobe University  
1-1 Rokkodai-cho, Nada-ku  
Kobe, 657-8501  
Japan  
kishimoto@lit.kobe-u.ac.jp

Anja Latrouite  
Department of General Linguistics  
Heinrich Heine University Düsseldorf  
Universitätsstr. 1  
40225 Düsseldorf  
Germany  
latrouite@phil.uni-duesseldorf.de

Chien-hung Lin  
International Graduate Program of Teaching Chinese as a Second  
Language  
Tunghai University  
No. 1727, Sec. 4, Taiwan Boulevard, Xitun District  
Taichung City 407224, Taiwan R.O.C.  
lngjhl@gmail.com

Wataru Nakamura  
Center for the Advancement of Higher Education and  
Graduate School of International Cultural Studies  
Tohoku University  
41 Kawauchi  
Aoba-ku, Sendai, Miyagi  
980-8576, Japan  
wnmtohoku@gmail.com

Rainer Osswald  
Department of Computer Linguistics  
Heinrich Heine University Düsseldorf  
Universitätsstr. 1  
40225 Düsseldorf  
Germany  
osswald@phil.hhu.de

Conor Pyle  
Centre for Language and Communication Studies,  
School of Linguistic, Speech and Communication Sciences,  
Trinity College Dublin,  
Dublin 2,  
Ireland  
pylec@tcd.ie

Koen Van Hooste  
Department of General Linguistics  
Heinrich Heine University Düsseldorf  
Universitätsstr. 1  
40225 Düsseldorf  
Germany  
koen.vanhooste@gmail.com

Robert D. Van Valin, Jr.  
Department of Linguistics  
University at Buffalo  
The State University of New York  
Buffalo, NY 14260 USA  
vanvalin@buffalo.edu

Nicolai Winther-Nielsen  
Hebrew Bible and Information and Communication Technology  
Fjellhaug International University College Copenhagen  
Leifsgade 33,6. DK-2300 Copenhagen S  
Denmark  
nwn@dbi.edu



# EDITOR'S INTRODUCTION

ROBERT D. VAN VALIN, JR.

*Challenges at the syntax-semantics-pragmatics interface: A Role and Reference Grammar perspective* addresses important issues in syntactic theory from the point of view of a parallel architecture theory which looks at grammar as reflecting the complex interaction of structure, meaning, and function. Role and Reference Grammar [RRG] is a monostatal, non-derivational theory which posits a linking algorithm which maps between a syntactic representation and a semantic representation, and discourse-pragmatics plays a role in the mapping.<sup>1</sup> This is summarized in Figure 1.

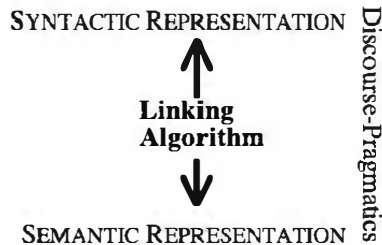


Figure 1: The organization of Role and Reference Grammar

The syntactic representation is concrete, in the sense that it represents the words and morphemes in their actual sequence and does not allow phonologically null elements, in particular no traces/unpronounced copies of words or phrases, null light verbs, null pronominals, null adpositions, etc. The semantic representation employs an *Aktionsart*-based system of lexi-

---

<sup>1</sup> For detailed introductions to RRG, see Van Valin and LaPolla (1997), Van Valin (2005), Pavey (2010), Mairal Usón, et al. (2012), Bentley, et al. (2022). For overviews of the theory, see Van Valin (2010, 2014, 2022a), Wu (2007), González Vergara (2006).

cal decomposition. The linking algorithm is bidirectional, mapping from semantics to syntax and from syntax to semantics. This reflects the fact that in language production speakers go from meaning to form to utterance, whereas in comprehension listeners go from utterance to form to meaning.<sup>2</sup> Discourse-pragmatics interacts with the linking algorithm in a variety of ways, and the interaction varies across languages, leading to significant differences in grammatical systems.

RRG has always been strongly cross-linguistically oriented, and this is a reflection of the 'founding question' that got work on it started back in the late 1970's: what would linguistic theory look like if it were based on languages with diverse structures, specifically Tagalog (Austronesian, Philippines), Dyirbal (Pama-Nyungen, Australia) and Lakhota (Siouan, North America), instead of English? This cross-linguistic orientation can be seen in the range of languages discussed in the papers in this volume, including Bamunka (Grassfields Bantu), Biblical Hebrew, Japanese, Persian, Pitjantjatjara (Pama-Nyungen, Australia), Russian, Tagalog, and Taiwan Sign Language. There is a comparably wide range of theoretical issues addressed, including verb classification and decomposition, light verb constructions, the instrument-subject construction, comparative constructions, case marking, adposition assignment, inversion constructions, clause linkage, and reference tracking. The contributions can be organized into four groups: (1) lexical representation and argument structure, (2) argument marking, (3) syntactic structure, and (4) information structure.

The first group addresses issues related to lexical representation and interesting cases of argument structure. In 'Activities, accomplishments and causation' Rainer Osswald explores the evolution of the RRG system of lexical decomposition in relation to the different proposals put forth in Dowty (1979), which is the basis of the original RRG system proposed in Foley and Van Valin (1984). The other two papers in the section deal with argument structure phenomena which deserve more attention in RRG. Jens Fleischhauer investigates light verb constructions in Persian, a language with a rich inventory of them, in 'Simplex and complex predicates in Persian – An RRG analysis'. Comparative constructions have not been analyzed from an RRG perspective, and Chien-hung Lin and Jung-hsing Chang present an RRG account of comparative constructions in Taiwan Sign Language in 'Linking syntax and semantics in comparatives of Tai-

---

<sup>2</sup> See Van Valin (2006, 2022b) for discussion of RRG as a component of a model of sentence processing.

wan Sign Language'. The emphasis is on their argument structure and the linking from semantics to syntax.

The second group of papers concerns the case and adpositional coding of direct and oblique arguments, focussing on dative and instrumental cases, the corresponding adpositions, and the syntactic expression of instrument arguments. Hideki Kishimoto investigates the contrast between dative case and postpositional marking of the third argument of ditransitive verbs in 'Dative case and three-place predicates in Japanese'. RRG makes an important distinction between *predicative* and *non-predicative* adpositions, i.e. adpositions that are assigned by rule (non-predicative) vs. those that contribute semantically to the clause and are part of the semantic representation (predicative); see Van Valin (2005:21-23, 49). The next two papers show that morphological cases, which are normally non-predicative, can also be predicative as well. The Western Desert language in central Australia has a dative-like case called the *purposive*, and it is the topic of Conor Pyle's contribution 'Purposive case and semantic arguments in Australian Western Desert dialects'. Wataru Nakamura examines RRG's claim that dative is the default case for non-macrorole direct core arguments in light of Jakobson's analysis of the instrumental as being the least marked case in Russian and reaches some interesting conclusions in his paper 'A Neo-Jakobsonian Account of Default Cases: Instrumental vs. Dative'. He also looks at the use of the English prepositions *with* and *by* in comparison with the Russian instrumental case. The next contribution, Koen Van Hooste's 'A Cross-Linguistic Survey of the Instrument-Subject Alternation', deals with variation in the morphosyntactic expression of instrument arguments, specifically with the possibility of expressing it as the subject, e.g. *The knife cut the bread* vs. *She cut the bread with the knife* in contrast to *\*The spoon ate the soup* vs. *She ate the soup with the spoon*, across a range of languages. The final paper of this section, 'Two-Theme Constructions and Preposition Assignment in Spanish' by Sergio Ibáñez Cerda, presents a contrastive analysis of the preposition assignment rules proposed for English in RRG with those posited for Spanish three-place predicates taking two theme arguments.

The third group of papers concerns aspects of syntactic structure. In her paper 'An Analysis of Non-Iconic Word Order in the Bamunka Reference Phrase', Ciara Anderson examines the structure of what are traditionally called 'noun phrases' in Bamunka, a Grassfields Bantu language spoken in Cameroon, in light of certain word order universals, giving an RRG account. The second paper concerns an important concept in the RRG theory of complex sentences, cosubordination. This notion has been subject to criticism, and there have been suggestions that it is not a valid construct.

Robert D. Van Valin, Jr. defends the concept, showing that the arguments against it lack validity in his paper 'Cosubordination'.

The final group investigates the syntax-pragmatics interface, commonly referred to as *information structure*. Anja Latrouite and Robert D. Van Valin, Jr. give an account of the rich inventory of inversion constructions in Tagalog in RRG terms in their contribution 'An RRG Account of Aspects of the Information Structure-Syntax Interface in Tagalog'. It reveals some unexpected interactions between syntactic positions and information structure. The final article in the volume concerns reference tracking in Biblical Hebrew: 'Why Eve Shouldn't Eat the Snake: An Intelligent Answer from Corpus-driven Information Structure and Reference Tracking in Biblical Hebrew' by Nicolai Winther-Nielsen. The author employs the discourse representations used in RRG as a crucial component of his analysis of the interpretation of potentially ambiguous utterances in context.

## References

- Bentley, Delia, Mairal Usón, Ricardo, Nakamura, Wataru & Van Valin, Robert D., Jr. (eds). 2022. *Cambridge handbook of Role and Reference Grammar*. Cambridge: Cambridge University Press.
- Dowty, David. 1979. *Word meaning and Montague Grammar*. Dordrecht: Reidel.
- Foley, William A. & Van Valin, Robert D. Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- González Vergara, Carlos. 2006. La Gramática del Papel y la Referencia: una aproximación al modelo. *Onomázein* 14: 101-140.
- Mairal Usón, Ricardo, Guerrero, Lilián & González Vergara, Carlos (eds.). 2012. *El funcionalismo en la teoría lingüística. La Gramática del Papel y la Referencia. Introducción, avances y aplicaciones*. Akal: Madrid.
- Pavey, Emma. 2010. *The structure of language: An introduction to grammatical analysis*. Cambridge: Cambridge University Press.
- Van Valin, Robert D. Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- . 2006. Semantic macroroles and language processing. In: Bornkessel-Schlesewsky, Ina, Schlewsky, Matthias, Comrie, Bernard & Friederici, Angela (eds.). *Semantic role universals and argument linking: Theoretical, typological and psycho-/neurolinguistic perspectives*, 263-302. Berlin: Mouton de Gruyter.

- 2010. Role and Reference Grammar as a framework for linguistic analysis. In: Heine, Bernd & Narrog, Heiko (eds.), *The Oxford handbook of linguistic analysis*, 703-38. Oxford: Oxford University Press.
  - 2014. Role and Reference Grammar. In: Carnie, Andrew, Siddiqi, Dan & Sato, Yosuke (eds.), *Routledge handbook of syntax*, 579-603. London: Routledge.
  - 2022a. Principles of Role and Reference Grammar. In Bentley, Delia et al. (eds.), in preparation.
  - 2022b. Grammatical aspects of language comprehension in the brain: A Role and Reference Grammar perspective. In Bentley, Delia et al. (eds.), in preparation.
- Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
- Wu, Joy. 2007. An introduction to Role and Reference Grammar [in Mandarin]. *Language and Linguistics*, 8: 5-70.





# **PART 1.**

## **LEXICAL REPRESENTATION AND ARGUMENT STRUCTURE**



# ACTIVITIES, ACCOMPLISHMENTS AND CAUSATION

RAINER OSSWALD

HEINRICH-HEINE-UNIVERSITÄT DÜSSELDORF

## Abstract

The *Aktionsart* system is a crucial component of the syntax-semantics interface of Role and Reference Grammar (RRG). The purpose of this paper is two-fold: (i) to review the development of RRG's *Aktionsart* system since its adaptation from Dowty (1979/1991) by Foley & Van Valin (1984) and to relate the various revisions introduced over the years to issues already mentioned in Dowty (1979/1991); (ii) to point out some issues in the latest formulation of the *Aktionsart* system in Van Valin (2018) and to sketch how to overcome them by a more expressive decomposition system.

## Keywords

Aktionsart, semantic decomposition, event semantics, telicity, causation

## 1. Introduction

The *Aktionsart* system and its representation by means of logical structures is one of the cornerstones of the syntax-semantics interface of Role and Reference Grammar (RRG). The logical structures are a core component of the semantic representation of a sentence, and the positions of the arguments in these structures play a crucial role for the realization of the arguments in the syntax as determined by the linking system. The system of *Aktionsart* classes, as well as the system of logical structures in general, has evolved considerably over the years since the publication of Foley & Van Valin (1984), which was largely based on Dowty's (1979/1991) adaptation of ideas from Generative Semantics.

The present paper has a two-fold objective: First, we will take a look at the issues of Dowty's proposal, many of which have been pointed out by Dowty himself, and show how they have been addressed by the various modifications of the decomposition system of RRG. This includes, among other things, the

Table 1: The early *Aktionsart* system of RRG

<i>Aktionsart</i> class	Logical structure
State	$P(x)$ or $P(x, y)$
Activity	$\mathbf{DO}(x, [P(x)])$ or $\mathbf{DO}(x, [P(x, y)])$ (agentive) $P(x)$ or $P(x, y)$ (non-agentive)
Achievement	$\mathbf{BECOME} P(x)$ or $\mathbf{BECOME} P(x, y)$
Accomplishment	$\phi$ CAUSE $\psi$ , with $\phi$ typically an activity structure and $\psi$ typically an achievement structure

independence of causation from telicity as a consequence of introducing active accomplishments, and the more recent decomposition of (simple) accomplishments into a process and a result component. The second objective of the paper is to address issues of the current decomposition system and to suggest possible ways of how to overcome them. In particular, it will prove useful to clarify the denotational meaning of the logical structures, that is, to spell out the reference to states and events more explicitly. To this end, a decompositional system will be proposed which builds crucially on attributes and which is closely related to frame-based representations.

## 2. The early *Aktionsart* system and Dowty's heritage

### 2.1. The early *Aktionsart* system

The original version of the RRG *Aktionsart* system proposed in Foley & Van Valin (1984) uses the four classes *state*, *activity*, *achievement*, and *accomplishment*, much in line with Dowty's adaptation of Vendler's (1957) classification.<sup>1</sup> Table 1 reviews the original system as summarized in Van Valin (1990: 224, 1993: 36). It is characteristic of the early system that accomplishments are consistently analyzed as causative, i.e., as having the logical structure  $\phi$  CAUSE  $\psi$ , in accordance with Dowty's aspect calculus (cf. Dowty 1979/1991: 91/124f). Achievements, on the other hand, are assumed to have the general form  $\mathbf{BECOME} \phi$ .

Note that the distinction between agentive and non-agentive activity shown in Table 1 is not strictly part of the calculus introduced in Chapter 2 of Dowty's

<sup>1</sup> The following brief review of the development of RRG's *Aktionsart* system draws considerably on Van Valin (2018).

book but is adapted from a revised classification sketched in a later chapter of that book (cf. Dowty 1979/1991: 184). The operator **DO** is in fact discussed rather controversially by Dowty with respect to its intended use for expressing agentivity, controllability and/or intentionality (cf. Section 2.2 below), and, for similar reasons, the use of **DO** has been strongly restricted in later versions of the RRG system (cf. Section 2.3). The early system represents unintentional/uncontrolled activities like falling by one-place predicates such as **fall'**, and it also introduces a generalized activity predicate **do'** which is unmarked for agency (Foley & Van Valin 1984: 53).

A remark on notation: The boldface+prime markup for primitive predicates like **fall'** was introduced in Foley & Van Valin (1984) and has been henceforth in use in the RRG literature. The notation was taken from Dowty (1979/1991), who uses boldface for words and employs the prime ' to indicate the translation from natural language into a logical language. So, for Dowty, **fall** is an English word while **fall'** is a predicate. Since the boldface type does not serve any markup purposes in the standard presentation of RRG, we can simply declare **fall** to be the predicate that (roughly) represents the meaning of the English verb *fall*. For reasons of notational parsimony we will therefore write predicates in boldface without an additional prime in the rest of this paper.

## 2.2. A short (de)tour through Dowty 1979

It is worthwhile to recall the general motivation behind Dowty's 1979/1991 book. His primary goal was to combine theories of word meaning with Montague's formalized compositional framework of the syntax-semantics interface. The chosen approach to word meaning was inspired by the decomposition analyses (in terms of **DO**, **CAUSE**, **BECOME**, etc.) proposed in the Generative Semantics literature, which he regarded as "the most highly structured version of decomposition analysis" available (Dowty 1979/1991: vi). The respective analyses go back to his dissertation, which was completed in 1972 when Generative Semantics was still flourishing. It was thus already in this thesis where he developed the formal *Aktionsart* classifications and the representations that served as the basis of the semantic structures of the early RRG system shown in Table 1.

While having included the decomposition analyses of his dissertation in the 1979 book, Dowty did not hold back with pointing out the many issues he meanwhile saw in this approach, and with sketching alternative solutions. He also proposed a second approach, *interval semantics*, for modeling the se-

manics of aspect (which was first published in Dowty 1977). Nevertheless, as Dowty (1979/1991: xix) emphasizes in the new preface to the second printing, the decompositional approach and the interval semantics account are not incompatible but can be fruitfully combined with each other, at least in principle, since they contribute to complementary domains: The decomposition approach is concerned with developing linguistically and cognitively motivated formalized representations for expressing aspectual and *Aktionsart* distinctions. Interval semantics, on the other hand, is concerned with the interpretation of such structures in the temporal domain, i.e., with respect to intervals, subintervals, lower and upper bounds of intervals, etc. That is, interval semantics is not to be seen as an alternative to decompositional semantics but as an attempt to spell out how the decompositional primitives are to be interpreted with respect to the time course or “dynamic contour” of the denoted events.

Among the operators **DO**, CAUSE and BECOME, only BECOME with its underlying change-of-state interpretation turns out as being useful in this respect. Roughly speaking, BECOME  $\phi$  is true at an interval  $I$  if  $\phi$  is not true at (an interval containing) the initial bound of  $I$  and  $\phi$  is true at (an interval containing) the final bound of  $I$ . It is far less obvious how to evaluate CAUSE within an interval semantics (see the discussion in Dowty 1979/1991: 191, footnote 17). And the notion of intentionality that is supposed to go along with **DO** is even more remote in this respect. In fact, Dowty (1979/1991: 121) casts serious doubts on the viability of taking **DO** as an indicator of intentionality and/or controllability and, moreover, finds it “doubtful that **DO** can really distinguish all activities from statives, after all” (Dowty 1979/1991: 119). Hence, while a decompositional analysis on the one hand and an interpretation of *Aktionsart* and aspect in the temporal dimension on the other hand are compatible with, and most probably even dependent on each other, it is questionable that the triad **DO**, CAUSE, BECOME is appropriate for this purpose. Moreover, the interval-based semantics calls for further refinements as well, as will be discussed in what follows.

In the new preface, Dowty (1979/1991: xxii) notes that the formal interpretation of BECOME in his interval semantics has the defect that “when an event of change takes place over an interval of time, the change in some sense does not »take place« until the end of the interval.” For example, the incremental change which goes along with an event as expressed by *paint the house red* is not captured by his treatment of BECOME. Dowty points to the work of Krifka and others who model the incremental character of such events in terms of an object-to-event homomorphism, and he more generally espouses the idea of developing his interval-based account into an event-based account

since the latter “is conceptually simpler, easier to formalize, and has substantive advantages” (Dowty 1979/1991: xxii).

### 2.3. Minor revisions of RRG’s original *Aktionsart* system

The original adaptation of Dowty’s system remained basically unchanged in Van Valin (1990, 1993) except for the following three minor qualifications: (i) The operator **DO** was restricted to the indication of agency as part of the lexical meaning of a verb as it is the case for *murder* in contrast to *kill*; see Van Valin & Wilkins (1996: 307ff) for arguments for restricting **DO** to lexicalized verbal agency. As mentioned in Section 2.2, there are also qualms about the correct interpretation of **DO** in Dowty (1979/1991). In particular, he notes that **DO** behaves like an *implicative verb* in the sense of Karttunen (1971) in that it does not affect the truth conditions of the embedded expression. Dowty (1979/1991: 118) concludes that the “contribution to meaning that **DO** makes is entirely in its conventional implicature.” RRG’s abandonment of **DO** as a general indicator of agency had the consequence that the argument of the generalized activity predicate **do** (cf. Section 2.1) may now be an agent or a pure effector, depending on the context.

(ii) The need for a distinction between punctual and durative achievements is briefly acknowledged (Van Valin 1993: 154, footnote 19). Not all achievements are punctual as attested by verbs such as *cool* and *dry*, which can be used in the progressive and even with durational adverbials. Dowty (1979/1991: 88ff) speaks of *degree achievements* in such cases but evades a thorough formal representation of them in his decomposition system. Likewise, **BECOME** was kept in Van Valin (1993) both for punctual and durative achievements.

(iii) The argument order in predicates of type **be-LOC** is reversed. While in Foley & Van Valin (1984), in line with Dowty (1979/1991: 210ff), **be-at**(*x*, *y*) means that *x* is (located) at *y*, the converse interpretation convention was established in Van Valin (1990): *x* is the location where *y* is located. This modification was obviously motivated by the desire of having a parallel representation of locative and possession predicates, which allows one to characterize the first argument of these predicates as “locative”, or, to put it differently, to have uniform argument positions available for the default macro-role assignment principles. From this perspective, the switch of arguments in the **be-LOC** predicates is mainly theory-driven. On the other hand, alienable possession is in many languages expressed by a locative construction (Foley & Van Valin 1984: 48; see also Stassen 2009: 48ff).



Table 2: Simple *Aktionsart* classes (textbook version)

<i>Aktionsart</i> class	Logical structure
State	$P(x)$ or $P(x, y)$
Activity	$\mathbf{do}(x, [P(x)])$ or $\mathbf{do}(x, [P(x, y)])$
Semelfactive	SEML $\phi$
Achievement	INGR $\phi$ with $\phi$ a state or activity structure
Accomplishment	BECOME $\phi$

### 3. The textbook version and recent developments

What we call the “textbook version” of the *Aktionsart* classes and their logical structures is the version presented in Van Valin (2005, 2010). The textbook version coincides basically with the representation system introduced in Van Valin & LaPolla (1997), except for the semelfactive operator SEML, which has been added later, and a number of minor modifications to be mentioned below. The *Aktionsart* classes of the textbook version are summarized in Tables 2 and 3.

#### 3.1. Simple *Aktionsart* classes

The *Aktionsart* classes listed in Table 2 are called *simple* since the corresponding logical structures are either primitive predicates representing states, or primitive predicates enclosed by  $\mathbf{do}(x, [\dots])$  representing activities, or expressions that result from applying one of the unary operators SEML, INGR or BECOME to a state or activity structure. (Due to lack of space, we put aside a discussion of the SEML operator.) The operator INGR (indicating *ingression*) has been introduced to make explicit the distinction between punctual and non-punctual change-of-state verbs (cf. Section 2.3). INGR is used for punctual verbs such as *shatter* and *explode* while BECOME is reserved for non-punctual, incremental verbs such as *melt* and *dry*. The members of the latter class are now referred to as *accomplishments*, in line with the original terminology of Vendler (1957), and in contrast to Dowty’s restriction of this term to causative structures (cf. Section 2.1).

The representation of activities follows basically Van Valin & Wilkins (1996), who show that agency is best understood as being derived from a number of interacting morphosyntactic, semantic and pragmatic factors, and that

Table 3: Complex *Aktionsart* classes (textbook version)

<i>Aktionsart</i> class	Logical structure
Active accomplishment	$\text{do}(x, [P(x, y)]) \& \text{INGR } \bullet(y)$ or $\text{do}(x, [P(x)]) \& \text{INGR } \bullet(y, x)$
Causative	$\phi \text{ CAUSE } \psi$ , with $\phi, \psi$ logical structures of any type

it is hence inadequate to anchor agency solely to a primitive operator **DO** in the semantic representation of the verb. The operator **DO** is therefore restricted to the representation of *lexicalized* agency (cf. Section 2.3) while the predicate **do** is assumed to be underspecified with respect to agency. However, instead of using **do** as a simple one-place predicate for activities, on a par with more specific activity predicates such as **walk**, Van Valin & Wilkins (1996) keep the decomposition structure of activities introduced by Ross (1972) and employed by Dowty (1979/1991), with **DO** now replaced by **do**. That is, an activity verb such as *walk* is represented by  $\text{do}(x, [\text{walk}(x)])$  and not just simply by  $\text{walk}(x)$ . This leads to the general semantic structure  $\text{do}(x, [P(x)])$  (or  $\text{do}(x, [P(x, y)])$ ) shown in Table 2, with the special case of  $\text{do}(x, \bullet)$  for unspecified activities. This way of representing activities apparently comes in handy for determining the first argument of an activity predicate (i.e., the effector) by the linking system since it can be directly read off from the logical structure. The semantic contribution of **do**, however, seems to be void, except for the case of  $\text{do}(x, \bullet)$ . We will return to this issue in Section 4.1 below.

### 3.2. Active vs. causative accomplishments

The most prominent change in the *Aktionsart* system of Van Valin & LaPolla (1997) is the introduction of *active accomplishments* as a separate class. The vast majority of this class consists of accomplishment uses of activity verbs. The examples in (1) provide a paradigmatic set of such uses in English.

- (1) a. Adam ate an apple. (consumption)  
 b. Mary drew a circle. (creation)  
 c. Kim walked to the store. (motion to goal)

According to the original decomposition system of Table 1, all of these accomplishment constructions are to be analyzed as causative (cf., e.g., Van Valin

1990: 224, Van Valin 1993: 38). For example, the sentences (1a) and (1c) would be assigned the logical structures (2a) and (2b), respectively.

- (2) a. [eat(Adam, apple)] CAUSE [BECOME consumed(apple)]  
 b. [walk(Kim)] CAUSE [BECOME be-at(store, Kim)]

Van Valin & LaPolla (1997: 100f) point out that a causative analysis of the accomplishments in (1) is untenable for at least two reasons: (i) It is difficult if not impossible to come up with a valid *causative paraphrase* for these constructions. (ii) Languages with *causative morphology* do not mark accomplishment constructions of this type as causative. As a consequence, Van Valin & LaPolla (1997: 111) propose the revised representations in (3), in which the CAUSE operator is replaced by the connective '&', which has the meaning 'and then'.

- (3) a. do(x, [eat(x, y)]) & BECOME consumed(y)  
 b. do(x, [walk(x)]) & BECOME be-at(y, x)

Accomplishment uses of activity verbs are thus conceived as denoting event sequences consisting of an activity immediately followed by an accomplishment. This analysis, however, has the undesirable implication that the accomplishment, which is a non-punctual change of state, does not start before the activity has ended. Therefore, the representations in (3) underwent a further revision in Van Valin (2005), where 'BECOME' is replaced by the punctual change of state operator 'INGR' in order to capture the fact that the resulting state sets in immediately with the activity's end. This has led to the two logical structures for active accomplishments shown in Table 3.

As noted in Van Valin (2005: 33/66), there are also a few lexicalized active accomplishments in English. The verb *devour* is one such example, which denotes an active accomplishment of consumption. Another candidate is the verb *repair* (cf. Rothstein 2012: 72). It is probably no accident that both, *devour* and *repair* originate from *prefixed* Anglo-French and eventually Latin verbs.

An important consequence of distinguishing active from causative accomplishments is that accomplishments are not bound to causation anymore, as they were in the original *Aktionsart* system. In the revised system, causation is basically orthogonal to the Vendlerian classification. In Table 3, this is reflected by fact that CAUSE can combine logical structures of any type.

### 3.3. Decomposing simple accomplishments

Van Valin (2005: 44) proposes a further decomposition of the BECOME operator in accomplishment structures. This step is motivated *inter alia* by data from Mparntwe Arrernte (Central Australia) which mark the difference between atelic and telic interpretations of deadjectival verbs overtly in the morphosyntax. Corresponding deadjectival verbs in English are *cool*, *darken*, and *dry*, which allow an atelic and a telic interpretation (*cool for an hour* vs. *cool in an hour*) without overt marking. Having this distinction appropriately represented in the semantics is of course not only relevant for languages which encode it overtly, such as Mparntwe Arrernte, but is equally important for languages like English. The class of verbs in question has been dubbed “degree achievements” in Dowty (1979/1991: 88) and has since then drawn considerable interest (e.g. Hay et al. 1999, Kearns 2007, Kennedy & Levin 2008)

(Simple) accomplishments (of states) consist by definition of a non-punctual, extended component and the final establishment of the resulting state. The solution proposed in Van Valin (2005: 44) is to make this event structure explicit by decomposing BECOME into a progression and an ingression component, that is, into the process (PROC) that leads to the result state and the final setting in of that state; in symbols, BECOME = PROC & INGR. For example, the semantic representation BECOME *cool*(*x*) of the (telic reading of the intransitive) English verb *cool* would be decomposed as in (4), with *cool* representing the meaning of the adjective *cool*.

$$(4) \text{ BECOME } \textit{cool}(x) = \text{PROC } \textit{cool}(x) \ \& \ \text{INGR } \textit{cool}(x)$$

The representation in (4) calls for a further elaboration in at least two respects. The first question is how exactly the expression PROC *cool*(*x*), and more generally PROC *P*(*x*) is to be interpreted. As to this question, there is already a tentative proposal in Dowty (1979/1991: 90) whose basic idea is that a degree predicate *P* such as *cool* is inherently *vague* and can be true at each time *t* during the process as expressed by the atelic *cool for an hour* if only *P* be resolved appropriately at each *t*. The assumption is that at each time, the undergoer has changed from  $\neg P$  to *P*, with varying resolutions of the vague predicate *P*. As a consequence, BECOME *P*(*x*) can be assumed to be true at all times within the interval denoted by *an hour*. From this perspective, it seems more adequate to apply the operator PROC not to *cool*(*x*) but to BECOME *cool*(*x*), understood properly. We will return to this point below in Section 3.4 when we discuss more recent proposals for the representation of active accomplishments, and later in Section 4.2.

A second problematic issue of the proposed decomposition in (4) is the connective ‘&’, which stands for ‘and then’ and thus expresses temporal succession. Taken literally, the logical structure in (4) says that the punctual change of turning cool (expressed by INGR cool) occurs *after* the process of cooling (expressed by PROC cool) has ended. This is clearly not an appropriate description since the object in question is cool *at the very moment* the cooling process comes to an end. In Section 4.2, we will propose an alternative representation that avoids this “and-then anomaly”.

### 3.4. Decomposing active accomplishments

As explained in Section 3.2, the textbook version of active accomplishment structures (cf. Table 3) does not make use of the BECOME operator anymore since the earlier proposal of Van Valin & LaPolla (1997), which had the form  $\text{do}(\dots) \& \text{BECOME} \dots$ , was replaced by  $\text{do}(\dots) \& \text{INGR} \dots$ . The reason is that it would be wrong to represent the non-punctual change of state as starting only after the activity has ended. However, as observed by Van Valin (2018), replacing ‘& BECOME’ by ‘& INGR’ in the representation of active accomplishments fails to take into account the incremental change of state (incremental consumption or creation, incremental motion to goal) that co-occurs with the activity. Van Valin (2018) therefore suggests reintroducing BECOME, but this time, the non-punctual, incremental component of the accomplishment is conflated with the activity. In the logical structure, this is realized by decomposing BECOME into PROC & INGR, as sketched in the previous section, and by conjoining the PROC component with the activity:

$$(5) [\text{do}(\dots) \wedge \text{PROC} \dots] \& \text{INGR} \dots$$

In this representation, the incremental change component of the accomplishment component is now temporally aligned with the activity component.

A not-so-obvious issue of the PROC & INGR decomposition is the choice of the predicates involved. As discussed in Section 3.3 in the context of degree achievements, a first question concerns the type of predicates the operator PROC applies to. Does PROC take a stative or a dynamic predicate? The first option would allow the straightforward decomposition of  $\text{BECOME } P$  into  $\text{PROC } P \& \text{INGR } P$ , with  $P$  a stative predicate. The predicate to which PROC is applied is identical to the one under BECOME and INGR in this case, but at the price of putting the burden of expressing the incremental change onto the operator PROC. The second option is to regard PROC as an operator that takes an incremental change description and turns it into the description of a

process that consists of ongoing incremental changes of this type.

In Van Valin (2018), the following tentative decompositions are proposed for the accomplishment components of the active accomplishment examples in (1):

- (6) a. [ $\text{do}(x, [\text{eat}(x, y)]) \wedge \text{PROC consume}(y)$ ] & INGR **consumed**(y)  
 b. [ $\text{do}(x, [\text{draw}(x, y)]) \wedge \text{PROC create}(y)$ ] & INGR **exist**(y)  
 c. [ $\text{do}(x, [\text{walk}(x)]) \wedge \text{PROC cover.path.distance}(x, (z))$ ]  
     & INGR **be-at**(y, x)

In these decompositions, the operator **PROC** is apparently applied to dynamic predicates and not to stative ones. We have **consume** instead of **consumed** in (6a), **create** instead of **exist** in (6b), and **cover.path.distance** instead of **be-at** in (6c). It is not fully transparent in each of these examples how the predicates under **PROC** and **INGR** are related to each other. In fact, it is not so clear what the predicates **consume** and **create** are supposed to mean in (6a) and (6b), respectively, since their only argument is the undergoer *y*, i.e., the entity being consumed or drawn. It seems that the predicates **being.consumed** and **being.created** would come closer to the intended meaning. The case of (6c) is slightly more intricate. As indicated by the predicate **cover.path.distance**, the incremental change concerns the position of the mover *x* on a path leading to the goal *y*.<sup>2</sup> That is, if **be-at** is to be part of the predicate under **PROC** then the location variable needs to be bound to varying locations that are getting successively closer to the goal (on the path taken).

Suppose the predicate under **PROC** should express the type of the ongoing incremental change. Then a systematic decomposition of **BECOME** into a **PROC** and an **INGR** component would require that the dynamic predicate under **PROC** is systematically related to the stative predicate under **INGR**. Moreover, the operator **PROC** should have basically the same effect in active accomplishment decompositions as it has in simple accomplishment structures. We will return to these issues in Sections 4.2 and 4.3.

### 3.5. Interim summary

If we look back at the development of the RRG *Aktionsart* system since its original adaption of Dowty (1979/1991), an interesting observation can be made. Recall that Dowty (1979/1991) was not satisfied with the decomposi-

<sup>2</sup> Note that the variable *z* in (6c) is intended to refer to the *distance* covered, not to the path traversed.

tion by **DO**, **CAUSE**, and **BECOME** as a basis of aspectual distinctions. He noted that the **DO** of agency is neither relevant for aspectual issues nor did he see how to give it a precise interpretation. The evaluation of **CAUSE** in the temporal dimension is also problematic. Only the change-of-state operator **BECOME** turned out to be useful for aspectual classification but was noted to lack the ability to express incremental changes.

In a sense, the successive modifications of the RRG system have addressed all of these issues in one way or another. To do largely away with **DO** was already proposed in Van Valin (1990, 1993). Moreover, the introduction of active accomplishments has released **CAUSE** from being responsible for defining accomplishments. In fact, causativity is now seen as a largely independent parameter when it comes to Vendlerian classes (cf. Van Valin 2005: 39). As to the deficiencies of **BECOME** with respect to capturing ongoing incremental changes, the suggested decomposition into **PROC** & **INGR** as well as the corresponding proposal for active accomplishments in Van Valin (2018) aim at coping with this problem as well.

In spite of these improvements, there are still a number of issues in the current representation system, as already mentioned in passing. We will address some of them in Section 4 and sketch possible modifications of the present system. The following section takes up again the important distinction between active and causative accomplishments.

### 3.6. Some intricacies of distinguishing active and causative accomplishments

Despite the two tests mentioned in Section 3.2, (i) existence of a causative paraphrase and (ii) morphological marking of causation in languages that provide such means, it is not always easy to distinguishing causative from active accomplishments. For instance, verbs of transportation like *carry* in motion to goal constructions such as (7) are potential candidates for active accomplishments even if the relevant logical structure, which has three arguments, is not covered by the two templates in Table 3.

(7) John carried the chair into the room.

As to criterion (ii), languages with causative morphology seem not to mark correlates of sentences like (7) as causative. In Georgian, for example, the concept of carrying someone or something somewhere is expressed by verbs of having (which vary with respect to the animacy of the object) plus a preverb of direction (Aronson 1990: 341f). Using verbs of having, taking or holding

combined with directed motion expressions is a widespread pattern for encoding transportation scenarios across languages (Wälchli, 2009). Causative morphology does not play a role here, in general. Possible exceptions are cases where the directed motion verb is transitive, in which case the verb is often analyzed as causative. In Japanese, for instance, the ‘carry’ verbs *hakobu* (‘carry, convey’), *motsu* (‘have, hold, carry’) and *katsugu* (‘carry on shoulders’) can combine with transitive verbs of directed motion such as *ireru* (‘cause to go in/take in’) and *dasu* (‘cause to go out/take out’) in lexical verb-verb compounds.<sup>3</sup> The use of such lexical compounds for expressing transport to goal scenarios is illustrated in (8a) for the compound verb *hakobiireru* (‘carry (and take) in(side)’).<sup>4</sup>

- (8) a. *Kagu o shinkyo ni hakobi-ire-ta.* Japanese  
 furniture ACC new home DIR carry-take.in-PAST  
 ‘They carried the furniture into their new home.’  
 b.  $\text{do}(x, [\text{carry}(x, y)]) \wedge [[\text{do}(x, \bullet)] \text{CAUSE} [\text{BECOME be-in}(z, y)]]$

Since the most obvious logical structure for *ireru* is  $[\text{do}(x, \bullet)] \text{CAUSE} [\text{BECOME be-in}(z, y)]$ , and since the carrying and the taking inside encoded by the compound *hakobiireru* occur simultaneously, the logical structure of the compound can be assumed to have the form (8b). In view of this structure, *hakobiireru* is probably best seen as a causative accomplishment verb (*ireru*, ‘take in’) with a lexically incorporated manner component (*hakobu*, ‘carry’).

The example in (8) does of course not provide evidence for a causative accomplishment analysis (9a) of the English construction in (7), nor does it preclude an active accomplishment analysis (9b) of that construction.

- (9) a.  $[\text{do}(x, [\text{carry}(x, y)])] \text{CAUSE} [\text{BECOME be-in}(z, y)]$   
 b.  $\text{do}(x, [\text{carry}(x, y)]) \& \text{INGR be-in}(z, y)$

The Japanese ‘carry’ verbs listed above can also combine with a number of intransitive directed motion verbs, in which case the moving entity is identified with the actor of the carrying (Matsumoto 1996: 206ff). Note that the resulting logical structure under an active accomplishment analysis differs from the one in (9b) in that not *y* but *x* is the theme of the location predicate.

We can conclude that while there is no morphosyntactic evidence for a

<sup>3</sup> The intransitive counterparts of *ireru* and *dasu* are *iru* (‘go into/enter’) and *deru* (‘come out/exit’), respectively.

<sup>4</sup> The example and its translation are taken from the *Compound Verb Lexicon* hosted by the National Institute of Japanese Language and Linguistics at [vlexicon.ninjal.ac.jp](http://vlexicon.ninjal.ac.jp).



causative accomplishment analysis of transport-to-goal constructions like (7), languages may differ with respect to which meaning components are realized overtly in the morphosyntax. It is part of the meaning of ‘carry’ verbs, and hence inherent in the predicate **carry**, that an actor *x* continuously and directly exerts force on an entity *y* with the effect that *y* is kept above the ground and (more or less) close to *x*. These conditions entail, or at least implicate, that if *x* is changing its location then *y* is moving along with *x* (putting aside the question of whether **carry** already entails the translocation of *x*). It thus basically makes no difference, with respect to what can be inferred about the final location of *y*, whether the (motion-to-)goal phrase adjoined to a ‘carry’ verb takes *x* or *y* as its argument. Making these meaning aspects of **carry** explicit, however, would require more expressive representations than the current logical structures of RRG (cf. Section 4 below and, e.g., Krifka 2004, Beavers 2011: 38).

The causative paraphrase criterion (i) gives mixed results when applied to a sentence like (7). On the one hand, *John caused the chair to become to be in room* seems an acceptable paraphrase in sense of Van Valin (2005: 38). But it does not support the causative accomplishment analysis in (9b), which says that John’s carrying the chair caused it to become to be in the room. It is not just John’s carrying the chair, it’s John’s carrying the chair into the room that caused the chair to become to be there. The ‘caused’ is to be understood here as indicating a *causal explanation* rather than as referring to a causal relation between events (cf. Davidson 1967: 703, Vendler 1962). In fact, the chair’s becoming to be in the room is an *entailment* of John’s carrying it there.

Caused motion to goal constructions introduce further intricacies, and it is instructive to see how different authors vary in their judgement with respect to the active/causative accomplishment distinction. Consider the following example taken from Rappaport Hovav & Levin (2001).

(10) We pulled the crate out of the water.

According to Rappaport Hovav & Levin (2001: 783), verbs like *pull* “describe the exertion of a force on a physical object, but do not lexically entail that the force displaces the object. However, when the force does cause a displacement, as in [(10)], then it must be exerted until the result location is attained. Thus, the exertion of force and the displacement occur in tandem.” Moreover, they note that (10) “is not so well paraphrased by ‘Our pulling the crate caused it to become out of the water.’” This line of reasoning, which is similar to the one employed in the above analysis of *carry*, suggests an active accomplishment analysis for (10). Goldberg & Jackendoff (2004: 558), on the other hand, disagree with Rappaport Hovav & Levin about the status of the causative para-

phrase. For them, “it is unclear why [(10)] is considered non-causal” since “it is naturally paraphrased by ‘we caused the crate to move out of the water by pulling; [...]’.” This causative paraphrase, however, is subject to the same objections as the one discussed for *carry*. While the phrase ‘by pulling’ correctly mentions one of causal factors involved, namely the exertion of force, it does not sufficiently describe the cause for the crate to become to be out of the water. The appropriate description would be ‘by pulling at the crate until it was out of the water’, which is a causal explanation or, rather, an elaboration.

The domain of caused motion descriptions contains further tricky aspects relevant to the distinction between active and causative accomplishments. For instance, verbs like *push* in motion to goal constructions are ambiguous with respect to whether they mean a continuous exertion of force on the moving object or a punctual push which causes the motion. The first case suggests an active accomplishment analysis along the lines of the *pull* example discussed above, the second case clearly calls for a causative accomplishment structure. A more thorough analysis of this issue is beyond the scope of the present paper, however.

Let us finally have another look at verbs of consumption in accomplishment constructions such as (1a). Consider the following paraphrase for (1a): *Adam's eating an apple caused the apple to become consumed*. This is surely not the most natural and elegant way to express the scenario in question, but on what grounds can it be qualified as inadequate or wrong? The crucial question is whether the relation between the eating activity and the becoming consumed of the affected object or stuff can be characterized as one between cause and effect. Reflecting on the use of the verb *eat* will most probably lead us to the conclusion that it is inherent in what we conceive as eating scenarios that the object eaten (at) gets (successively) consumed. By using the verb *eat*, we do not spell out any of the (possibly causal) chains of events involved (such as separating pieces of food from some bigger chunk by means of the teeth, a knife, a fork, a spoon, etc., putting it into the mouth, and swallowing it, maybe after chewing it). It is thus part of every eating event that the food eaten gets consumed; put differently, that food is becoming consumed when being eaten is a simple entailment, and this entailment holds already for the activity interpretation of *eat*. The accomplishment interpretation of the construction in (1a) then comes about through the implicature that the apple is consumed *completely*.

## 4. Possible modifications of the *Aktionsart* representations

### 4.1. Activities revisited

In the current version of RRG, activities are represented as primitive predicates enclosed by  $\text{do}(x, [\dots])$  (cf. Section 3.1). Examples are  $\text{do}(x, [\text{walk}(x)])$  and  $\text{do}(x, [\text{carry}(x, y)])$ . The purpose of this **do** envelope is not fully clear, however. If the embedded predicate denotes an activity then adding **do** appears to be redundant from a semantic point of view. We could treat  $\text{do}(x, [\dots])$  simply as a fairly clumsy diacritic which marks off activity predicates and which in addition highlights the actor argument. As mentioned in Section 3.1, having this information readily available in this way might be useful for the formulation of the linking principles. But there are other ways of representing and accessing the relevant information. Instead of referring to the first argument of  $\text{do}(x, [\dots])$ , the linking algorithm can simply refer to the first argument of the activity predicate itself. Moreover, in order to single out activity predicates, we may, for instance, assume that such predicates entail a general activity predicate **do**. That is, the lexicon contains conditionals like  $\text{walk}(x) \rightarrow \text{do}(x)$ . In the following sections, we will therefore write  $\text{walk}(x)$  instead of  $\text{do}(x, [\text{walk}(x)])$ .

There is a limited number of activity verbs for which the **do** envelope is not redundant since the embedded predicate is stative. The only cases mentioned in Van Valin (2005) are verbs of directed perception such as *listen* and *look at*, *watch* which are analyzed as activity versions of stative perception predicates and are represented as  $\text{do}(x, [\text{hear}(x, y)])$  and  $\text{do}(x, [\text{see}(x, y)])$ , respectively. These examples are already discussed in Dowty (1979/1991: 113f) as one of the few cases for which there is a difference between the activity structure and the embedded (stative) predicate. Another of Dowty's examples is the difference between stative and agentive readings of certain adjectives and nouns in predicative constructions. Consider, for instance, the contrast between *John is polite/a hero* and *John is being polite/a hero*. It seems reasonable to argue that the 'is being' construction and likewise the verbs of directed perception encode agentivity and are thus candidates for the use of **DO**. If this is true then **do** is redundant here as well.

### 4.2. Accomplishments revisited

The decomposition of BECOME into PROC & INGR introduced in Sections 3.3 and 3.4 leaves open the following questions: (i) Which kinds of pred-

icates does PROC apply to? (ii) How is the predicate under PROC related to the one under BECOME (and INGR) in the decomposition? (iii) How can we resolve the “and-then anomaly” that comes with ‘&’? Let us begin with the last question. Since the final or result state sets in at the very moment the process ends, it seems appropriate to replace ‘& INGR  $P(x)$ ’ by something like ‘and finally  $P(x)$ ’ or, more formally, ‘ $\wedge$  FINAL  $P(x)$ ’, where ‘ $\wedge$ ’ is ordinary conjunction. The term FINAL  $P(x)$  is to be taken as true of an event just in case  $P(x)$  holds at the end of the event. The underlying idea is that every bounded event has an attribute FIN(AL) whose value is the *final stage* of the respective event.<sup>5</sup> An event  $e$  satisfies FIN  $P(x)$  if the FIN value  $s$  of  $e$  satisfies  $P(x)$ , i.e., if  $s$  is a stage of  $x$ ’s being  $P$ .<sup>6</sup> Note that FIN  $P(x)$  does not exclude *per se* the possibility that  $P(x)$  already holds at earlier stages of the event. One way to cope with this point is to derive  $\neg P(x)$  for all non-final stages from an appropriately defined representation of the process component.<sup>7</sup> In fact, the version proposed below would allow such an inference. Another option is to represent the accomplishment interpretation of degree achievement verbs simply as INIT  $\neg P(x) \wedge$  FIN  $P(x)$ , which is basically Dowty’s explication of BECOME  $P(x)$ , where INIT(IAL) is an attribute whose value is the initial stage of a bounded event. The representation of the process component would be regarded as irrelevant in this case, for one could argue that the expression *the soup cooled in ten minutes* is true if the soup was not cool at a contextually given moment and it was cool ten minutes later.

As to question (i), the decomposition of BECOME in the active accomplishment representations (6) of Section 3.4, as well as Dowty’s sketch of the representation of degree achievements mentioned in Section 3.3, has led us to the conclusion that the operator PROC applies to dynamic predicates, i.e., to predicates which denote an incremental change. When applied to such a predicate, PROC returns an expression which denotes processes consisting of ongoing incremental changes of the type denoted by that predicate. Question (ii) asks how the dynamic predicate under PROC is defined in terms of the stative predicate that underlies degree achievements such as *cool*. The dynamic predicate needs to express the incremental change towards becoming cool in

<sup>5</sup> Stages are here conceived as *instantaneous parts* or “snap-shots”; cf. Sider (2001).

<sup>6</sup> Here and in the following, expressions such as  $F \phi$  are to be read as predicates of the form  $\lambda x \exists y (F(x, y) \wedge \phi(y))$ , where  $F$  is denoted a *functional* relation and  $\phi$  is a one-place predicate. (Alternatively, we could write  $\lambda x \exists y (F(x) = y \wedge \phi(y))$  or more succinctly  $\lambda x \phi(F(x))$  if we put aside the complication that  $F$  is not a total but only a partial function, in general.)

<sup>7</sup> See also Pustejovsky (1991, 2000).

this case. According to Dowty's tentative proposal mentioned in Section 3.3, the vague predicate **cool** is resolved differently at each time during the cooling process in such a way that **BECOME cool**( $x$ ) holds throughout the process. This idea has been taken up by Abusch (1986), who assumes that predicates such as **cool** depend on a contextual parameter and that the *Aktionsart* interpretation of the corresponding degree achievement verb varies according to whether this parameter is determined by the context of utterance or is existentially bound. This leads to **BECOME (cool**( $c_{uti}$ ))( $x$ ) in the first case and to a predicate of the form  $\lambda e \exists c (\text{BECOME } (\text{cool}(c))(x)(e))$  in the second case.<sup>8</sup> Following Dowty and Abusch, the latter predicate can be taken as equivalent to  $\lambda e \exists c ((\text{INIT} \neg (\text{cool}(c))(x) \wedge \text{FIN } (\text{cool}(c))(x))(e))$ . Since the parameter  $c$  can be chosen appropriately for every event, the predicate is true of every subevent of an event in which  $x$  is cooling continuously.<sup>9</sup> By comparison, if the parameter is fixed by the context of utterance then the change-of-state predicate corresponds to an accomplishment reading of the degree achievement verb.

Analyzing the process interpretation of (deadjectival) degree achievement verbs in terms of a variable, context-dependent resolution of the underlying gradable adjective is not very convincing, however. The final stage of the incremental change is obviously better described *comparatively* with respect to the initial stage of the change; the change is one from *less cool* to *more cool* rather than from *not cool* to *cool* with *cool* resolved appropriately.<sup>10</sup> Put simply, the incremental change is an increase of the coolness of the undergoer. Specifying the increase presupposes a *scale*, i.e. an *ordered* set of *degrees*. Since an increase of coolness amounts to a decrease of temperature, and since *cool* is commonly analyzed as a dimensional adjective with negative polarity, with temperature as the dimension of measurement, the scale associated with *cool* consists of degrees of temperature with an inverted ordering (i.e., higher degrees of temperature are lower with respect to the scale ordering) (cf. e.g. Bierwisch 1989, Kennedy & McNally 2005). Suppose **temp-stg**( $x, d$ ) denotes stages of  $x$ 's having temperature  $d$ . Then changes of state in which  $x$  is becoming cooler can be characterized by the predicate in (11).

<sup>8</sup> The presentation follows basically Kennedy & Levin's (2008: 158) brief exposition of Abusch's approach in an event semantics framework.

<sup>9</sup> Note that it would be wrong to conclude that if the predicate under discussion is true of an event  $e$  then it is true of all subevents of  $e$ , for an intermediate rise of the temperature is not excluded.

<sup>10</sup> Kennedy & Levin (2008: 171) come to a similar conclusion (but take a different route in their formalization).

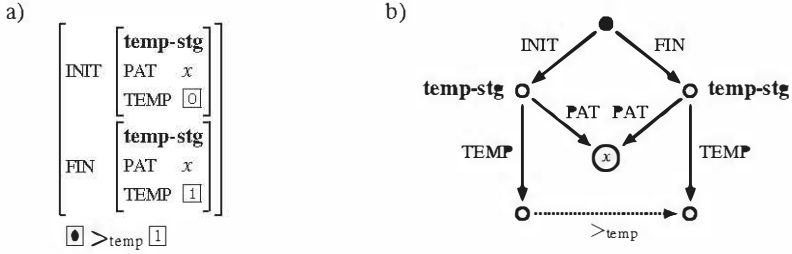


Figure 1: Attribute-value matrix (a) and (minimal) frame model (b) for (12)

$$(11) \lambda e \exists d_{\bullet} \exists d_1 ((\text{INIT } \mathbf{temp-stg}(x, d_{\bullet}) \wedge \text{FIN } \mathbf{temp-stg}(x, d_1))(e) \wedge d_{\bullet} >_{\text{temp}} d_1)$$

The representation in (12) below is essentially equivalent to the one in (11) but employs more systematically the idea of a representation based on attributes. To this end, **temp-stg** is considered as a stage predicate to which the arguments  $x$  and  $d$  are bound by the functional relations **TH(EME)** and **PAT(IENT)**, respectively.<sup>11</sup>

$$(12) \text{INIT } (\mathbf{temp-stg} \wedge \text{PAT} \dot{=} x) \wedge \text{FIN } (\mathbf{temp-stg} \wedge \text{PAT} \dot{=} x) \wedge \text{INIT TEMP } >_{\text{temp}} \text{FIN TEMP}$$

A semantic decomposition based on attributes (i.e. functional relations) naturally leads to *frame-based* representations.<sup>12</sup> Figure 1b) depicts the frame structure defined by the attribute-value description in (12). The shaded node at the top represents the referential node of the frame; the attribute-value matrix on the left of the figure is a notational variant of the expression in (12).

We started off with the question of how the dynamic predicate under **PROC** is related to the stative predicate that underlies *cool*. Given that the representations in (11) and (12) are candidates for the predicate under **PROC**, what is their relation to the respective stative predicate? A primitive stative predicate such as **cool** is obviously not helpful in this respect. What is needed

<sup>11</sup> In addition to the notation introduced in Footnote 6, the following conventions are used here:  $F \dot{=} y$  stands for  $\lambda x F(x, y)$ ,  $FG$  is short for  $\lambda x \lambda y \exists z (F(x, z) \wedge G(z, y))$ , and  $\text{Fr } G$  expands into  $\lambda x \exists y \exists z (F(x, y) \wedge G(x, z) \wedge \mathbf{r}(y, z))$ , where  $F$  and  $G$  are functional relations and  $\mathbf{r}$  is an arbitrary binary relation.

<sup>12</sup> Cf. Osswald & Van Valin (2014) for a general outline of decompositional frame semantics and Kalhneier & Osswald (2013) for an introduction to the formal framework.

is a lexical semantic representation of dimensional adjectives, as proposed in Bierwisch (1989), Kennedy & McNally (2005), and others, which takes into account the dimension of measurement as well as the scale structure of possible degrees and their ordering. With these semantic ingredients available, the simple positive use of *cool* in expressions such as *the soup is cool* can be roughly characterized by the stage predicate in (13), where  $\text{std}_{\text{coolness}}$  is a (contextually specified) standard of comparison associated with the coolness scale.<sup>13</sup>

$$(13) \text{ temp-stg} \wedge \text{PAT} \doteq x \wedge \text{std}_{\text{coolness}} >_{\text{temp}} \text{TEMP}$$

A full-fledged frame-semantic representation of *cool* would provide attributes for accessing the underlying stage type as well as the coolness scale and its components, including the degree ordering and the standard of comparison. These components are then employed in different ways at the interface to morphology and syntax for deriving the representations in (12) and (13), among others. While spelling this out in detail is beyond the scope of the present paper, it should be clear from the foregoing discussion that a decomposition of  $\text{BECOME } P(x)$  requires a proper decomposition of  $P(x)$ , irrespective of the chosen semantic formalism.

We conclude this section with a brief sketch of how to cope with the operator  $\text{PROG}$  in a semantic framework that is primarily based on attributes. Recall that if  $\phi$  is a change-of-state predicate such as (12) then  $e$  satisfies  $\text{PROG } \phi$  if  $e$  consists of ongoing changes of type  $\phi$ . A straightforward solution is to characterize  $e$  by an attribute  $\text{PROG}(\text{RESSION-TYPE})$  whose value is the type of change that goes on in the course of  $e$ .<sup>14</sup> In the example discussed above, the respective type  $T$  would be an abstract representation of the frame in Figure 1. An additional axiom would ensure that if  $e$  satisfies  $\text{PROG} = T$  then any segment of  $e$  is an instance of  $T$ .<sup>15</sup>

### 4.3. The structure of active accomplishments

The different versions of active accomplishment structures discussed in Section 3 all suffer the “and-then anomaly”. This is also true of the latest proposal

<sup>13</sup> As to the last conjunct in (13), expressions of the form  $c \mathbf{r} G$  are to be interpreted as  $\lambda x \exists y (G(x, y) \wedge \mathbf{r}(c, y))$ , where  $c$  is a variable or constant,  $G$  is a functional relation and  $\mathbf{r}$  is an arbitrary binary relation.

<sup>14</sup> Galton (2012) seems to advocate a similar view of the nature of processes and activities.

<sup>15</sup> Cf. Balogh & Osswald (2020) for a more elaborate explanation of this approach.

(cf. Section 3.4), which conflates the process component of the accomplishment with the activity component. The logical structures in (6) express that the activity (or process) ends *and then* a punctual change of state occurs. For example, (6b) means that the drawing activity and the simultaneously ongoing creation come to an end *and then* the resulting object comes into existence. Likewise, in the motion-to-goal case (6c), the given analysis says that the motion activity stops *and then* a punctual change of state from not being at the goal to being at the goal occurs. This is obviously not what's happening. Rather, the drawn figure exists at the very moment the drawing activity stops. Likewise, the mover is at the goal at the very moment the motion stops; reaching the goal co-occurs with coming to a final halt – it is not something which happens afterwards.<sup>16</sup> It thus seems advisable to avoid the temporal connective '&' when combining process or activity expressions with accomplishment or achievement expressions in active accomplishment structures.

Instead of starting with  $\phi$  & BECOME  $\psi$  as proposed in Van Valin & LaPolla (1997), we may simply start with  $\phi \wedge$  BECOME  $\psi$ . This representation does not have the above-mentioned temporal succession problem of Van Valin & LaPolla's original proposal, which lead to  $\phi$  & INGR  $\psi$  in Van Valin (2005) and eventually to the reintroduction of BECOME as PROC & INGR in Van Valin (2018). Applied to the motion-to-goal case (1c), we get  $\text{walk}(x) \wedge$  BECOME  $\text{be-at}(y, x)$ , which describes walking activities of  $x$  that are simultaneously describable as events of  $x$ 's coming to be at  $y$ . For a decomposition of BECOME, we can follow the strategy of Section 4.2 and use the predicate FIN  $\text{be-at}(y, x)$  to describe events that have final stages at which  $\text{be-at}(y, x)$  holds. The revised version of (6c) then looks as follows, where the expression under PROC remains to be explicated.

(14)  $\text{walk}(x) \wedge \text{PROC} \dots \wedge \text{FIN } \text{be-at}(y, x)$

Recall that BECOME  $\text{be-at}(y, x)$  is intended to capture the meaning of the English preposition *to* in motion-to-goal constructions like *walk to the store* (1c). It is commonly taken as part of the meaning of *to* in such a context that the undergoer  $x$  moves continuously towards the goal until the target area is reached. In other words,  $x$  traverses a path that leads to the goal (keeping aside the possibility of teleportation). The incremental change component should capture something like 'move (forward) along a path that leads to  $y$ '. Thus a possible candidate of the predicate under PROC could

---

<sup>16</sup> In fact, the motion does not even have to stop at the goal but can continue. But it definitely does not stop before reaching the goal.



be **move.along.path.to**( $x, y$ ). Since the incremental steps of a given motion event are all bound to one and the same path, it seems more appropriate to explicitly introduce the path as a constant  $p$  on a par with  $x$  and  $y$ , as in **move.along.to**( $x, p, y$ ), or, more simply, **move.along**( $x, p$ ), with the additional constraint that the path  $p$  ends at  $y$ . The incremental change expressed by **move.along**( $x, p$ ) is that  $x$  (continuously) changes its location on  $p$  in the direction of  $p$ . In particular, this requires that the path  $p$  is directed, that is, its points are linearly ordered. An event (segment)  $e$  of type **move.along**( $x, p$ ) then satisfies INIT **be-at**( $u, x$ )  $\wedge$  FIN **be-at**( $v, x$ ), where  $u$  and  $v$  are points on the path  $p$  and  $u$  precedes  $v$  with respect to the direction of  $p$ , in symbols,  $u <_p v$ . That is,  $x$  is located at  $u$  at the initial stage of  $e$ , and at  $v$  at the final stage of  $e$ . We arrive at the following representation, which shows the same pattern as (12).

$$(15) \text{ INIT } (\mathbf{be-at} \wedge \text{THEME} \doteq x) \wedge \text{FIN } (\mathbf{be-at} \wedge \text{THEME} \doteq x) \\ \wedge \text{INIT LOC} <_p \text{FIN LOC}$$

In contrast to cases like (12), the scalar structure that underlies (15) is not lexicalized but compositionally derived. This is a general property of the active accomplishment constructions in (1) since the activity verb does not *per se* provide a scale. The scalar structures are either path scales, as in the motion-to-goal constructions above, or “extent” or “volume” scales in the creation/consumption constructions, in which case the scale is defined in terms of the incremental theme object (Rappaport Hovav 2008, Kennedy 2012). In all of these cases, the scalar structure is dependent on the event denoted by the activity component of the construction.

#### 4.4. Causative structures

We conclude this section by showing how causative structures fit into the picture developed so far. Dowty (1979/1991) treats causal expressions such as (16a) as statements (like all other formulas of his aspect calculus). If we interpret (16a) as a statement which expresses the causal relatedness of two events then a logical explication of (16a) could look like (16b), where the variables  $e_1$  and  $e_2$  refer to events and  $\phi(e_1)$  and  $\psi(e_2)$  are sentences and thus fully tensed clauses.

$$(16) \text{ a. } \phi \text{ CAUSE } \psi \\ \text{ b. } \exists e_1 \exists e_2 (\mathbf{cause}(e_1, e_2) \wedge \phi(e_1) \wedge \psi(e_2))$$

From an RRG point of view, it seems more reasonable to read (16a) as an event description, which is typically associated with a nuclear syntactic structure (at least in the case of direct causation) and thus not yet specified with respect to time and space. This means that the causal expression denotes *complex events* of type *causation*, which have two event components, namely the *cause* component and the *effect* component. A formalization of this analysis is given in (17a), where CAUS(E) and EFF(E)CT are the two attributes of **causation** events whose values are the cause and the effect component, respectively.

- (17) a.  $\lambda e \exists e_1 \exists e_2 (\text{causation}(e) \wedge \text{CAUS}(e, e_1) \wedge \text{EFFCT}(e, e_2) \wedge \phi(e_1) \wedge \psi(e_2))$   
 b. **causation**  $\wedge$  CAUS  $\phi$   $\wedge$  EFFCT  $\psi$

Note that in (17), CAUS is *not* meant to denote the relation between cause and effect but rather the functional relation between complex events of type causation and their cause components. By applying the notational conventions introduced in the previous sections, we can replace the expression in (17a) by the compact version in (17b). The latter expression can be shortened further to CAUS  $\phi$   $\wedge$  EFFCT  $\psi$  if we assume that having the attributes CAUS and EFFCT implies being of type **causation**, in symbols: CAUS  $\top$   $\wedge$  EFFCT  $\top \Rightarrow$  **causation**, where  $\top$  is the *universal* predicate, which is true of everything.<sup>17</sup> It is also possible to state general constraints on the temporal relation between cause and effect. Let  $<_{\bullet}$  be the relation that holds between two events (or states) if the first starts earlier than the second. Then we have **causation**  $\Rightarrow$  CAUS  $<_{\bullet}$  EFFCT.

In (18), the revised representation is applied to a causative resultative construction. The corresponding standard RRG representation is shown in (18c) for comparison. The complex causation scenario described in (18a) has as its CAUS component a **kick** activity of Kim at the door, and as its EFFCT component a change-of-state of the door becoming **shut**.

- (18) a. Kim kicked the door shut.  
 b. CAUS (**kick**  $\wedge$  EFFCTR  $\doteq x$   $\wedge$  PAT  $\doteq y$ )  
 $\wedge$  EFFCT FIN (**shut**  $\wedge$  PAT  $\doteq y$ )  
 c. [**do**( $x$ , **kick**( $x$ ,  $y$ ))] CAUSE [**BECOME shut**( $y$ )]

As mentioned before, an attribute-based analysis of event predicates suggests a “Neo-Davidsonian” representation of the relation between an event and its

<sup>17</sup>  $\phi \Rightarrow \psi$  is short for  $\forall x(\phi(x) \rightarrow \psi(x))$ , where  $\phi$  and  $\psi$  are one-place predicates. Note that  $F \top$  is equivalent to  $\lambda x \exists y F(x, y)$  by definition of  $\top$ .

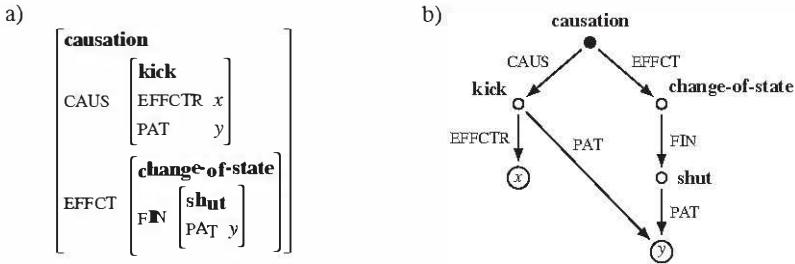


Figure 2: Attribute-value matrix (a) and frame model (b) for (18b)

participants in terms of (functional) thematic relations. Hence, (18b) uses  $\mathbf{kick} \wedge \text{EFFCTR} \doteq x \wedge \text{PAT} \doteq y$  instead of  $\mathbf{kick}(x, y)$ , with the thematic relations  $\text{EFF(E)CT(OR)}$  and  $\text{PAT(IENT)}$ . Figure 2a) shows a (slightly enriched) presentation of (18b) as an attribute-value matrix; Figure 2b) depicts the corresponding frame model.

Since the logical structures of RRG play a crucial role in the linking system, a few words might be in order about how the proposed representations can be integrated into that system. Recall that it is the position of the arguments in the logical structures that determines the macrorole assignment and thereby, together with other factors, the realization of the arguments in the morphosyntax. With respect to primitive predicates, the representations proposed above differ insofar as they rely on thematic relations instead of the linear ordering of the arguments. However, these two notions are basically equivalent (cf. Van Valin 2005: 55). The depth of embedding of an argument in the representations, on the other hand, which is equally important for the macrorole assignment, is preserved in the attribute-based representations.

## 5. Conclusion

In the first part of the paper, we have seen that many of the revisions of the RRG *Aktionsart* system go back to deficiencies of Dowty's original Generative Semantics-inspired decomposition system, most of which have been pointed out by Dowty himself, in one way or another. These deficiencies comprise, among others, problems of coming up with a proper denotational interpretation of the operators, as in the case of  $\mathbf{DO}$ , and issues related to developing a sufficiently fine-grained representation of the temporal course of an event.

The current version of the *Aktionsart* system has resolved a good part of these deficiencies but is still not without flaws. We have discussed the “and-

then anomaly” and the problems of characterizing the process components of simple and active accomplishments. One of our insights was that the decomposition of accomplishments into a process and a result component calls for a decompositional analysis of the underlying predicate. We sketched a system of frame-based representations which allows such a decomposition, has a clear denotational interpretation, and preserves the structural properties of the logical structures. Among the many issues to be spelled out in more detail is the frame-based description of the compositional scale constitution for active accomplishments mentioned at the close of Section 4.3. Another such issue is the proper integration of causal factors in the transitive motion cases discussed in Section 3.6.

It remains to point out that the attribute-based approach sketched in Section 4 is of course not the only way of turning the *Aktionsart* system and its recent extensions into a formal representation language. Formalization as such, however, is indispensable if a precise semantic interpretation is intended to be achieved. Dowty (1979/1991) was meticulous in specifying the formal makeup (i.e., the syntax) and the model-theoretic interpretation of his representation language (dubbed by him the ‘translation language’, following Montague). When the decompositional elements of Dowty’s system were taken over by Foley & Van Valin (1984), formal rigor did not play a similar prominent role, nor did it in the subsequent modifications and extensions of the RRG *Aktionsart* system. The main concern of Foley & Van Valin and later work on RRG was rather to show how semantic structures can contribute to explaining, among others, how verbal arguments and interclausal semantic relations are grammatically encoded across languages.

## Acknowledgements

The research presented in this article was supported by the Collaborative Research Centre 991 “The Structure of Representations in Language, Cognition, and Science” funded by the German Research Foundation (DFG) and the project TreeGraSP funded by the European Research Council (ERC). I would like to thank Robert D. Van Valin, Jr. for valuable comments on earlier drafts of this article.

## References

- Abusch, Dorit. 1986. Verbs of change, causation and time. Technical Report 86-50, CSLI, Stanford, CA.

- Aronson, Howard I. 1990. *Georgian: A Reading Grammar. Corrected Edition*. Slavica Publishers, Inc., Columbus, Ohio.
- Balogh, Kata & Rainer Osswald. 2020. A frame-based analysis of verbal particles in Hungarian. In Sebastian Löbner, Thomas Gamerschlag, Tobias Kalenschner, Markus Schrenk & Henk Zeevat, eds., *Concepts, Frames and Cascades in Semantics, Cognition and Ontology*. Springer, Dordrecht.
- Beavers, John. 2011. An aspectual analysis of ditransitive verbs of caused possession in English. *Journal of Semantics*, 28:1–54.
- Bierwisch, Manfred. 1989. The semantics of gradation. In Manfred Bierwisch & Ewald Lang, eds., *Dimensional Adjectives. Grammatical Structure and Conceptual Interpretation*, 71–261. Springer, Berlin.
- Davidson, Donald. 1967. Causal relations. *Journal of Philosophy*, 64(21): 691–703.
- Dowty, David. 1977. Toward a semantic analysis of verb aspect and the English imperfective ‘progressive’. *Linguistics and Philosophy*, 1:45–77.
- Dowty, David. 1979/1991. *Word Meaning and Montague Grammar*. D. Reidel, Dordrecht.
- Foley, William A. & Robert D. Van Valin, Jr. 1984. *Functional Syntax and Universal Grammar*. Number 38 in Cambridge Studies in Linguistics. Cambridge University Press, Cambridge.
- Galton, Antony P. 2012. The ontology of states, processes, and events. In Mitsuhiro Okada & Barry Smith, eds., *Proceedings of the Fifth Interdisciplinary Ontology Conference*, 35–45, Tokyo, Keio University.
- Goldberg, Adele E. & Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language*, 80:532–568.
- Hay, Jennifer, Christopher Kennedy & Beth Levin. 1999. Scalar structure underlies telicity in “degree achievements”. In T. Matthews & D. Strolovitch, eds., *Proceedings of SALT IX*, 127–144, Ithaca, NY, CLC Publications.
- Kallmeyer, Laura & Rainer Osswald. 2013. Syntax-driven semantic frame composition in Lexicalized Tree Adjoining Grammars. *Journal of Language Modelling*, 1(2):267–330.
- Karttunen, Lauri. 1971. The logic of English predicate complement constructions. Distributed by the Indiana University Linguistics Club.
- Kearns, Kate. 2007. Telic senses of adjectival verbs. *Lingua*, 117:26–66.
- Kennedy, Christopher. 2012. The composition of incremental change. In Violeta Demonte & Louise McNally, eds., *Telicity, Change, and State. A Cross-Categorical View of Event Structure*, Oxford Studies in Theoretical Linguistics, 103–121. Oxford University Press, Oxford.
- Kennedy, Christopher & Beth Levin. 2008. Measure of change: The adjectival

- core of degree achievements. In Louise McNally & Christopher Kennedy, eds., *Adjectives and Adverbs. Syntax, Semantics, and Discourse*, 156–182. Oxford University Press, Oxford.
- Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification and the semantics of gradable predicates. *Language*, 81:345–381.
- Krifka, Manfred. 2004. Semantic and pragmatic conditions for the Dative Alternation. *Korean Journal of English Language and Linguistics*, 4:1–32.
- Matsumoto, Yo. 1996. *Complex Predicates in Japanese*. CSLI Publications, Stanford, CA.
- Osswald, Rainer & Robert D. Van Valin, Jr. 2014. FrameNet, frame structure, and the syntax-semantics interface. In Thomas Gamerschlag, Doris Gerland, Rainer Osswald & Wiebke Petersen, eds., *Frames and Concept Types*, 125–156. Springer, Dordrecht.
- Pustejovsky, James. 1991. The syntax of event structure. *Cognition*, 41:47–81.
- Pustejovsky, James. 2000. Events and the semantics of opposition. In Carol Tenny & James Pustejovsky, eds., *Events as Grammatical Objects*, 445–482. CSLI Publications, Stanford, CA.
- Rappaport Hovav, Malka. 2008. Lexicalized meaning and the internal temporal structure of events. In Susan Rothstein, ed., *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, 13–42. John Benjamins, Amsterdam.
- Rappaport Hovav, Malka & Beth Levin. 2001. An event structure account of English resultatives. *Language*, 77(4):766–797.
- Ross, John Robert. 1972. Act. In Donald Davidson & Gilbert Harman, eds., *Semantics of Natural Language*, 70–126. D. Reidel, Dordrecht.
- Rothstein, Susan. 2012. Another look at accomplishments and incrementality. In Violeta Demonte & Louise McNally, eds., *Telicity, Change, and State. A Cross-Categorial View of Event Structure*, Oxford Studies in Theoretical Linguistics, 60–102. Oxford University Press, Oxford.
- Sider, Theodore. 2001. *Four-Dimensionalism. An Ontology of Persistence and Time*. Clarendon Press, Oxford.
- Stassen, Leon. 2009. *Predicative Possession*. Oxford University Press, Oxford.
- Van Valin, Robert D., Jr. 1990. Semantic parameters of split intransitivity. *Language*, 66(2):221–260.
- Van Valin, Robert D., Jr. 1993. A synopsis of Role and Reference Grammar. In Robert Van Valin, ed., *Advances in Role and Reference Grammar*, 1–164. John Benjamins, Philadelphia.
- Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge University Press, Cambridge.

- Van Valin, Robert D., Jr. 2010. Role and Reference Grammar as a framework for linguistic analysis. In Bernd Heine & Heiko Narrog, eds., *The Oxford Handbook of Linguistic Analysis*, 703–738. Oxford University Press, Oxford.
- Van Valin, Robert D., Jr. 2018. Some issues regarding (active) accomplishments. In Rolf Kailuweit, Eva Staudinger & Lisann Künnel, eds., *Applying and Expanding Role and Reference Grammar*, (NIHIN Studies), 355–378. Albert-Ludwigs-Universität, Universitätsbibliothek, Freiburg.
- Van Valin, Robert D., Jr. & Randy J. LaPolla. 1997. *Syntax*. Cambridge University Press, Cambridge.
- Van Valin, Robert D., Jr. & David P. Wilkins. 1996. The case for ‘Effector’: Case roles, agents, and agency revisited. In Masayoshi Shibatani & Sandra A. Thompson, eds., *Grammatical Constructions*, 289–322. Clarendon Press, Oxford.
- Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review*, 66(2): 143–160.
- Vendler, Zeno. 1962. Effects, results and consequences. In R. J. Butler, ed., *Analytic Philosophy*, 1–14. Barnes and Noble, New York.
- Wälchli, Bernhard. 2009. Motion events in parallel texts. A study in primary-data typology. Universität Bern. Habilitationsschrift.

# SIMPLEX AND COMPLEX PREDICATES IN PERSIAN – AN RRG ANALYSIS

JENS FLEISCHHAUER

HEINRICH HEINE UNIVERSITY DÜSSELDORF

## Abstract

Persian only has a limited number of simplex verbal predicates. Instead, Persian makes frequent use of complex predicates, mostly light verb constructions consisting of a semantically light verb and a second predicational element. Light verb constructions bear superficial resemblance to cases of pseudo-incorporation. Although various authors do not distinguish different types of complex predicates, I argue that light verb constructions should be strictly distinguished from cases of pseudo-incorporation. The two constructions are syntactically alike but exemplify different patterns of semantic composition. With respect to these two constructions, I claim that they have the same syntax, but different semantics.

## Keywords

light verb constructions, pseudo-incorporation, Persian, RRG

## 1. Introduction<sup>1</sup>

Persian – as well as various other Iranian languages – only possesses a small number of lexically full verbs. Instead, Persian makes use of com-

---

<sup>1</sup> List of abbreviations: ACC = accusative, CP = complex predicate, DEM = demonstrative, EZ = Ezâfe, ICE = idiomatically combining expression, INDEF = indefiniteness, INF = infinitive, LV = light verb, LVC = light verb construction, NVE = non-verbal element, PI = pseudo-incorporation, PL = plural, PRS = present tense, PST = past tense, PP = past participle, RP = referential phrase, RRG = Role and Reference Grammar



plex predicates consisting of a light verbal head and a non-verbal element (NVE). Such complex predicates are called ‘light verb constructions’ (LVCs); two example LVCs are shown in (1). A basic characteristic of LVCs is that the verb is semantically light, whereas the predicational content is mainly contributed by the NVE.

- (1) a. *šir*        *dâdan*  
       milk        give.INF  
       ‘to breastfeed’  
    b. *gul*        *xordan*  
       deceit     eat.INF  
       ‘to be deceived’

The different components of an LVC do not form a morphological word. In (2), the NVE *guš* ‘ear’ is separated from the light verb (L.V) by the future tense auxiliary *xâhad*. Persian LVCs show a degree of syntactic flexibility, which speaks against an analysis in terms of compounding or morphological incorporation.

- (2) *Be man guš xâhad kard.*  
       to me ear will do.PRS  
       ‘S.he will listen to me.’ (Mohammad & Karimi 1992: 197)

Besides LVCs, Persian frequently licenses bare uses of undergoer arguments. *Goošt* ‘meat’ – in (3) – is a nominal stem, rid of any functional morphology (e.g. number marking or case marking). The example in (3) superficially looks like the LVC in (1b). In both cases, the verb *xordan* is combined with a bare noun.

- (3) *goošt*        *xordan*  
       meat        eat.INF  
       ‘to meat-eat’ (i.e. ‘meat-eating’)

Some authors (e.g. Ghomeshi & Massam 1994, Vahedi-Langrudi 1996, Mahmoodi-Bakhtiari 2018: 295) argue that the examples in (1) and (3) exemplify the same kind of complex predicate. Others (e.g. Mohammad & Karimi 1992, Lazar 1992, Megerdoomian 2012, Modarresi 2014) argue that the examples in (1) and (3) have different properties and therefore instantiate different kinds of complex predicates. Whereas (1) represents light verb constructions, the example in (3) is subsumed under the notion of ‘pseudo-incorporation’ (the notion goes back to Massam 2001).

The current paper adopts the position that the examples in (1) and (3) represent different types of complex predicates. Within the paper, I present evidence in support of this view. First, I will show that the two types of complex predicates differ remarkably with regard to their nominal element. Pseudo-incorporation is restricted to bare nouns, whereas LVC-formation is not. Second, I argue that the two types of complex predicates differ with respect to their semantic composition. At the same time, the two types of complex predicates show clear similarities. The most important one is that the nominal element is an argument of the respective verb.

The whole discussion of differences and similarities between the two types of complex predicates has consequences for the syntactic analysis of LVCs. Within RRG, LVCs are usually analyzed in terms of nuclear cosubordination (e.g. Saeedi 2009, 2016, 2017). Such an analysis, as I will show, has to be rejected as it leads to wrong predictions. Instead, I analyze LVCs syntactically like regular argument-predicate constructions as in (4). Thus, I propose that there is no essential syntactic difference between regular argument-predicate constructions (4), pseudo-incorporation (3) and light verb constructions.

- (4) *in*        *goošt-râ*        *xordan*  
 DEM    meat-ACC<sup>2</sup>        eat.PRS  
 ‘eat the/this meat’

The structure of the paper is as follows: In Section 2, I start by briefly introducing the notion of a light verb construction. The treatment of LVCs within RRG is the topic of Section 3. I review the basics of Saeedi’s (2009, 2016, 2017) analysis and show which predictions can be derived from it. Section 4 investigates the grammatical properties of the non-verbal element of Persian LVCs. The syntactic flexibility of LVCs is addressed in Section 5. In this section, I discuss Karimi-Doostan’s analysis of Persian separable (i.e. syntactic flexible) LVCs and argue that he makes wrong predictions with respect to which NVEs can be separated from their LV. In Section 6, I argue for a relationship between syntactic flexibility and semantic compositionality. Finally, I propose a revised syntactic analysis of Persian LVCs in Section 7.

---

<sup>2</sup> There exists no consensus on whether the case marker *-râ* is a bound morpheme or not, Yousef (2018) describes it as being an enclitic particle, whereas Ghomeshi (1997) presents evidence for *-râ* being a phrasal affix. Throughout the paper, I adopt Ghomeshi’s analysis.

## 2. Light verb constructions

The notion ‘light verb’ goes back to Jespersen’s (1942) discussion of the grammar of Modern English. Jespersen introduced this notion for verbs like *give* or *have* occurring in constructions like *give a kiss* or *have a swim*. According to him, the verb only serves a grammatical function and is used for realizing tense and agreement morphology. Subsequent discussions have shown that light verbs are not semantically empty but only semantically reduced (e.g. Butt & Geuder 2001, Butt 2010, Butt & Lahiri 2013). The main predicational content is provided by the non-verbal element (e.g. *a kiss* or *a swim*); nevertheless, the light verb makes a (often subtle) semantic contribution to the LVC. This can be shown by contrasting the two Persian LVCs in (5). In both LVCs, the NVE *sedâ* ‘sound’ is used, the two only differ with respect to the light verb. The choice of light verb results in a subtle difference in meaning. *Sedâ kardan* (6a) can be used for situation in which a sound is produced intentionally, whereas *sedâ dâdan* (6b) cannot. If the LV would be semantically empty, a contrast like in (6) would be unexpected.

- (5) a. *sedâ dâdan*                      b. *sedâ kardan*  
 sound give.INF                      sound do.INF  
 ‘to produce a sound’                      ‘to produce a sound’

- (6) a. *Baçče amdan sedâ kard.*  
 child intentionally sound do.PST  
 ‘The child produced a sound intentionally.’ (Fleischhauer & Neisani 2020: 71)
- b. *#Baçče amdan sedâ dâd.*  
 child intentionally sound give.PST  
 ‘The child produced a sound intentionally.’ (Fleischhauer & Neisani 2020: 7)

According to Mohammad & Karimi (1992: 195), contemporary Persian possesses a limited number of lexically full verbs (around 115). Others, e.g. Samvelian (2018: 256), mention higher numbers (around 250) but state that only around half of them are still in use. Light verb constructions compensate for the lack of full verbs and fill a lexical gap. The verbs used as light verbs show a dual nature as they are also used as heavy verbs,

which is illustrated by the examples in (7).<sup>3</sup> In (7a), *dādan* ‘give’ denotes an event of giving: the book is transferred from the referent of *mard* ‘man’ to the referent of *zan* ‘woman’. In contrast, (7b) does not denote an event of giving; rather the LVC denotes an event of producing a sound. In this particular case, the produced sound is not directed towards someone, thus the situation cannot be interpreted as involving a transfer of a sound from a sender to a recipient. Most importantly, in case of a heavy verb, the verb determines the denoted situation. In case of an LVC, on the other hand, the denoted situation is dependent on the NVE. The NVE provides the semantic core of the LVC, whereas the light verb is its grammatical head.

- (7) a. *Ān mard be ān zan yek ketāb dād.*  
 DEM man to DEM woman INDEF book give.PST  
 ‘The man gave a book to the woman.’  
 b. *Motor sedā dād.*  
 engine sound give.PST  
 ‘The engine produced a sound.’

The non-verbal element is not restricted in its lexical category but can be a noun (7b), an adjective/participle (8a) or a prepositional phrase (8b).<sup>4</sup> For a more extensive overview on the combinatorial possibilities, the reader is referred to Dabir-Moghadam (1997) and Karimi (1997) respectively. In the remainder, I focus on the first type of LVCs, leaving LVCs with non-nominal NVEs aside.

- (8) a. *dekor kārđān* b. *be donya amāđān*  
 annoyed do.INF to world come.INF  
 ‘to annoy’ ‘to be born’  
 (Dabir-Moghadam 1997: 31)

In the next section, I introduce previous syntactic analyses of LVCs proposed in the framework of Role and Reference Grammar.

<sup>3</sup> I use the notions ‘light verb’ and ‘heavy verb’ for different uses of lexically full verbs. A ‘heavy verb’ is a use of a full verb, in which it retains its full lexical meaning and provides the main predicational content. A ‘light verb’ is, accordingly, the use of a lexically full verb, in which it does not provide the main predicational content.

<sup>4</sup> In Indo-Aryan languages such as Hindi/Urdu and Bengali, the non-light element can also be a verb (e.g. Butt 1995, 2010, Butt & Lahiri 2013).

### 3. LVCs within RRG

Within Role and Reference Grammar (RRG), complex constructions – complex sentences, complex referential phrases (RPs; Van Valin 2008) but also complex predicates – are analyzed in terms of juncture and nexus types. The three nexus relations are coordination, subordination and cosubordination, whereas the juncture types correspond to the three basic layers: nuclear juncture, core juncture and clausal juncture. Coordination consists in joining two units of equal status and size (Van Valin 2005: 183). Subordination refers to the embedding of one unit into another (Van Valin 2005: 183), whereas cosubordination is somehow in between; as Van Valin (2005: 187) puts it: “units of equal size are joined together in a coordination-like relation, but share some grammatical category”. Sharing of a grammatical category – also referred to as ‘operator sharing’ – is mandatory in cosubordination but not in coordination. In the case of coordination and cosubordination, the linkage is necessarily symmetrical, meaning that the joined units are of equal size. Asymmetrical linkage is only possible in case of subordination. An example of asymmetrical linkage is the sentence *That she arrived late shocked everyone* (Van Valin 2005: 199); the clause *she arrived late* is subordinated to the core *shocked everyone* and functions as a core argument.

There is no uniform treatment of LVCs within RRG. Romero-Méndez (2007) analyzes Spanish LVCs as a core juncture. Saeedi (2009, 2016, 2017) – for Persian –, Nolan (2014) – for Modern Irish – and Staudinger (2018) – for Spanish and French – propose an analysis in terms of nuclear juncture. More specifically, the authors propose an analysis in terms of nuclear cosubordination. Under such an analysis, two nuclei are combined to form a single complex nucleus. Applying such an analysis to the Persian LVC *sedâ dâdan* ‘to produce a sound’ gives the syntactic tree in Figure 1.

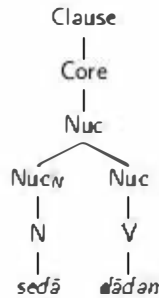


Fig. 1: Syntactic tree of the LVC *sedâ dâdan* 'to produce a sound' following the nuclear cosubordination analysis.

In her analysis of Persian LVCs, Saeedi (2016: 171) argues that the light verb fuses with the NVE "through morphological derivation to generate the new verb of the sentence or core [...], which acts as one single word, or verb." Saeedi treats LVCs as morphological compounds. Her analysis faces a serious problem. LVCs do not behave like morphological compounds; rather the two components of an LVC behave like two syntactically independent elements. For example, the LV can be separated from the NVE by functional (e.g. future tense auxiliary) as well as lexical material (e.g. attributive adjective or relative clauses). Saeedi accounts for this by assuming that the NVE is not just a nominal nucleus but can be an NP as well. Thus, she is speaking of NP-verb compounds rather than noun-verb compounds. This contradicts her analysis of Persian LVCs in terms of nuclear cosubordination since cosubordination is a symmetrical linkage type. In her final analysis (Saeedi 2009: 287) she assumes the embedding of an NP under a PRED node, as schematically represented in Figure 2.

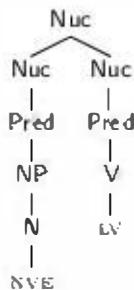


Fig. 2: Schematic representation of Saeedi's syntactic analysis of Persian LVCs

In the following chapters, I will present evidence against Saeedi's analysis. I will show that Saeedi's notion of compound is too broad and does not do justice to the language data. In addition, I will show that the grammatical evidence argues against an analysis in terms of nuclear cosubordination but requires treating the NVE as a core argument of its light verb.

#### 4. Grammatical properties of the NVE

In this section, I want to address two questions. First, is the NVE really a nucleus<sub>RP</sub> or can it host more structure? The question can be paraphrased as 'Are NVEs always bare nouns or can they host functional morphology?' This question is investigated in Section 4.1. The second question, investigated in Section 4.2, is whether the NVE fills an argument position of the light verb or not.

##### 4.1 Bare and non-bare nouns

Persian nouns are frequently used bare. This holds for NVE's (9a) as well as for core arguments of heavy verbs. In (9b), *qazā* 'food' is the undergoer argument of *xordan* 'eat', which is used as a heavy rather than a light verb. This is an instance of pseudo-incorporation.

- (9) a. *Motor*            *sedā*    *dād.*  
 engine            sound    give.PST  
 'The engine produced a sound.'
- b. *Bačče-hā*        *qazā*    *xord-and.*  
 child-PL        food    eat.PST-3PL  
 'The children were food-eating' (lit. 'The children ate food.')

Pseudo-incorporated nouns (PI-nouns) show similar properties cross-linguistically (for an overview see Borik & Gehrke 2015), among which are: (i) PI-nouns are number neutral, (ii) PI-nouns are discourse opaque and (iii) PI-nouns obligatorily show narrow scope with respect to scope bearing elements (e.g. modals or negation). For reasons of space, I only illustrate the second property but refer the reader to Megerdooian (2012) as well as Modaresi (2014, 2015) and Krifka & Modaresi (2016) for a more detailed discussion of pseudo-incorporation in Persian. The example in (10a) demonstrates that the PI-noun *sib* 'apple' does not introduce a discourse referent. Similarly, (10b) shows that *ketāb* 'book' does not serve

as an antecedent for (zero) anaphora.<sup>5</sup> Furthermore, the pseudo-incorporated nouns are number neutral, as the English translation indicates.

- (10) a. *Mân sib xârid-âm. #Xeili xošmazeh ast.*  
 1SG apple buy.PST-1SG very tasty be.PST.3SG  
 ‘I bought an apple/apples. It was very tasty.’ (Modaressi 2014: 25)  
 b. *Ali ketâb xârid va #oon-o xoond.*  
 Ali book buy.PST.3SG and DEM-ACC read.PST.3SG  
 ‘Ali bought a book/books and read it.’ (Modaressi 2014: 26)

The properties attributed to PI-nouns are also found with the nominal NVEs. The NVE *sedâ* ‘sound’ does not introduce a discourse referent (11).

- (11) a. *Âbgarmkon sedâ dâd. #Ân (sedâ) boland bud.*  
 boiler sound.give.PST DEM sound loud be.PST.3SG  
 ‘The boiler made a sound. It was loud.’  
 b. *Âbgarmkon sedâ dâd, #sedâ mesle raadobarq.*  
 boiler sound.give.PST sound like thunderstorm  
 ‘The boiler produced a sound like a thunderstorm.’  
 (Fleischhauer & Neisani 2020: 57)

It seems as if the nominal NVE behaves like a pseudo-incorporated noun, which might lead to the conclusion that pseudo-incorporation is not different from LVC-formation (such a position is – as already mentioned in the introduction – defended by various authors). However, the two differ in two important respects. First, there is a crucial difference with respect to the verb. *Dâdan* ‘give’ in (9a) is used a light verb, whereas *xordan* ‘eat’ in (b) is not. *Sedâ dâdan* does not mean ‘to give someone a sound’, thus the verb is semantically reduced. *Xordan*, on the other hand, simply means ‘to eat’ in (9b); the denoted situation is only determined by the verb but not by the noun *qazâ* ‘food’. Thus, *qazâ* is pseudo-incorporated into a heavy verb rather than a light verb.

A second difference obtains with respect to the nominal element. The pseudo-incorporated noun is necessarily bare; nominal morphology blocks pseudo-incorporation. A referential noun is interpreted as a regular macro-role core argument of the verb. The case marked noun *qâza-râ* ‘the food’

---

<sup>5</sup> Modaressi (2014, 2015) shows that PI-nouns are not discourse opaque but ‘discourse translucent’ (a term going back to Farkas and de Swart 2003); under certain conditions, they introduce a discourse referent, which can be picked up anaphorically.



in (12) is discourse transparent and can be picked up anaphorically. Non-referential undergoer arguments of heavy verbs are semantically incorporated, whereas referential ones are not.

- (12) *Mân* *Dišâb* *qâza-râ* *xord-âm* *vâ* *kâmi*  
 1SG last-night food-ACC eat.PST-1SG and some  
*âz* *an-ra* *hâm* *be gorbe* *dad-âm.*  
 of it-ACC also to cat give.PST-1SG  
 ‘Last night I ate the food and gave some of it to the cat too.’  
 (Megerdoomian 2012: 186)

Saeedi (2009: 76) proposes that nominal NVEs – very much like pseudo-incorporated nouns – can neither take one of the markers of indefiniteness (*ye* or *-i*) nor the plural marker *-hâ*. Contrary to her claim, functional morphology is licensed at the nominal NVE. As (13) shows, the NVE *sedâ* ‘sound’ can be marked for number (plural) as well as indefiniteness. Saeedi (2009: 76) also claims that NVEs are non-referential but *sedâ*, in (13), is used referentially as demonstrated by the fact that it can be picked up anaphorically.

- (13) *Âbgamkon* *sedâ-hâ-i* *dâd.* *Xeili boland bud-and*  
 boiler sound-PL-INDEF give.PST very loud is.PST-3PL  
 ‘The boiler produced some [specific] sounds. They were loud.’  
 (Fleischhauer & Neisani 2020: 57)

We can identify three predicational construction types, each of which have specific morphosyntactic as well as semantic properties. Pseudo-incorporation constructions (PI-constructions) consist of a non-referential undergoer argument and a heavy verb; the undergoer argument is semantically incorporated into the verb. A morphosyntactic requirement is that the undergoer argument is realized as a nucleus<sub>RP</sub>, i.e. bare noun. Regular predicate-argument constructions consist of a referential undergoer argument and a heavy verb. The undergoer argument is a full RP and cannot be used bare. This is the stereotypical predicate-argument relation as it is found in standard compositional patterns. LVCs, finally, consist of a NVE and a light verb. The NVE can be referential and is realized as a full RP. The properties of the three predicational construction types are summarized in Table 1.

Table 1: Summary of semantic and morphosyntactic properties of the predicational construction types

Predicational construction type	Noun	Verb
pseudo-incorporation construction	semantics: non-referential undergoer argument morphosyntax: nucleus <sub>RP</sub> (bare noun)	heavy verb
regular predicate-argument construction	semantics: referential undergoer argument morphosyntax: RP	heavy verb
light verb construction	semantics: no restriction with respect to referentiality morphosyntax: RP or bare noun	light verb

With respect to the syntactic analysis of Persian LVCs, it follows that they cannot be analyzed in terms of nuclear cosubordination. Cosubordination is a symmetrical linkage relation; the joined units are of equal size. However, as the discussion revealed, this does not turn out to be the case for Persian LVCs. The NVE is not a nucleus<sub>RP</sub> but (can be) a full-fledged RP.

## 4.2 Are NVEs arguments of their LVs?

The second question is whether the NVE is an argument of its respective light verb or not? If the answer is no, we would gain a crucial difference between LVCs on the one hand and the other predicational types. Evidence for the argument status of the nominal NVE would be if it is syntactically treated like regular arguments. To answer the above mentioned question, I compare NVEs with the undergoer arguments of transitive predicates. Two properties are crucial for our discussion: (i) case marking and (ii) behavior under passivization.

Persian shows definiteness-based differential object marking (e.g. Bosson 1985, Lazar 1992, Ghomeshi 1997) and restricts the case marker *-râ* to those undergoer arguments having definite reference.<sup>6</sup> By comparing

<sup>6</sup> Aissen 2003 shows that definiteness is not the only feature determine case marking in Persian. A recent overview in different analyses of *-râ* is found in Samvelian (2018).

(14a) and (b), one sees that the presence of *-râ* correlates with definite reference of the undergoer argument.

- (14) a. *Ân mard be ân zan yek ketâb*  
 DEM man to DEM woman INDEF book  
*dâd.*  
 give.PST  
 ‘The man gave a book to the woman.’
- b. *Ân mard ketâb-râ be ân zan*  
 DEM man book-ACC to DEM woman  
*dâd.*  
 give.PST  
 ‘The man gave the book to the woman.’

Megerdooomian (2012: 193ff.) argues that NVEs normally cannot take the accusative case marker *-râ*. Exceptions are, according to her, modified NVEs. Karimi-Doostan (2011) as well as Samvelian & Faghiri (2016) show that case marking on NVEs is not dependent on modification. The noun *latme* ‘damage’ in (15) takes the accusative case marker, although it occurs without further modifiers. The NVE is preceded by the demonstrative determiner *în*, which obligatorily requires accusative case marking on the undergoer argument of transitive predicates.

- (15) *Tegarg-e diruz în latme-râ be bâq-e*  
 hail-EZ yesterday DEM damage-ACC to garden-EZ  
*man zad.*  
 1SG hit.PST  
 ‘The yesterday’s hail caused this damage to my garden.’  
 (Karimi-Doostan 2011: 71; slightly adapted)

Examples like the one in (15) speak in favor of the view that *latme* is the undergoer argument of the light verb. A second piece of evidence supporting such a view is provided by passivization (a similar argument is proposed by Romero-Méndez 2007: 15 for Spanish NVEs). Persian expresses passive by a combination of the past participle of a full verb and the auxiliary *šodan* ‘become’. In Persian, passivization is restricted to verbs having two macrorole core arguments, of which the lower one is realized as the single macrorole core argument of the passivized verb (cf. Mace 2003: 128). The examples in (16) show that the LVC *mesâl zadan* ‘to give an example’ (lit. ‘example hit’) can be passivized. The crucial fact is that the NVE *mesâl* ‘example’ is the subject of the passive sentence.

- (16) a. *Moâlem e âlmâni ye mân mesâl*  
 teacher EZ German EZ 1PL example  
*e xub-i zad*  
 EZ good-INDEF hit.PST  
 ‘Our German teacher gave a good example.’
- b. *Mesâl e xub-i zade šod*  
 example EZ good-INDEF hit.PP become.PST.3SG  
 ‘A good example was given.’

Samvelian (2018: 262) also states that NVEs can become the subject under passivization and concludes “the nominal element of the CP [complex predicate] has exactly the same syntactic status as a bare direct object.” Support for this view is gained by the fact that even NVEs, which resist case marking by *-râ*, can be realized as the single macrorole argument of a passive. An illustrative example – taken from Samvelian (2018: 262) – is shown in (17). The LVC is *ahammīyat dâdan* ‘give importance to’, consisting of the NVE *ahammīyat* ‘importance’ and the LV *dâdan* ‘give’. Adding the case marker *-râ* to *ahammīyat* would turn (17a) into an ungrammatical sentence.

- (17) a. *Matbu’ât be in mas’ale xeili ahammīyat*  
 press to DEM issue very importance  
*dâd-and*  
 give.PST-3PL  
 ‘The press gave much importance to this issue.’
- b. *be in mas’ale xeili ahammīyat dâde*  
 to DEM issue very importance give.PP  
*šod*  
 become.PST.3SG  
 ‘Much importance was given to this issue.’

Samvelian’s analysis is in line with Vahedi-Langrudi’s (1996) and Müller’s (2010: 633) view that NVEs “are subcategorized arguments of the respective light verbs”. I adopt this view assuming that the NVE is a core argument of its light verb. This does not entail that the NVE is always a macrorole core argument. Certain LVCs realize a different argument than the NVE as the undergoer argument.<sup>7</sup> Example (18) shows the LVC *šekast*

<sup>7</sup> Samvelian & Faghiri (2016: 215) argue for a relationship between case marking of the NVE and the agentivity of the LVC. They propose that NVEs of agentive LVCs can take – under certain conditions – *-râ*, whereas NVEs of non-agentive LVCs never do.

*dâdan* ‘to defeat’ (lit. ‘defeat give’); it is not the NVE *šekast* but the RP *došman* ‘enemy’ which receives accusative case marking. Thus, *dâdan* ‘gives’ takes three arguments: *ânhâ* ‘they’ is the actor argument, *došman* ‘enemy’ is the undergoer argument and *šekast* is a non-macrorole argument. *Šekast* is realized differently from regular non-macrorole core arguments of ditransitive predicates, which usually take a preposition like *be* ‘to’ (cf. (19)). One possible analysis could be that NVEs are always realized as direct core arguments, i.e. they are either unmarked or take accusative case marking.

(18) *Ân-hâ došman-râ šekast dâd-and.*  
 DEM-PL enemy-ACC defeat give.PST-3PL  
 ‘They defeated the enemy.’

(19) *Ân mard be ân zan yek ketâb dâd.*  
 DEM man to DEM woman INDEF book give.PST  
 ‘The man gave a book to the woman.’

For the current paper, I adopt the view that NVEs are direct core arguments of their LVs. To substantiate this position, further work analyzing a larger sample of Persian LVCs is planned. Under a view such as the one adopted in the current paper, Persian allows the realization of three direct core arguments: one actor and two non-actor arguments. Among the non-actor arguments, the NVE is usually placed in preverbal position, the other argument – irrespective whether it is the undergoer argument or not – precedes the NVE.

In the next section, I discuss the syntactic flexibility of Persian LVCs, showing that the NVE can be separated from the LV and – at the same time – can linearly precede the other arguments.

## 5. Syntactic flexibility of Persian LVCs

The last section has shown that NVEs license functional morphology. Certain markers, e.g. the plural marker *-hâ* as well as the indefiniteness marker *-i*, are suffixed to the noun and therefore intervene between the nominal stem and the light verb. Beside functional elements – irrespective whether bound (*-hâ* or *-i*) or not (the future tense auxiliary *xâhad* (2)) – lexical material can also intervene between the two components of an LVC. An example is shown in (20); the attributive modifier *vez vez* ‘droning’ is placed between the NVE *sedâ* and the light verb *dâdan*. Attributive modifiers are realized postnominally and require a linking element, which

is called ‘ezâfe’.<sup>8</sup> Thus, beside functional morphology, the nominal NVE also licenses restrictive modifiers.<sup>9</sup>

- (20) *Motor sedâ ye vez vez dâd.*  
 engine sound EZ droning give.PST  
 ‘The engine made a droning sound.’

Attributive modification of the NVE is one of the properties Karimi-Doostan (2011) uses for discrimination between separable and non-separable Persian LVCs. Karimi-Doostan (2011: 71f.) classifies light verb constructions as separable if they show the properties in (21). From his discussion, it is not obvious whether an LVC needs to show all of these properties to be classified as ‘separable’ or whether a subset of these properties is sufficient.

- (21) (i) The NVE can be modified by an adjective,  
 (ii) the NVE can function as a direct object,  
 (iii) the NVE can be relativized,  
 (iv) the NVE can be focused on by *Wh*-interrogatives,  
 (v) the NVE can be scrambled.

The first two properties have already been discussed above and in the last section. The third property is illustrated in (22). The particle *ke* introduces a relative clause, which modifies the NVE *latme* ‘damage’.

- (22) *Latme-i ke tegarg be bâq-hâ zad*  
 damage-INDEF that hail to garden-PL hit.PST  
*jobrân nâpazir ast.*  
 irretrievable is  
 ‘The damage caused by the hail to the gardens is irretrievable.’  
 (Karimi-Doostan 2011: 71; slightly adapted)

<sup>8</sup> See Ortman (2002: 66) and Fleischhauer & Neisani (2020: 61f.) for a discussion of the semantics of the *ezâfe*-morpheme.

<sup>9</sup> Megerdoomian (2012) argues that attributive modifiers within LVCs function adverbially, i.e. having the whole LVC rather than just the NVE in its scope. This is – as the author argues – different in case of pseudo-incorporated nouns, indicating that the NVE is not an internal argument of the light verb, whereas the pseudo-incorporated noun is an internal argument of the incorporating verb. As argued above *vez vez* ‘droning’ functions as an attributive rather than adverbial modifier, contradicting Megerdoomian’s claim; see Section 6 as well as Fleischhauer & Neisani (2020) for further evidence supporting this view.

Property (iv) – the NVE can be focused on by *Wh*-interrogatives – is illustrated by the example in (23).

- (23) *Ali ĉe latme-i be šomâ zad?*  
 Ali what damage-INDEF to 2SG hit.PST  
 'What loss did Ali cause to you?'  
 (Karimi-Doostan 2011: 71; slightly adapted)

The last property – scrambling of the NVE – is shown in (24). The RP *in latme-râ* 'this damage' is in sentence initial position and therefore precedes all other arguments.

- (24) *In latme-râ degarg-e diruz be*  
 DEM damage-ACC hail-EZ yesterday to  
*bâġ-e mân zad.*  
 garden-EZ 1SG hit.PST  
 'Yesterday's hail caused this damage to my garden.'  
 (Karimi-Doostan 2011: 72; slightly adapted)

The data show that LVCs form a rather loose syntactic construction. This is not specific to LVCs, since pseudo-incorporated nouns also show a certain degree of syntactic flexibility. As (25) reveals, the bare noun *ketâb* 'book' can be scrambled and realized in sentence initial position. Such a construction results in a contrastive reading, as the English translation suggests. Morphologically, *ketâb* is singular but – as also revealed by the translation – it receives (in this particular context) a plural interpretation. Thus, *ketâb* is number neutral, which gives evidence for the view that it is pseudo-incorporated.

- (25) *KETÂB Parviz barâ Kimea xarid.*  
 book Parviz for Kimea buy.PST  
 'Parviz bought BOOKS for Kimea (as opposed to journals)'.  
 (Karimi 2018: 165; slightly changed)

Karimi-Doostan proposes a correlation between the separability of the two components of an LVC and the nominal type of NVE. He argues that the properties in (21) are restricted to NVEs which belong to the class of eventive nouns. Other types of nominal NVEs – verbal nouns and non-eventive nouns – are, according to him, not separable from the LV. The three noun types can be distinguished on the basis of two properties: (i) what kind of entity does the noun refer to (eventuality vs. object)?, and (ii) does the

noun show nominal features (e.g. is it compatible with the demonstrative determiner)?

The examples in (26) show that the nouns *sedâ* ‘sound’ (a) and *jâde* ‘street’ (b) can be preceded by the demonstrative determiner *in*, whereas *anjâm* ‘performing’ (c) cannot. Thus, the first two nouns show nominal features, whereas the latter does not.

- (26) a. *in sedâ*      b. *in jâde*      c. \**in anjâm*  
 DEM sound      DEM street      DEM performing  
 ‘this sound’      ‘this street’

Eventive nouns – in contrast to non-eventive ones – can be temporally located. This is shown for the two nouns *sedâ* and *anjâm* in (27). The noun *jâde* cannot be temporally located.

- (27) a. *Sedâ*    *ye*      *tir*      *andâzi*    *dar*      *âsemân*  
 sound    EZ      bullet    shooting    in      sky  
*e*      *tehrân*    *sobh*      *e*      *jomʔe*    *ettefâq*  
 EZ      Teheran    morning    EZ      Friday    happen  
*ofiâd*  
 fall.PST  
 ‘The sound of shooting in the sky of Teheran occurred on Friday morning.’ (Fleischhauer & Neisani 2020: 73)
- b. *Anjâm*      *e*      *tarh*      *dar*      *tul*  
 performing    EZ      sketch    in      length  
*e*      *vaq̄t*      *e*      *esterâhat*    *surat*    *gereft*  
 EZ      time      EZÂFE    rest      form      get.PST  
 ‘The performance of the sketch took place during the break (of the game).’ (Fleischhauer & Neisani 2020: 52)

The outcome of the two test criteria is summarized in Table 2-2. *Sedâ* qualifies as being an eventive noun; *anjâm* shows the characteristics of a verbal noun and *jâde* is a non-eventive noun. Therefore, *sedâ* should be separable from its LV. This is in fact borne out, as the example in (20) shows.



Table 2: Classification of nouns (based on Karimi-Doostan 2011)

Noun type	refers to...	nominal features?	Example
Eventive nouns	Eventuality	yes	(26a)/(27a)
Verbal nouns	Eventuality	no	(26b)/(27b)
Non-eventive nouns	Object	yes	(26c)

Karimi-Doostan's analysis predicts that the non-eventive noun *jâde* – if used as an NVE – should not be separable from its LV. In fact, *jâde* can be separated from the LV *kešidan* 'pull' by an attributive modifier (28). Given the interpretation of the modified LVC, there is no doubt that the adjective *sangi* 'stony' modifies *jâde* and does not function as an adverbial modifier of the whole LVC.

- (28) *jâde ye sangi kešidan*  
 street EZ stony pull.inf  
 'to build a stony road'

In addition, (29) shows that *jâde* can be modified by an adjunct PP, which is placed between the NVE and the LV. Furthermore, the NVE licenses functional morphology.

- (29) *In âqâyan jade-hâ-i be forudgâh kešid-and.*  
 DEM men road-PL-INDEF to airport pull.PST-3PL  
 'These men built roads to the airport.'

Finally, *jâde* is realized as the undergoer argument of the LV *kešidan* since it can receive accusative case marking (30a) and becomes the single marcorole core argument under passivization (b).

- (30) a. *In âqâyan in jade-râ be forudgâh*  
 DEM men DEM road-ACC to airport  
*kešid-and.*  
 pull.PST-3PL  
 'These men built the road to the airport.'
- b. *Jâde-hâ-i be forudgâh kešide*  
 road-PL-INDEF to airport pull.PP  
*šod-and.*  
 become.PST-3PL  
 '(Some) roads to the airport were build.'

The syntactic flexibility of Persian LVCs is not – contrary to Karimi-Doostan’s proposal – dependent on the nominal type of NVE. Instead, as I will argue in the next section, separability is dependent on semantic compositionality.

## 6. Compositionality of Persian LVCs

Various researchers assume that light verb constructions are semantically non-compositional (e.g. Goldberg 2002, Family 2011, Modarressi 2014, Moezzi-pour & Ghandhari 2017; see Müller 2010 for arguments against a construction grammar analysis of Persian LVCs). The aim of the current section consists in presenting evidence in favor of a semantically compositional analysis of separable LVCs. To gain evidence for this view, I follow Nunberg et al.’s (1994) argument that ‘idiomatically combining expressions’ are a semantically compositional subtype of idioms. In fact, I will argue – in line with Samvelian & Faghiri (2016) – that separable LVCs are idiomatically combining expressions.

Nunberg et al. (1994) propose a distinction between ‘idiomatic phrases’ (e.g. *kick the bucket*) on the one hand and ‘idiomatically combining expressions’ (ICEs) on the other hand. ICEs – like English *pull strings* or *spill the beans* – count as phrasal idioms but – as the authors argue – are semantically compositional. The individual components of an ICE have an identifiable meaning and combine in a non-arbitrary way. This claim is not intended to mean that a speaker is capable of deriving the meaning of e.g. *spill the beans* based on the literal meaning of its components. Rather, the meaning of an ICE is composed on basis of a probable non-literal interpretation of its components. Thus, after the meaning of the various components of an ICE is identified, the meaning of the whole is derived compositionally.

In their discussion of ICEs, Nunberg et al. mention various LVCs without using that term. In the following discussion, I will demonstrate that LVCs show the same properties as ICEs and therefore qualify as being idiomatically combining expressions. A similar proposal is found in Samvelian & Faghiri (2016), although the authors integrate it into a Construction-based approach.

To support the compositional view on ICEs, Nunberg et al. present several pieces of evidence, of which I restrict the discussion to the following:

- (i) ICEs come in families,
- (ii) ICEs license internal modification.

Regarding ICE families, Nunberg et al. (1994: 504) write that we “find families of idioms, where for instance, the same verb can occur in different environments to form distinct, but semantically related, idioms.” For illustration, they cite examples like *hit the hay/ hit the sack* or *take a leak/ take a piss/ take a crap* (p. 504). With respect to LVCs, one can define a family as a set of LVCs headed by the same light verb and exemplifying the same interpretational pattern. Two families headed by the Persian light verb *kešidân* ‘pull’ are listed in (31).<sup>10</sup>

- (31) a. *šigar kešidân* ‘smoke cigarettes’ (lit. ‘cigarette pull’), *hāšiš kešidân* ‘smoke hash’ (lit. ‘hash pull’), *qālyan kešidân* ‘smoke a hooka’ (lit. ‘hookas pull’), *pip kešidân* ‘smoke a pipe’ (lit. ‘pipe pull’)
- b. *jāde kešidân* ‘build a road’ (lit. ‘road pull’), *divār kešidân* ‘build a wall’ (lit. ‘wall pull’), *nārde kešidân* ‘put up a fence’ (lit. ‘fence pull’)
- (Family 2011: 13)

The members of the first family in (31a) can be paraphrased as ‘to smoke N’, whereas the members of the second family (31b) are roughly paraphrased as ‘to build N’. Systematic interpretation patterns arise depending on the type of noun (in (31) the interpretation depends on the difference between a noun denoting an ingredient used for smoking – a smokeable substance or an instrument of smoking – and a noun denoting a build object). The existence of such families would be surprising, as Nunberg et al. (1994: 504) state, if each expression had an idiosyncratic meaning. Instead we find compositional patterns for different subtypes of the construction ‘N + *kešidân*<sub>LIGHT</sub>’; the compositional meaning of the construction depends on the semantic type of the NVE. This is an instance of what Löbner (2012) calls ‘subcompositionality’. He states that “[A] syntactic construction is subcompositional if there is no uniform rule of semantic composition for it” (Löbner 2012: 224). Rather, each subcompositional pattern requires its own rule of semantic composition. Thus, if one advocates the view that LVCs are semantically compositional, then the very existence of LVC-families speaks in favor of subcompositionality.

---

<sup>10</sup> See Romero-Méndez (2007: 22ff.) for a discussion of some Spanish LVC-families. Family (2011, 2014) proposes an analysis of Persian LVC-families, which she calls ‘verbal islands’, within a constructional grammar approach. Thus, in her view the existence of ‘verbal islands’ does not provide evidence for a compositional analysis of LVCs.

The second piece of evidence supporting a compositional view is that ICEs license internal modification (a notion going back to Ernst 1981). In case of internal modification, only a part of the expression, rather than the whole expression, is in the scope of the modifier. Nunberg et al. (1994: 500) argue: “In order to modify part of the meaning of an idiom by modifying a part of the idiom, it is necessary that the part of the idiom have a meaning which is part of the meaning of the idiom”. An illustrative example, discussed by Nunberg et al. (1994: 500), is *leave no legal stone unturned*. As the authors state, this does not mean ‘legally leave no stone unturned’ but “that all legal methods are used” (Nunberg et al. 1994: 500).

Internal modification is also found in Persian LVCs. This is illustrated with respect to the light verb constructions *jâde kešidân* ‘build a road’ (lit. ‘road pull’), which belongs to the family listed in (31b). The attributive modifiers in (32) specify a property of the build road; in (a) it is the price of the road, whereas in (b) the modifier specifies the road’s trajectory.

- (32) a. *jâde ye gerâni kešidan*  
 road EZ expensive pull.INF  
 ‘to build an expensive road’  
 b. *jâde ye mârpič kešidan*  
 road EZ twisted pull.INF  
 ‘to build a twisted road’

A further telling example is shown in (33a). The adjective *masxare* ‘funny’ is realized as an attributive modifier of the NVE *sedâ* ‘sound’. The resulting interpretation is that the sound, which is produced by the dog, is funny. As (33b) shows, *masxare* can also be used adverbially. In this case, it is not placed post-nominally and it does not require the linking element. Crucially, the interpretation of (33b) is different from the one in (a). The sentence means ‘it was funny that the dog produced a sound’. It can be funny that the dog produced a sound, without the produced sound being funny (and vice versa). Thus, *masxare* really functions as an internal modifier in (33a) having only scope over the NVE, whereas it scopes over the whole LVC in (b). This is in clear opposition to Megerdooomian’s (2012: 196) claim that attributive modifiers of the nominal element within the NVE always function as adverbial modifiers of the whole LVC.

- (33) a. *Sag sedâ ye masxare-i dâd.*  
 dog sound EZ funny-INDEF give.PST  
 ‘The dog made a funny sound.’

- b. *Sag masxare sedâ dâd.*  
 dog funny sound give.PST  
 ‘Funnily, the dog made a sound.’

Nunberg et al. discuss further properties of ICEs – related to the syntactic separation of its components – which I am not going to discuss here. The relevant conclusion for us is that separable LVCs behave like ICEs showing that they are semantically compositional. The position adopted in the current paper is that LVCs are always compositional, whereas non-compositional N-V constructions are idiomatic expressions. Thus, I propose a distinction between compositional LVCs on the one hand and similarly looking but non-compositional constructions on the other hand.

## 7. A revised syntactic analysis of Persian separable LVCs

In Section 4, I concluded that the NVE of a separable LVC is realized as a direct core argument of its light verb. This makes separable LVCs syntactically similar to pseudo-incorporation constructions and regular predicate-argument constructions. Pseudo-incorporation is restricted to undergoer arguments, other bare nouns (e.g. actor arguments) do not show the properties of pseudo-incorporated nouns. The only difference between the three predicational construction types is that the NVE is not necessarily a macrorole core argument. The exact conditions under which the NVE is realized as a macrorole core argument is not clearly understood yet.

Since I am analyzing the NVE as a core argument of the LV, there does not seem to be any need for proposing a clause linkage analysis of Persian LVCs. Instead I propose a syntactic analysis as depicted in a tree like in Figure 3 for the sentence *Sag sedâ dâd* ‘The dog produced a sound’. Structurally, the NVE is represented as a regular core argument, its status as specifying the main predicational content is not indicated in the syntactic structure.

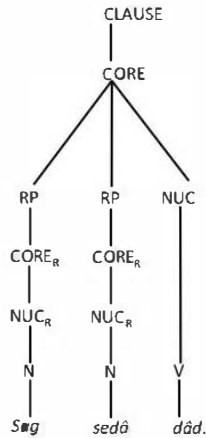


Fig. 3: Syntactic tree of the sentence *Sog sedâ dâd* ‘The dog produced a sound.’

Within the current analysis, I assume the same syntactic structure for LVCs, pseudo-incorporation constructions and regular predicate-argument constructions. This is in opposition to claims made by authors, like Megerdooomian (2012). She writes that “the distinct interpretations in the two constructions [i.e., pseudo-incorporation constructions and LVCs] clearly point to a difference in structure between predicates formed with a light verb and those composed of a thematic [i.e. heavy] verb” (Megerdooomian 2012: 189). Her argumentation is based on the determination of event structure and argument structure. Megerdooomian adopts Folli et al.’s (2005) position that “the light verb contributes eventive and aspectual properties to the complex predicate, the NV[E] determines aktionsart (or inner aspect) and substantive information, while both components contribute to the argument structure” (Megerdooomian 2012: 190). Heavy verbs, on the other hand, include “all of the mentioned properties of event, aspect, and argument structure, as well as the core meaning” (Megerdooomian 2012: 190). She claims that the NVE cannot be an argument of the LV since it contributes to aktionsart, argument structure as well as lexical content. Leaving the issue of argument structure aside, I want to comment on the composition of event structure and lexical content. Starting with event structure, I demonstrate that LVCs are not special in having a compositionally derived event structure.

Starting with Verkuyl (1972), it has become widely known that event structure is not determined by the predicate alone but that the predicate’s arguments and even adjuncts contribute to event structure as well. A

prominent case of compositionally derived event structure are incremental theme verbs (a notion going back to Dowty 1991) like *eat* and *drink*. Such verbs show variable telicity depending on the referential properties of its undergoer argument. The bare noun *pizza* in (34a) gives rise to an atelic reading, whereas the use of the RP *the pizza* results in a telic one (b). In its atelic use, *eat* is an activity predicate, whereas it functions as an active accomplishment in its telic use (Van Valin 2005: 47).<sup>11</sup> Filip (2011) argues that these verbs are unspecified for telicity, which highlights the fact that a certain aktionsart property – i.e., telicity – is only determined after the composition of the verb with its undergoer argument. Thus, the fact that event structure is compositionally derived does not demarcate LVCs from regular heavy uses of lexically full verbs.

- (34) a. *Carl ate pizza (for ten minutes).*  
 b. *Carl ate the pizza (in ten minutes).*

How do the interpretational differences between these three predicational construction types arise? Although the three predicational construction types have the same syntax, their semantic composition proceeds differently.<sup>12</sup>

I just want to briefly discuss the differences in the semantic composition of the three predicational construction types. Thus, the following discussion is rather schematic and illustrates the basic line of argumentation without going into the details of semantic composition (but see Fleischhauer & Neisani 2020 for a more detailed discussion of the semantic composition of Persian LVCs). The starting point for this discussion is the example in (35), which is an instance of regular predicate-argument construction. The verb *xordan* ‘eat’ has two direct core arguments. *Sib*

---

<sup>11</sup> Van Valin (2005: 47) proposes a lexical rule, which derives the active accomplishment use of creation/consumption verbs from an activity form. The contrast between atelic (activity) and telic (active accomplishment) uses of such verbs is in fact not only dependent on referential properties of the referent of the undergoer argument, for a discussion of further relevant properties see Van Valin & Latrouite (2014), Czardybon & Fleischhauer (2014), Fleischhauer & Czardybon (2016) as well as Fleischhauer & Gabrovská (2019).

<sup>12</sup> Saeedi treats pseudo-incorporation very much like the formation of LVCs and argues even that some LVCs “seem to form through the incorporation process” (Saeedi 2017: 404). This view is in direct opposition to the one advocated in the current paper, as I argue that the essential difference between pseudo-incorporation and the formation of LVCs consists in their respective mood of composition. In my view, LVCs do not result from semantically incorporating a nominal element within the light verb.

‘apple’ is the undergoer argument and combines with the verb via functional application. The resulting interpretation is that the referent of the actor argument denoted by *pesar* ‘boy’ eats the individual denoted by *sib*.

- (35) *Pesar sib-râ xord.*  
 boy apple-ACC eat.PST  
 ‘The boy ate the apple.’

An instance of pseudo-incorporation is shown in (36). The examples in (35) and (36) use the same heavy verb *xordan*, *sib* (in (35)) as well as *qazā* (in (36)) fulfill the selectional restrictions the verb imposes on its theme argument, i.e., denoting something edible. *Qazā* is used non-referentially and modifies the verb. The PI-construction *qazā xordan* is interpreted as denoting a specific type of activity, which is executed by the actor argument’s referent (‘food-eating’ rather than a different type of activity).

- (36) *Bačče-hā qazā xord-and.*  
 child-PL food eat.PST-3PL  
 ‘The children were food-eating’. (lit. ‘The children ate food.’)

Chung & Ladusaw (2004) argue that in the case of pseudo-incorporation, semantic composition does not proceed via functional application. They introduce a further compositional mood called ‘restrict’. ‘Restrict’ results in composing a predicate with a property-denoting expression and restricts the original function denoted by the predicate to a subdomain. Thus, *qazā* restricts the function EAT denoted by *xordan* to FOOD-EAT.<sup>13</sup> The use of ‘restrict’ is triggered by having a property-denoting expression realized as the argument of a heavy verb.

With respect to LVCs, Romero-Méndez (2007) made an RRG-specific proposal how this particular type of complex predicate is built up semantically. RRG proposes a system of logical structures, which are intended to capture grammatically relevant aspects of meaning. Among the grammatically relevant meaning components are aktionsart and causativity (see Van Valin 2005). Logical structures consists of a small set of operators (BECOME, INGR, CAUSE) and predicates (e.g. rain’, hit’, see’, tall’), which either function as argument or modifier of one of the operators.<sup>14</sup> With re-

<sup>13</sup> See Chung & Ladusaw (2004) for the formal details of the ‘restrict’ operation.

<sup>14</sup> CAUSE is an operator, which takes logical structures rather than predicates as its arguments.



spect to Spanish LVCs, Romero-Méndez (2007) proposes the following LVF-formation rules:

- (37) Formation of LVCs (based in Romero-Méndez 2007: 36):
- (i) The verb should provide the lexical template or part of it, and
  - (ii) the qualia structure of the nominal replaces part of the lexical template provided by the verb.

The light verb contributes a partial logical structure such as *do'* ( $x, [\text{pred}' (\ )]$ ). Part of the structure – the *pred'* ( $\ )$  component – is unspecified in the LV and is specified by the NVE. Thus, the logical structure of the LVC is derived by combining the LVC's components (Romero-Méndez proposes an analysis in terms of co-composition; e.g. Pustejovsky 2012). Within this proposal, the primary function of the light verb consists in introducing the aktionsart operators.

The basic idea of Romero-Méndez's proposal is in line with Fleischhauer & Neisani's (2020) compositional analysis of Persian LVCs. Fleischhauer & Neisani analyze light verbs as 'defective' verbs, which only come with a rudimentary lexical semantics. Light verbs – following Butt & Geuder (2001) – do not denote eventualities of their own but need combining with a NVE to form a full-fledged predicational structure. In line with the syntactic analysis proposed in this paper, the NVE is syntactically realized as a direct core argument; semantically it saturates a predicate variable (rather than an individual variable). This allows the NVE to contribute to the predicational content and not just introducing an individual over which the verb is predicating. Semantically, the NVE combines with the LV via regular functional application. Thus, pseudo-incorporation differs from the formation of LVC in terms of their underlying modes of composition. Pseudo-incorporation is achieved via the 'restrict'-operation, whereas the meaning of LVC are composed by functional application.

## 8. Conclusion

At this stage, I want to summarize the central claims put forward in this paper with respect to Persian predicational construction types (regular argument-predicate constructions, PI-constructions and LVCs). First, the three predicational construction types have the same syntax. In particular, I reject the view that the NVE is a nominal nucleus cosubordinated to a verbal nucleus and treat the NVE as a direct core argument of its light verb. Thus, the syntactic relation between the NVE and the LV is the same

than the syntactic relation between the non-actor direct core arguments of heavy verbs (irrespective of whether it is pseudo-incorporated or not). Second, the predicational construction types differ in their semantics. The semantic differences are due to different modes of composition (saturation or functional application in case of regular predicate-argument constructions and LVCs, and restrict in case of pseudo-incorporation) and different uses of the lexical full verbs (heavy vs. light uses). Thus, modelling the differences between regular predicate-argument constructions on the one hand and the two types of complex predicates on the other, an elaborate representational format for a richer semantic de-/composition is required. A promising step towards that direction is presented in Osswald & Kallmeyer's (2018) work, which combines RRG with frame semantics.

Various questions are still open for future work. Among those questions are: Under which conditions is the NVE realized as a macrorole argument? Answering the first question requires a corpus-based analysis to determine the grammatical properties of a huge number of Persian LVCs. This task is directly related to the second question: How is the argument structure (compositionally) derived? There exists an extensive literature on that issue (e.g. Butt 2010, Müller 2010), which should be taken into account in the development of an RRG perspective on it.

## Acknowledgments

I like to thank Robert D. Van Valin Jr. and Anja Latrouite for helpful comments on a previous version of the paper. I especially like to thank Mozghan Neisani for her help with the language data.

## References

- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21:435-483.
- Borik, Olga & Berit Gehrke. 2015. *An introduction to the syntax and semantics of pseudo-incorporation*. In Olga Borik & Berit Gehrke (eds.), *The syntax and semantics of pseudo-incorporation*, 1-43. Leiden/Boston: Brill.
- Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr.
- Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Stanford: CSLI Publications.

- Butt, Miriam. 2010. *The light verb jungle: still hacking away*. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex predicates: cross-linguistic perspectives on event structure*, 48–78. Cambridge: Cambridge University Press.
- Butt, Miriam & Wilhelm Geuder. 2001. *On the (semi)lexical status of light verbs*. In Norbert Corver & Henk van Riemsdijk (eds.), *Semi-lexical categories: On the content of function words and the function of content words*, 323–370. Berlin: Mouton.
- Butt, Miriam & Aditi Lahiri. 2013. Diachronic pertinacity of light verbs. *Lingua* 135: 7–29.
- Chung, Sandra & William Ladusaw. 2004. *Restriction and saturation*. Cambridge, MA: MIT Press.
- Czardybon, Adrian & Jens Fleischhauer. 2014. *Definiteness and perfectivity in telic incremental theme predications*. In Doris Gerland, Christian Horn, Anja Latrouite, & Albert Ortman (eds.), *Meaning and grammar of nouns and verbs*, 373–400. Düsseldorf: Düsseldorf University Press.
- Dabir-Moghaddam, Mohammad. 1997. Compound verbs in Persian. *Studies in the Linguistic Sciences* 27 (2): 25–59.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.
- Ernst, Thomas. 1981. Grist for the linguistics mill: Idioms and ‘extra’ adjectives. *Journal of Linguistic Research* 1: 51–68.
- Farkas, Donka & Henriette de Swart. 2003. *The semantics of incorporation: From argument structure to discourse transparency*. Stanford: CSLI Publications.
- Family, Neiloufar. 2011. Verbal islands in Persian. *Folia Linguistica* 45 (1): 1–30.
- Family, Neiloufar. 2014. *Semantic Spaces of Persian Light Verb Constructions: A Constructionist Account*. Leiden: Brill.
- Filip, Hana. 2011. *Aspectual class and Aktionsart*. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, 1186–1217. Berlin/New York: Mouton de Gruyter.
- Fleischhauer, Jens & Adrian Czardybon. 2016. The role of verbal prefixes and particles in aspectual composition. *Studies in Language* 40 (1): 176–203.
- Fleischhauer, Jens & Ekaterina Gabrovskaja. 2019. *Perfectivity and atelicity – The role of perfective aspect in aspectual composition*. In Emmanuelle Roussel, Adeline Patar & Rea Peltola (eds.), *Cross-linguistic per-*

- spectives on the semantics of grammatical aspect*, 97-126. Leiden: Brill.
- Fleischhauer, Jens & Mozghan Neisani. 2020. Adverbial and attributive modification of Persian separable light verb constructions. *Journal of Linguistics* 56: 45-85.
- Folli, Raffaella, Heidi Harley & Simin Karimi. 2005. Determinants of eventtype in Persian complex predicates. *Lingua* 115: 1365-1401.
- Ghomeshi, Jila. 1997. Topics in Persian VPs. *Lingua* 102.2-3: 133-167.
- Ghomeshi, Jila & Diane Massam. 1994. Lexical/syntactic relations without projections. *Linguistic Analysis* 24.3-4: 175-217.
- Goldberg, Adele. 2002. *Words by default: The Persian complex predicate construction*. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: Form-Function Incongruency and the Architecture of Grammar*, 84-112. Stanford: CSLI Publications.
- Jespersen, Otto. 1942. *A modern English grammar on historical principles, part VI, morphology*. Copenhagen: Ejnar Munksgaard.
- Karimi, Simin. 1997. Persian complex verbs: idiomatic or compositional. *Lexicology* 3: 273-318.
- Karimi, Simin. 2018. *Generative approaches to syntax*. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford handbook of Persian linguistics*, 161-204. Oxford: Oxford University Press.
- Karimi-Doostan, Gholamhossein. 2011. Separability of light verb constructions in Persian. *Studia Linguistica* 65 (1): 70-95.
- Krifka, Manfred & Fereshteh Modarresi. 2016. Number neutrality and anaphoric update of pseudo-incorporated nominals in Persian (and weak definites in English). *Proceedings of SALT* 26: 874-891.
- Lazard, Gilbert. 1992. *A grammar of contemporary Persian*. Costa Mesa, Cal./New York: Mazda Publishers.
- Löbner, Sebastian. 2012. *Sub-compositionality*. In Markus Werning, Wolfgang Hinzen & Edouard Machery (eds.), *The Oxford handbook of compositionality*, 220-241. Oxford: Oxford University Press.
- Mace, John. 2003. *Persian Grammar*. London: Routledge.
- Mahmoodi-Bakhtiari, Behrooz. 2018. *Morphology*. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford handbook of Persian linguistics*, 273-299. Oxford: Oxford University Press.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19: 153-197.
- Megerdoomian, Karine. 2012. The status of the nominal in Persian complex predicates. *Natural Language & Linguistic Theory* 30: 179-216.

- Modarresi, Fereshteh. 2014. *Bare nouns in Persian: Interpretation, grammar and prosody*. Berlin: Humboldt-Universität zu Berlin. (Doctoral dissertation.)
- Modarresi, Fereshteh. 2015. *Discourse properties of bare noun objects*. In Olga Borik & Berit Gehrke (eds.), *The syntax and semantics of pseudo-incorporation*, 189-221. Leiden: Brill.
- Mohammad, Jan & Simin Karimi. 1992. *Light verbs are taking over: Complex verbs in Persian*. In J. A. Nevis & V. Samiian (eds.), *Proceedings of the Western Conference on Linguistics*. Vol. 5, pp. 195-212. Fresno: California State University.
- Moezzi-pour, Farhad & Mina Ghandhari. 2017. *Concept structuring in Persian PP-centric complex predicates*. In Brian Nolan and Elke Die-drichsen (eds.), *Argument realisation in complex predicates and complex events*, 413-447. Amsterdam/ Philadelphia: John Benjamins.
- Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3): 601-655.
- Nolan, Brian. 2014. Complex predicates and light verb constructions in Modern Irish. *Spanish Journal of Applied Linguistics* 27 (1): 140-167.
- Nunberg, Geoffrey, Ivan Sag & Thomas Wasow. 1994. Idioms. *Language* 70.3: 491–538.
- Ortmann, Albert. 2002. *Kategorien des Nomens. Schnittstellen und Ökonomie*. Tübingen: Niemeyer.
- Osswald, Rainer and Laura Kallmeyer. 2018. *Towards a formalization of Role and Reference Grammar*. In Rolf Kailuweit, Eva Staudinger and Lisann Künkel (eds.), *Applying and expanding Role and Reference Grammar*, 355-378. Freiburg: FRIAS.
- Pustejovsky, James. 2012. *Co-compositionality in grammar*. In Markus Werning, Wolfram Hinzen & Édouard Machery (eds.), *The Oxford handbook of compositionality*, 371-384. Oxford: Oxford University Press.
- Romero-Méndez, Rodrigo 2007. Spanish light verb constructions: Copredication with syntactically formed complex predicates. (<http://linguistics.buffalo.edu/people/faculty/vanvalin/rrg.html>)
- Saeedi, Zari. 2009. *Persian light verb constructions: A Role & Reference Grammar account*. Dissertation: Trinity College Dublin.
- Saeedi Zari. 2016. *Complex predicates in modern Persian*. Sheffield: Equinox.
- Saeedi, Zari. 2017. *Nominal predication in Persian. A functional characterization*. In Brian Nolan and Elke Die-drichsen (eds.), *Argument realisation in complex predicates and complex events*, 373-412. Amsterdam/ Philadelphia: John Benjamins.

- Samvelian, Pollet. 2018. *Specific features of Persian syntax*. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford handbook of Persian linguistics*, 226-269. Oxford: Oxford University Press.
- Samvelian, Pollet & Pegah Faghiri. 2016. Re-thinking compositionality in Persian complex predicates. In *Proceedings of Berkeley Linguistic Society 39th Annual Meeting, February 16-17 2013*, 212-227. Berkeley: Berkeley Linguistics Society.
- Staudinger, Eva. 2018. *French and Spanish 'MAKE/GIVE + FEAR'-type LVCs – an RRG Constructional Account*. In Rolf Kailuweit, Lisann Künkel & Eva Staudinger (eds.), *Applying and expanding Role and Reference Grammar*, 237-261. Freiburg: FRIAS.
- Vahedi-Langrudi, Mohammad-Mehdi. 1996. *The syntax, semantics and argument structure of complex predicates in modern Farsi*. Ottawa: University of Ottawa. (Doctoral dissertation.)
- Van Valin, Robert D., Jr. 2005 *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2008. *RP's and the nature of lexical and syntactic categories in Role and Reference Grammar*. In Robert D. Van Valin, Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 161-178. Amsterdam/ Philadelphia: John Benjamins.
- Van Valin, Robert D. & Anja Latrouite. 2014. *Referentiality and Telicity in Tagalog and Lakota*. In Doris Gerland, Christian Horn, Anja Latrouite & Albert Ortman (eds.), *Meaning and grammar of nouns and verbs*, 410-426. Düsseldorf: Düsseldorf University Press.
- Verkuyl, Henk. 1972. *On the compositional nature of the aspects*. Dordrecht: Reidel.
- Yousef, Saeed. 2018. *Persian: A comprehensive grammar*. London New York: Routledge.



# LINKING SYNTAX AND SEMANTICS IN COMPARATIVES OF TAIWAN SIGN LANGUAGE: A ROLE AND REFERENCE GRAMMAR ACCOUNT

CHIEN-HUNG LIN<sup>1</sup> & JUNG-HSING CHANG<sup>2</sup>  
TUNG HAI UNIVERSITY<sup>1</sup>  
& NATIONAL CHUNG CHENG UNIVERSITY<sup>2</sup>

## Abstract

This paper discusses comparatives in Taiwan Sign Language (TSL) within the framework of Role and Reference Grammar, with the goal of finding out how TSL signers exploit the signing space to express comparatives. In TSL, no particular morpheme (e.g., locative morpheme or particle) is recruited to identify the roles that the RP arguments play in the comparatives (i.e., comparee and standard). Instead, TSL signers exploit the signing space to set the relation of arguments. This paper focuses on the three issues: (a) how the syntactic structures of TSL comparatives are represented in RRG, (b) how the structure of the signing space used to represent the TSL comparatives is incorporated into RRG, and (c) how the roles of RP arguments are determined in the logical structure of RRG.

## Keywords

Comparatives, referential loci, Taiwan Sign Language, signing space

## 1. Introduction

This paper discusses comparative constructions in Taiwan Sign Language (TSL) within the framework of Role and Reference Grammar (RRG) (Van Valin & LaPolla 1997; Van Valin 2005), with the goal of finding out how TSL signers exploit the signing space to express comparatives. The concept of comparison can be defined as a mental act by which two objects



are assigned a position on a predicative scale (Stassen 2005: 490). The linguistic codification for comparison is named ‘comparative constructions’ or simply ‘comparatives’. Comparative constructions in spoken languages have attracted much attention in typological studies (Beck et al. 2009; Bobaljik 2012; Dixon 2008; Stassen 1985; Ultan 1972) and in studies concerned with the interplay between syntax and semantics (Kennedy 1997; Klein 1982; Lin 2009; Liu 2011; Tsao 1989; Xiang 2006). However, comparative constructions in signed languages have not yet been explored, leaving a gap in current frameworks of comparatives. The study pursues the question of how the concept of comparison is encoded through the visual-gestural channel, and how the syntactic structures of TSL comparatives are represented in RRG.

In previous research on sign languages, it has been shown that signed languages and spoken languages share many linguistic properties and that both possess complex grammatical structures for encoding relationships between form and meaning (Chang et al. 2005; Chen 2012; Fischer 1975; Klima & Bellugi 1979; Liddell 1980; Meir & Sandler 2008; Sandler & Lillo-Martin 2006; Su 2011; Tai 2008; Tai & Su 2006). The major difference between signed languages and spoken languages resides in how their messages are produced and perceived, namely, the auditory-vocal channel of spoken languages and the visual-gestural channel of signed languages. The unique nature of the visual-gestural channel has allowed for the development of certain features distinctive to sign languages.

TSL is characterized by conveying information through the visual-gestural channel. TSL signers exploit the signing space, the three-dimensional space in front of them, to construct messages (Lin 2011, 2014). Plenty of linguistic phenomena make reference to the signing space (Engberg-Pedersen 1993; Friedman 1975; Liddell 2003; Meir & Sandler 2008; Padden 1990). This study aims to cast light on the interplay between TSL comparatives and the signing space within the framework of RRG, with the goal of finding out how TSL signers exploit the signing space to express comparatives. This paper is organized as follows: Section 2 introduces the background of comparatives. Section 3 explores the syntactic representation of TSL comparatives. Section 4 investigates the correlation between the semantic and syntactic representations of TSL comparatives, while Section 5 concludes this paper.

## 2. Background on comparative constructions

Typological studies on spoken language have shown that a typical comparative construction usually consists of five essential elements: (a) a comparee, (b) a standard, (c) a marker of the standard, (d) a parameter of

comparison, and (e) a degree marker (Dixon 2008; Dixon 2012; Heine 1997; Stassen 1985), as elucidated in (1).

- (1) Elements of a typical comparative construction in spoken languages
- a. Comparee: the item which is being compared.
  - b. Standard: the item which provides a reference point for the comparee to be compared against.
  - c. Marker of standard: the marker which indicates the grammatical relation of the standard.
  - d. Parameter of comparison: the predicate which indicates the attribute being compared.
  - e. Degree marker: the marker which designates the degree of the compared attribute.

With respect to grammatical relations, it has been generalized that the RP argument serving as the comparee is usually realized as the subject, while the RP argument serving as the standard may be realized as the object or an oblique argument (Dixon 2008; Lin 2014). This is demonstrated in the English comparative shown in example (2).

- (2) a. *Harley is smarter than Floyd.*
- b. *Harley is smarter than Floyd.*
- 
- Comparee    Parameter    Degree marker    Marker of standard    Standard

The comparative in (2) contains two RP arguments.<sup>1</sup> The RP argument occurring in the subject position (i.e., *Harley*) is identified as the comparee. The morpheme *than* is a specific particle in English comparatives that functions as the marker of the standard. The RP argument introduced by the comparative particle *than* (i.e., *Floyd*) is recognized as the standard. The predicative adjective *smart* serves as the parameter of the comparison, denoting the attribute compared. The adjective is inflected by the morpheme *-er* which serves as a degree marker to convey the comparative degree of the attribute.

Drawing on extensive data from a variety of spoken languages, many linguists delve into the issue concerning the types of comparative construction, with an intention to tease apart similarities and differences across spoken languages (Beck et al. 2009; Dixon 2008; Stassen 1985; Ultan 1972). According to the grammatical means recruited to identify the

<sup>1</sup> Abbreviations: NMR 'non-macrorole argument', NOM 'nominative', PRED 'predicate', RP 'reference phrase', SC 'spatial coreference'.

roles that the RP arguments play in a comparative construction (i.e., comparee and standard), five types of comparative constructions are proposed: (a) locational comparatives, (b) verbal comparatives, (c) particle comparatives, (d) bi-clausal comparatives, and (e) topic-comment comparatives (Dixon 2008; Heine 1997; Stassen 1985; Ultan 1972).

In brief, a typical comparative contains two compared RP arguments, in which one plays the role of the comparee and the other fills the role of the standard. In spoken languages, a diverse range of grammatical strategies are employed to identify the roles that the RP arguments play in a comparative. These strategies can be generally classified into two types: (a) grammatical morphemes, and (b) syntactic structures (Lin 2014). Grammatical morphemes include the use of locative morphemes, particular verbs and comparative particles, while syntactic structures include bi-clausal constructions and topic-comment constructions.

In TSL, no particular morpheme (e.g., locative morpheme or particle) is recruited to identify the roles that the RP arguments play in the comparatives in TSL. For instance, sentence (3a) compares IRON and PLASTIC in terms of their hardness. These two items being compared are encoded as RP arguments. Sentence (3a) shows that there is no grammatical morpheme recruited to determine the roles that the RP arguments play within this comparative. However, it is worth noting that each lexical sign in sentence (3a) is manipulated to be associated with a certain locus in the signing space. The omission of such manipulation will lead to a different interpretation, such as exemplified in (3b). Further, in sentence (3c), the RP argument IRON occurs before the RP argument PLASTIC. Interestingly, such exchange of the syntactic positions does not affect the interpretation of this comparative.

- |        |                                |                              |  |
|--------|--------------------------------|------------------------------|--|
| (3) a. | <i>PLASTIC</i> <sub>-L</sub>   | <i>IRON</i> <sub>-R</sub>    | <i>HARD</i> <sub>-R</sub> . <sup>2</sup> |
|        | plastic-NOM                    | iron-NOM                     | hard-PRED-SC                             |
|        | ‘Iron is harder than plastic.’ |                              |  |
| b.     | <i>PLASTIC</i> <sub>-N</sub>   | <i>IRON</i> <sub>-N</sub>    | <i>HARD</i> <sub>-N</sub> .              |
|        | plastic-NOM                    | iron-NOM                     | hard-PRED                                |
|        | ‘Iron and plastic are hard.’   |                              |  |
| c.     | <i>IRON</i> <sub>-R</sub>      | <i>PLASTIC</i> <sub>-L</sub> | <i>HARD</i> <sub>-R</sub> .              |
|        | iron-NOM                       | plastic-NOM                  | hard-PRED-SC                             |
|        | ‘Iron is harder than plastic.’ |                              |  |

---

<sup>2</sup> The capitalized notations of R and L indicate the loci of the referents. R stands for the right side of the space in front of the signer, and L stands for the left side. The notation N stands for the neutral space which refers to the area in center of the signing space.

Most of the studies in the previous literature on signed languages focus on issues of how comparative discourses are expressed and how signers make use of the signing space to construct a coherent discourse (Winston 1991; Winston 1993; Winston 1995). Specifically, it has been suggested that making use of two sides of the signing space, particularly the right side and the left side, is a convenient visual demonstration for expressing comparisons or contrasts in discourse (Engberg-Pedersen 1993; Johnston & Schembri 2007; Lin 2011; Meir & Sandler 2008; Winston 1993, 1995). However, less attention has been paid to the syntactic and the semantic levels, and not much research addresses how comparatives are constructed in signed languages. The study of comparative constructions in signed languages will help shed more light on the universality of human languages and the characteristics attributed to the effects of the visual-gestural channel. The following section introduces the syntactic representation of TSL comparatives.

### 3. TSL comparative constructions

Conceptually, a typical comparative construction contains two compared items, one functioning as the item being compared and the other as the reference point for the comparison. In terms of the function they possess in a comparative construction, the former item has the role of comparee, while the latter item has the role of standard. The two compared items are usually encoded as RP arguments in a comparative construction. To determine the roles that the RP arguments play within a comparative, spoken languages employ various grammatical strategies. As suggested by Stassen (1985, 2005), most of the strategies are concerned with ways of marking the RP argument that serves as the role of standard.

Drawing on TSL data collected, Lin (2014) proposed that TSL signers employ two fundamental comparative constructions: (a) comparative constructions with one verb and (b) comparative constructions with two verbs. All other TSL comparatives are derived from these two types. The syntactic structures and grammatical properties of the two types of constructions are introduced in Section 3.1 and Section 3.2.

#### 3.1 Comparative constructions with one verb

In TSL, when an adjectival verb is used predicatively to form a sentence, there is no copular verb occurring at the position between the adjectival verb and the subject. As in (4), the RP argument MATH occurs before the

adjectival verb EASY. Both signs are articulated in neutral space, as illustrated in (4a) and (4b).

- (4)  $MATH_{i-N}$                        $EASY_{-N}$   
 math-NOM                      easy-PRED  
 'Math is (very) easy.'

a.  $MATH_{i-N}$ b.  $EASY_{-N}$ 

When an adjectival verb is used predicatively to form comparatives, the syntactic structure becomes more complex because there is more than one RP argument involved in comparative. The sentence (5) is a comparative construction with one verb in which two items, namely, math and English, are compared in terms of the degree of difficulty. The two compared items are presented as RP arguments. The RP arguments are signed first and the adjectival verb is expressed later. Unlike the declarative sentence given in (4), the RP arguments are articulated at the right side and the left side of the space in front of the signer respectively, as shown in (5a) and (5b). The attribute being compared is denoted by the adjectival verb EASY, which is articulated at the right side rather than in neutral space, as shown in (5c).

- (5)  $ENGLISH_{i-L}$                        $MATH_{j-R}$                        $EASY_{-R}$   
 English-NOM                      math-NOM                      easy-PRED-SC  
 'Math is easier than English.'

a.  $ENGLISH_{i-L}$ b.  $MATH_{j-R}$ c.  $EASY_{-R}$

Sentence (5) shows that there is no grammatical morpheme recruited to determine the roles that the RP arguments play within this comparative. Further, it is worth noting that each lexical sign in sentence (5) is manipulated to be associated with a certain locus in the space in front of the signer. The omission of such manipulation will lead to a different interpretation, as illustrated in (6).

- (6) *ENGLISH*<sub>i-N</sub>      *MATH*<sub>j-N</sub>      *EASY*<sub>N</sub>  
 English-NOM      math-NOM      easy-PRED  
 ‘Both Math and English are easy.’

Sentence (6) is composed of the same lexical signs and obeys the same word order as sentence (5). However, unlike sentence (5), all the lexical signs in sentence (6) are articulated in neutral space. As a result, the sentence is interpreted as a declarative sentence rather than a comparative one. The contrast between sentences (5) and (6) shows that the spatial association is crucial in forming TSL comparatives. In TSL comparatives, the syntactic positions of RP arguments do not directly correlate with the determination of the roles of the RP arguments. Rather, Lin (2014) proposed that there are two crucial factors in determining the roles of RP arguments and the interpretation of comparatives: (a) the spatial locations of RP arguments and (b) the spatial location of adjectival verb, as illustrated in (7) to (8).

- (7) *ELDER SISTER*<sub>j-R</sub>    *YOUNGER SISTER*<sub>i-L</sub>    *PRETTY*<sub>-R</sub>  
 elder.sister-NOM    younger.sister-NOM    pretty-PRED-SC  
 ‘The elder sister is prettier than the younger sister.’



a. ELDER SISTER<sub>j-R</sub>



b. YOUNGER SISTER<sub>i-L</sub>



c. PRETTY<sub>-R</sub>

The sentence in (7) compares the elder sister with the younger sister in terms of their appearance. Both compared items are presented as RP arguments. The RP argument ELDER SISTER is signed by the right hand, while the RP argument YOUNGER SISTER is signed by the left hand, as shown in (7a) and (7b). While expressing the adjectival verb PRETTY, the

signer's body leans toward the right side, as shown in (7c). In terms of spatial location, the RP argument ELDER SISTER and the adjectival verb PRETTY are placed at the right side of the space, while the RP argument YOUNGER SISTER is at the left side. The sentence is interpreted as 'The elder sister is prettier than the younger sister.' If the syntactic positions of these lexical signs remain the same, the exchange of the spatial locations of the RP arguments will affect the interpretation of the given comparative, as exemplified in (8).

- (8) *ELDER SISTER*<sub>j-L</sub>    *YOUNGER SISTER*<sub>i-R</sub>    *PRETTY*<sub>-R</sub>  
 elder.sister-NOM    younger.sister-NOM    pretty-PRED-SC  
 'The younger sister is prettier than the elder sister.'

a. ELDER SISTER<sub>j-L</sub>b. YOUNGER SISTER<sub>i-R</sub>c. PRETTY<sub>-R</sub>

Sentence (8) is composed of the same lexical signs and obeys the same word order as sentence (7). The only difference between these two sentences lies in the spatial locations of RP arguments. In (8), the RP argument ELDER SISTER is placed at the left side of the signing space and the RP argument YOUNGER SISTER at the right side, as shown in (8a) and (8b). As a result, the sentence is interpreted as 'The younger sister is prettier than the elder sister.'

In TSL comparatives, the syntactic positions of RP arguments do not directly correlate with the determination of the roles of the RP arguments in a comparative (i.e., comparee or standard); rather, the spatial location of adjectival verb plays a vital role in determining the roles of the RP arguments, as illustrated in (9) and (10).

- (9) *MR. WANG*<sub>i-L</sub>    *MR. CHEN*<sub>j-R</sub>    *FAT*<sub>-L</sub>  
 Mr. Wang-NOM    Mr. Chen-NOM    fat-PRED-SC  
 'Mr. Wang is fatter than Mr. Chen.'

a. MR. WANG<sub>i-L</sub>b. MR. CHEN<sub>j-R</sub>c. FAT<sub>L</sub>

(10) MR WANG<sub>i-L</sub> MR. CHEN<sub>j-R</sub> FAT<sub>R</sub>  
 Mr. Wang-NOM Mr. Chen-NOM fat-PRED-SC  
 'Mr. Chen is fatter than Mr. Wang.'

a. MR. WANG<sub>i-L</sub>b. MR. CHEN<sub>j-R</sub>c. FAT<sub>R</sub>

The comparatives in (9) and (10) are composed of the same lexical signs and obey the same word order, but they denote antithetical meanings from each other. For instance, the RP argument MR. WANG is interpreted as the comparee in (9), whereas it is the standard in (10). There is no grammatical morpheme (e.g., locative morphemes or particles) being used to determine the relations of the RP arguments. The only difference between these two sentences resides in the spatial locations of adjectival verb FAT. The adjectival verb FAT in sentence (9) is articulated on the left side to be associated with the RP argument MR. WANG, while the lexical sign FAT in (10) is articulated on the right side to be associated with the RP argument MR. CHEN, as shown in (9c) and (10c), respectively. In brief, in comparative constructions with one verb, the RP argument associated with adjectival verbs is identified as the comparee, while the other argument is the standard.

Within an English comparative, the adjectival verb must be marked by the comparative morphemes such as *-er* or *more* to convey comparative degree of the adjective. These comparative morphemes are also termed 'degree markers' (Ullian 1972; Heine 1997) or 'index of comparison' (Dixon 2008). It has been pointed out that some spoken languages employ



degree markers obligatorily (i.e., English), others optionally (i.e., Mopan and Turkish), and still others do not make use of them at all (i.e., Chinese, Japanese, Swahili) (Heine 1997). According to our observation of TSL comparatives, TSL does not have comparative morphemes specific for the adjectival verbs in comparatives. Instead, reinforcing adverbs (e.g., VERY and EXCESSIVELY) are recruited optionally to intensify the degree of adjectival verbs. It is also found that sometimes the determination of the roles of RP arguments may hinge on scalar adverbs rather than adjectival verbs, as illustrated in (11) and (12).

- (11) *YOUNGER SISTER*<sub>i-L</sub> *ELDER SISTER*<sub>j-R</sub> *PRETTY*<sub>N</sub> *VERY*<sub>R</sub>  
 younger.sister-NOM elder.sister-NOM pretty-PRED very-ADV-SC  
 ‘The elder sister is prettier than the younger sister.’

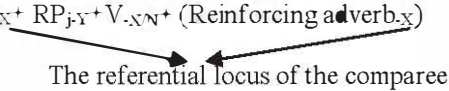
a. *YOUNGER SISTER*<sub>i-L</sub>b. *ELDER SISTER*<sub>j-R</sub>c. *PRETTY*<sub>↑</sub>d. *VERY*<sub>R</sub>

- (12) *YOUNGER SISTER*<sub>i-L</sub> *ELDER SISTER*<sub>j-R</sub> *PRETTY*<sub>N</sub> *VERY*<sub>L</sub>  
 younger.sister-NOM elder.sister-NOM pretty-PRED very-ADV-SC  
 ‘The younger sister is prettier than the elder sister.’

Sentence (11) and sentence (12) are composed of the same lexical signs and are constructed in the same word order. Both sentences recruit the reinforcing adverb VERY to intensify the degree of adjectival verb PRETTY. The only difference between these two sentences lies in the spatial locations of the reinforcing adverb VERY; they are interpreted differently. In sentence (11), the adjectival verb PRETTY is articulated in neutral space, without being associated with the right or the left, as in

(11c). The relations of the RP arguments remain indeterminate. This reinforcing adverb VERY is articulated at the right side to be associated with the RP argument ELDER SISTER, indicating that this argument serves as the comparee in the comparative, as shown in (11d). Conversely, if the adverb VERY is articulated at the left side, the RP argument ELDER SISTER will be identified as the role of standard, as exemplified in (12). That is to say, the reinforcing adverbs (i.e., VERY) have the function of setting the relations of the RP arguments in a comparative.

The above examples demonstrate that TSL comparatives manipulate either adjectival verbs or reinforcing adverbs to make association with the spatial locus of the comparee RP to distinguish between the comparative roles of RP arguments. Based on the above discussion, the syntactic structure of the comparatives with one verb can be generalized as (13) (Lin 2014).

- (13) Constructional schema for comparative constructions with one verb  
 $RP_i.X + RP_j.Y + V.N.N + (\text{Reinforcing adverb}_X)$   
  
 The referential locus of the comparee

In (13), the two compared RP arguments are expressed first and the adjectival verb is signed later. The RP arguments are assigned to the right and the left of the signing space respectively (i.e., X and Y). Thus, there are two referential loci established in the signing space. The adjectival verb (i.e., V) and the reinforcing adverb are modified to be associated with one of the referential loci. The RP argument associated with these two signs is identified as the comparee. The following section will discuss comparative constructions with two verbs.

### 3.2 Comparative constructions with two verbs

The second type of comparative construction contains two adjectival verbs. One of the adjectival verbs is related to the RP argument that is the comparee, while the other adjectival verb is associated with the RP argument that functions as the standard. These two adjectival verbs are formed by means of reduplicating the adjectival verb at issue; however, they have to be modified into different forms to convey different degrees of intensity, as illustrated in (14).

- (14) *MISS WANG*<sub>i-R</sub>                      *MISS CHEN*<sub>j-L</sub>  
 Miss Wang-NOM                      Miss Chen-NOM  
*LONG TONGUED*<sub>[lax]-L</sub>                      *LONG TONGUED*<sub>[tense]-R</sub><sup>3</sup>  
 long.tongued-PRED-SC                      long.tongued-PRED-SC  
 ‘Miss Wang is more long-tongued than Miss Chen.’

a. *MISS WANG*<sub>i-R</sub>b. *MISS CHEN*<sub>j-L</sub>c. *LONG TONGUED*<sub>[lax]-L</sub>d. *LONG TONGUED*<sub>[tense]-R</sub>

Sentence (14) compares Miss Wang and Miss Chen with respect to their talkativeness. These two compared items are presented as RP arguments in the comparative. The RP argument *MISS WANG* is articulated at the right side of the signing space, while the RP argument *MISS CHEN* is at the left side, as shown in (14a) and (14b). In this comparative construction, the adjectival verb *LONG TONGUED* is reduplicated, and then is expressed in different forms to convey different degrees of intensity. While expressing the adjectival verb undergoing lax modulation, the signer's body is oriented toward the left side of the signing space, namely, the referential locus of the RP argument *MISS CHEN*, as shown in (14c). While expressing the adjectival verb

<sup>3</sup> Following Klima and Bellugi (1979), [tense] refers to tense modulation which is characterized by an elongated movement, tensed hand muscles, and a long final hold, simultaneously accompanied by certain non-manual features, such as furrowed brows and squinted eyes, to emphasize the degree of adjectival verbs. [lax] refers to the lax modulation which is characterized by the hand muscles being lax and the movement of the given sign being shortened, to express approximate degree of adjectival verbs.

undergoing tense modulation, the signer's body is oriented toward the right side, namely, the referential locus of the RP argument MISS WANG, as shown in (14d).

Note that it is obligatory to present the reduplicated adjectival verbs in the forms that convey different degrees of intensity to show a contrast between the compared items; otherwise the comparative will become ungrammatical, as shown in (15a). Further, the spatial association of each constituent in (14) is crucial grammatical strategy. The omission of such manipulation will lead to the ungrammaticality of the comparative, as exemplified in (15b).

- (15) a. \*MISS WANG<sub>F-N</sub>      MISS CHEN<sub>J-N</sub>  
 Miss Wang-NOM      Miss Chen-NOM  
 LONG TONGUED<sub>L</sub>      LONG TONGUED<sub>R</sub>  
 longtongued-PRED-SC      longtongued-PRED-SC
- b. \*MISS WANG<sub>F-N</sub>      MISS CHEN<sub>J-N</sub>  
 Miss Wang-NOM      Miss Chen-NOM  
 LONG TONGUED<sub>[lex]-N</sub> LONG TONGUED<sub>[tense]-N</sub>  
 longtongued-PRED      longtongued-PRED

Sentence (15a) and sentence (15b) are composed of the same lexical signs as those in sentence (14). In sentence (15a), the reduplicated adjectival verbs do not undergo aspectual modulations; thus, the comparative becomes ungrammatical. In sentence (15b), all the lexical signs are articulated in neutral space; as a result, the sentence becomes ungrammatical. With respect to word order, this comparative construction possesses flexible word order, as illustrated in (16).

- (16) a. MR. LIN<sub>F-L</sub>      MR. CHEN<sub>J-R</sub>      FAT<sub>[tense]-L</sub>      FAT<sub>[lex]-R</sub>  
 Mr. Lin-NOM      Mr. Chen-NOM      fat-PRED-SC      fat-PRED-SC  
 'Mr. Lin is fatter than Mr. Chen.'
- b. MR. CHEN<sub>J-R</sub>      MR. LIN<sub>F-L</sub>      FAT<sub>[tense]-L</sub>      FAT<sub>[lex]-R</sub>  
 Mr. Chen-NOM      Mr. Lin-NOM      fat-PRED-SC      fat-PRED-SC  
 'Mr. Lin is fatter than Mr. Chen.'
- c. MR. LIN<sub>F-R</sub>      MR. CHEN<sub>J-L</sub>      FAT<sub>[tense]-L</sub>      FAT<sub>[lex]-R</sub>  
 Mr. Lin-NOM      Mr. Chen-NOM      fat-PRED-SC      fat-PRED-SC  
 'Mr. Chen is fatter than Mr. Lin.'

In (16b), all the lexical signs are placed at the same spatial locations as those in (16a). The only difference between sentences (16a) and (16b) lies in the syntactic positions of RP arguments. Unlike (16a), in (16b), the RP

argument MR. CHEN occurs before the RP argument MR. LIN, but the comparative still retains the same interpretation. However, if the spatial locations of the RP arguments are exchanged, the comparative will be interpreted in a different way, as shown in (16c). In (16c), all the lexical signs are arranged in the same word order as that of sentence (16a). The major difference between sentences (16a) and (16c) resides in the spatial locations of RP arguments. In (16c), the RP argument MR. LIN is placed at the right side, whereas the RP argument MR. CHEN is at the left side. As a result, the sentence is interpreted differently, namely, ‘Mr. Chen is fatter than Mr. Lin.’

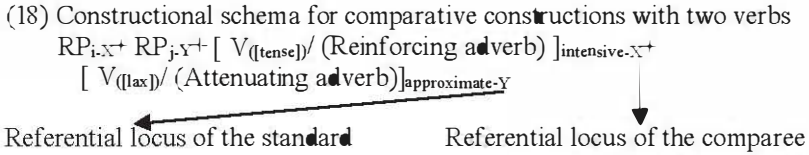
The above examples show that the adjectival verb denoting the attribute compared is reduplicated, and they undergo different aspectual modulations to convey varying degrees of intensity. As pointed out by Klima and Bellugi (1979), not all adjectival verbs can undergo aspectual modulations. In TSL, some adjectival verbs are incompatible with aspectual modulations due to their articulatory limitations (e.g., the adjectival verbs SMART, PRETTY, RICH, and POOR in TSL). If a comparative construction contains the adjectival verb that is incompatible with aspectual modulations, two types of scalar adverbs, namely, reinforcing adverbs (i.e., VERY and EXCESSIVELY) and attenuating adverbs (i.e., A LITTLE, COMMON, MEDIocre, and ORDINARY), will be recruited to show different degrees of intensity, as exemplified in (17).

- (17) *CAT*<sub>i-R</sub>    *DOG*<sub>j-L</sub>    *SMART*<sub>R</sub>    *VERY*<sub>R</sub>  
 cat-NOM    dog-NOM    smart-PRED-SC    very-ADV-SC  
*SMART*<sub>L</sub>    *MEDIocre*<sub>L</sub>  
 smart-PRED-SC    mediocre-ADV-SC  
 ‘Cats are smarter than dogs.’

Sentence (17) compares two animals, namely, cats and dogs, in terms of their intelligence. The compared items are presented as RP arguments. The RP argument CAT is assigned to the right side of the signing space, while the RP argument DOG is assigned to the left side. The attribute being compared is denoted by the adjectival verb SMART. This adjectival verb is expressed twice: the first time on the right side, and the second time on the left side. In this sentence, two types of scalar adverbs are recruited: reinforcing adverbs and attenuating adverbs. In terms of syntactic distribution, the reinforcing adverb VERY occurs in the syntactic position immediately following the first adjectival verb, and the attenuating adverb MEDIocre occurs in the position immediately following the second adjectival verb. In terms of spatial location, the

adverb VERY is articulated at the right side, while the adverb MEDIOCRE is at the left side.

Based on the comparative sentences discussed above, the syntactic structure of comparative constructions with two verbs can be generalized as (18) (Lin 2014).



With respect to comparative constructions with two verbs, the RP arguments are assigned on the right and the left sides of the signing space respectively (i.e., X and Y). After expressing the two compared RP arguments, there are two referential loci established in the signing space. The constituents of the predicate denoting intensive degree are modified to be associated with one of the referential loci, whereas those of the predicate denoting approximate degree are modified to be associated with the other referential locus. The RP argument associated with the former predicate (i.e., RP<sub>i</sub>) is identified as the comparee, whereas the RP argument associated with the latter predicate (i.e., RP<sub>j</sub>) is identified as the standard.

### 3.3 Comparatives derived from the two fundamental constructions

The above discussions have pointed out that in TSL there are two basic comparative constructions, namely, comparatives with one verb and comparatives with two verbs, and that they have a close relationship with the signing space. These two constructions are the fundamental constructions from which all TSL comparatives are derived. It is found that certain lexical signs, such as SURPASS, DEFEAT, and OUTRANGE, are frequently employed in TSL comparatives. These lexical signs behave like verbs and denote the meaning of ‘to exceed’ or ‘to surpass’. Thus, these verbs are called ‘exceed’ verbs (Lin 2014). See the comparative containing the lexical sign DEFEAT exemplified in (19).

- (19) *MR. CHEN*<sub>i-L</sub>    *MR. LIN*<sub>j-R</sub>    *HANDSOME*<sub>N</sub>    *DEFEAT*<sub>j→i</sub>.  
 Mr. Chen-NOM    Mr. Lin-NOM    handsome-PRED    defeat-PRED  
 ‘Mr. Lin is more handsome than Mr. Chen.’

a. *MR. CHEN*<sub>i-L</sub>b. *MR. LIN*<sub>j-R</sub>c. *HANDSOME*<sub>N</sub>d. *DEFEAT*<sub>j→i</sub>

- (20) *MR. CHEN*<sub>i-L</sub>    *MR. LIN*<sub>j-R</sub>    *HANDSOME*<sub>N</sub>    *DEFEAT*<sub>i→j</sub>.  
 Mr. Chen-NOM    Mr. Lin-NOM    handsome-PRED    defeat-PRED  
 ‘Mr. Chen is more handsome than Mr. Lin.’

Sentence (19) compares Mr. Lin with Mr. Chen in terms of their looks. The RP argument *MR. CHEN* is assigned to the left side of the signing space, while the RP argument *MR. LIN* is assigned to the right side. The attribute being compared is denoted by the adjectival verb *HANDSOME*. Note that this adjectival verb is articulated in neutral space rather than being associated with a specific referential locus, as shown in (19c). Thus, the relations of the RP arguments remain indeterminate. In such situations, reinforcing adverbs are supposed to be recruited to set the relations; however, no adverb appears in sentence (19). Instead, the sign *DEFEAT* is employed, and it is placed at the sentence-final position. This sign contains a path movement which moves from the locus associated with the RP argument *MR. LIN* to that of the RP argument *MR. CHEN*, as shown in (19d). Through the path movement, the roles that the RP arguments play in this comparative can be identified. If the path movement of the sign *DEFEAT* is moved in the opposite direction, the comparative will be interpreted differently, as shown in (20).

Verbs such as SURPASS, DEFEAT, and OUTRANGE in TSL are subsumed as agreement verbs because these verbs contain a path movement which allows them to indicate the grammatical relations of the RP arguments in a sentence (i.e., subject and object). When these verbs are employed in comparatives, they maintain their ability to indicate the relations of RP arguments remains. In summary, the syntactic structures of two types of comparative constructions with ‘exceed’ verbs are generalized in (21) and (22) (Lin 2014).

(21) Comparative constructions with one verb employing ‘exceed’ verbs

$RP_{i-x} + RP_{j-y} + V_{-x} + (\text{reinforcing adverb}_{-x}) + \text{‘Exceed’ } V_{i \rightarrow j}$

(22) Comparative constructions with two verbs employing ‘exceed’ verbs

$RP_{i-x} + RP_{j-y} + [V_{(\text{tense})}/(\text{reinforcing adverb})]_{\text{intensive-x}^+} + [V_{(\text{tense})}/(\text{attenuating adverb})]_{\text{approximate-y}^+} + \text{‘Exceed’ } V_{i \rightarrow j}$

In these two constructions, the ‘exceed’ verb (i.e., ‘exceed’ V) can occur in syntactic position either following or preceding the verbal predicate(s). The ‘exceed’ verb contains a path movement. The RP argument associated with the initial point of the path movement is identified as the comparee (e.g.,  $RP_i$ ), while the RP argument associated with the endpoint is identified as the standard (e.g.,  $RP_j$ ).

#### 4. The linking of arguments in comparatives to syntax

The comparatives in TSL belong to state verbs. The two basic logical structures of comparatives in TSL are proposed as in (23) and (24). In (23), MORE THAN refers to comparison and predicate contains two RP arguments. The arguments with subscripts such as x, y, and z indicate whether the arguments are co-indexed. In (24), exceed verbs are employed in comparatives, they can indicate the relations of RP arguments.

(23) [ $(\text{pred}_1(x))$  MORE THAN  $(\text{pred}_2(y))$ ]

(24) [ $(\text{pred}_1(x)) \wedge (\text{pred}_2(y))$ ]  $\wedge$  exceeds<sup>3</sup>(x, y)

Based on RRG, the three steps for linking semantics to syntax are: (a) to constitute the logical structure and then replace the variables in it with referring expressions; (b) to determine which argument is actor and which is undergoer, based on the Actor-Undergoer Hierarchy, and (c) to link the arguments into syntax. From the observation of the surface forms, the



constructional schemata for TSL comparatives are given in (25) and (26).

(25) Comparative constructions with one verb

$PSA_{i-X} + \bullet BJ_{j-Y} + V_{1-N/N^+} (V_{2-X \rightarrow Y})$

(26) Comparative constructions with two verbs

$PSA_{i-X} + \bullet BJ_{j-Y} + V_{1-X} + V_{2-Y} + (V_{3-X \rightarrow Y})$

As for comparative constructions with one verb,  $V_1$  stands for the adjectival verbs denoting the attribute being compared.  $V_2$  refers to ‘exceed’ verbs that move from the locus of comparee RP to that of standard RP. Note that due to flexible word order of TSL, PSA position is not limited to the sentential initial position. As for comparative constructions with two verbs,  $V_1$  stands for the adjectival verbs with intensive modulation and  $V_2$  stands for the adjectival verbs with approximate modulation.  $V_3$  refers to ‘exceed’ verbs that clarify the grammatical relations of RP arguments. Further, the linking principles, as proposed in (27), account for how the RP arguments in TSL comparatives are linked to syntactic representation. Notably, in addition to identifying the macroroles of RP arguments play in a comparative, the arguments have to be assigned to the right side or the left side, respectively, to establish referential loci to which subsequent constituents may refer, as illustrated in Principle C.

(27) The linking Principles for TSL Comparatives:

Principle A: The Undergoer of  $V_1$ , namely Undergoer<sub>1</sub>, is linked to the PSA position.

Principle B: The role, namely non-macrorole direct core argument (NMR), is linked to the position immediately preceding or following The Undergoer<sub>1</sub>.

Principle C: Undergoer<sub>1</sub> is to be associated with a designated referential locus X, the right side or the left side of the signing space. NMR is to be associated with the locus that is contralateral to locus X.

TSL comparatives given in (28) to (30) demonstrate how the semantic representation is linked to the syntactic representation.

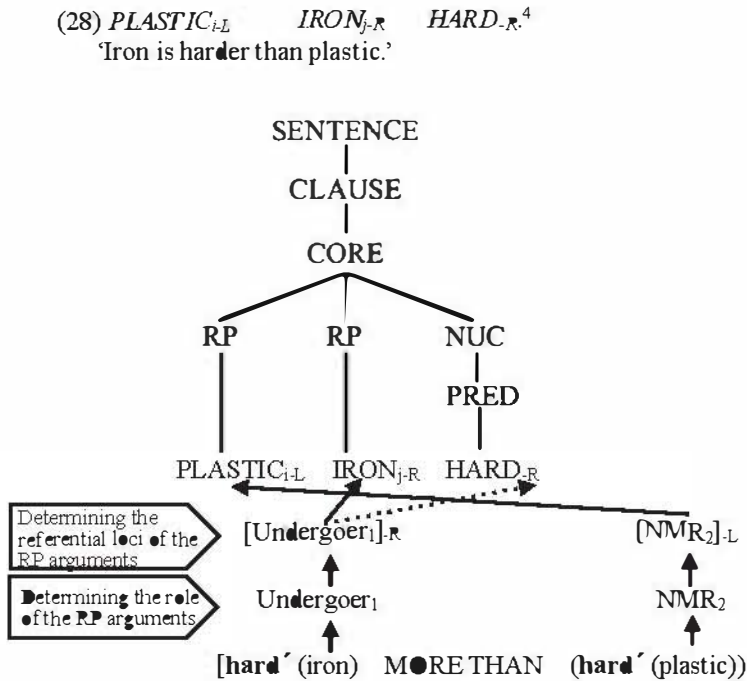


Figure 1: Semantics-to-syntax linking in (28)

The comparative with one verb is exemplified in (28). Following the linking principle A and principle B, the first argument of  $V_1$  (i.e., *IRON*) is selected as  $\text{Undergoer}_1$ , and the first argument of  $V_2$  (i.e., *PLASTIC*) is selected as  $\text{NMR}_2$ . Following the linking principle C,  $\text{Undergoer}_1$  is assigned to the right side of the signing space and is linked to PSA position. As for  $\text{NMR}_2$ , it is assigned to the left side of the signing space and is linked to the position immediately next to PSA position. Further, the adjectival verb *HARD* is manipulated to be associated with the referential locus of  $\text{Undergoer}_1$ , as the dotted arrow illustrated in (28). Example (29) demonstrates how the linking algorithm accounts for the comparatives with two verbs.

<sup>4</sup> The macroroles subscripted with 1 or 2 indicate whether the given macrorole is denoted by  $V_1$  or  $V_2$ .

- (29) *MR. LIN*<sub>i-L</sub> *MR. CHEN*<sub>j-R</sub> *FAT*<sub>[tense]-L</sub> *FAT*<sub>[lax]-R</sub>  
'Mr. Lin is fatter than Mr. Chen.'

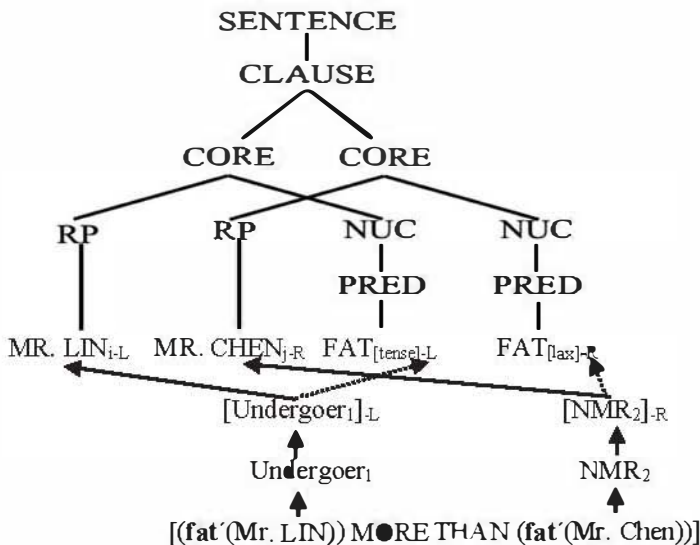


Figure 2: Semantics-to-syntax linking in (29)

The comparative with two verbs is exemplified in (29).  $RP_1$  agrees with  $FAT_{[tense]}$  and  $RP_2$  with  $FAT_{[lax]}$ , which suggests that  $RP_1$  is a core argument of  $FAT_{[tense]}$  and  $RP_2$  a core argument of  $FAT_{[lax]}$ , which in turn suggests that each predicate heads its own core. Thus, the syntactic representation of this sentence involves two cores in a clause, namely a core juncture. Following the linking principle A and principle B, the first argument of  $V_1$  (i.e., MR. LIN) is selected as  $Undergoer_1$ , and the first argument of  $V_2$  (i.e., MR. CHEN) is selected as  $NMR_2$ . Following the linking principle C,  $Undergoer_1$  is assigned to the right side of the signing space and is linked to PSA position.  $NMR_2$  is assigned to the contralateral side to the locus of  $Undergoer_1$  and is linked to the position immediately next to PSA position. Further, the two adjectival verbs, FAT, are expressed. The verb FAT with tense modulation is manipulated to be associated with the locus of  $Undergoer_1$ , and the one with lax modulation is manipulated to be associated with the locus of  $NMR_2$ , as the dotted arrow shown in (29). Example (30) demonstrates how linking algorithm accounts for the comparatives with two verbs employing 'exceed' verbs.

(30) MR CHEN<sub>i-L</sub> MR LIN<sub>j-R</sub> HANDSOME<sub>N</sub> DEFEAT<sub>j→i</sub>.  
 'Mr. Lin is more handsome than Mr. Chen.'

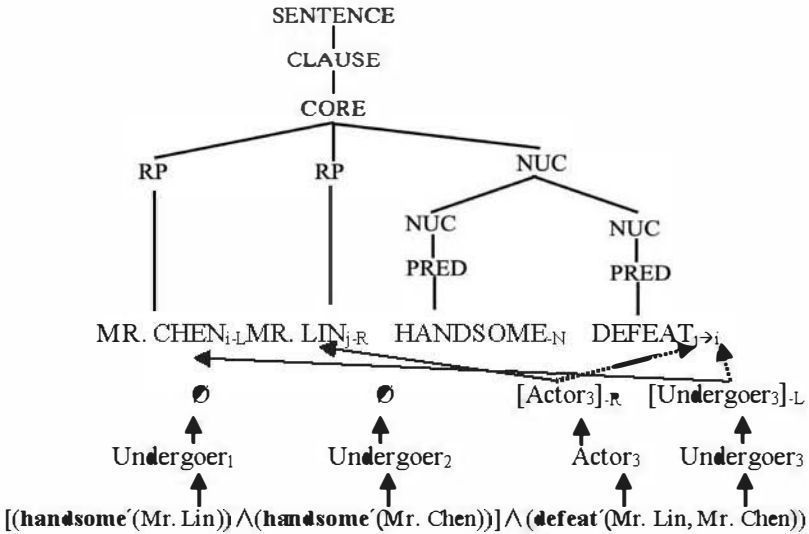


Figure 3: Semantics-to-syntax linking in (30)

The comparative with one verb employing the ‘exceed’ verb DEFEAT is exemplified in (30). Following the linking principle A and Principle B, the first argument of V<sub>1</sub> (i.e., MR. LIN) is selected as Undergoer<sub>1</sub>, the first argument of V<sub>2</sub> (i.e., MR. CHEN) is selected as Undergoer<sub>2</sub>. As for the verb DEFEAT, namely V<sub>3</sub>, the first argument of V<sub>3</sub> (i.e., MR. LIN) is selected as Actor<sub>3</sub>, the second argument of V<sub>3</sub> (i.e., MR. CHEN) is selected as Undergoer<sub>3</sub>. As suggested by privileged syntactic argument selection hierarchy, when the two arguments refer to the same entity, only the RP argument with the macrorole in a higher prominence hierarchy is syntactically expressed. Thus, Actor<sub>3</sub> is assigned to the right side of the signing space and is linked to PSA position. The rightmost argument in the logical structure, namely, Undergoer<sub>3</sub>, is expressed and is linked to the position immediately next to PSA position; additionally, it is assigned to the locus that is opposite to the locus of Actor<sub>3</sub>. Then, the adjectival verb HANDSOME is expressed in the neutral space. Further, the ‘exceed’ verb DEFEAT accompanies a path movement which moves from the locus of Actor<sub>3</sub> to that of Undergoer<sub>3</sub>, indicating that Mr. Lin defeats Mr. Chen with respect to their appearance.

The above examples show that the use of signing space for spatial reference is recognized as an essential spatial device for TSL comparatives. After the semantic macrorole of each argument is determined, these arguments must be associated with the right side or the left side of the signing space for the following verbal predicates to make spatial association with them.

## 5. Concluding remarks

This study has demonstrated that like many spoken languages, TSL possesses particular constructions to express the concept of comparison, and these constructions are composed of the essential elements, including a comparee, a standard, a parameter of comparison, and a degree marker. This paper showed that TSL signers employ two fundamental comparative constructions: (a) comparative constructions with one verb and (b) comparative constructions with two verbs. All other TSL comparatives are derived from these two types.

TSL does not contain grammatical morphemes that consistently serve as markers of the standard. Instead, TSL signers exploit signing space for that function. Ascribing to visual-gestural channel of TSL, the structure of signing space plays a significant role in TSL comparative constructions. Thus, adopting the signing space as a parameter in the framework of RRG allows us to provide a comprehensive account of TSL comparatives. While expressing comparatives, TSL signers divide the signing space into two parts: the right side and the left side. It is required to assign the RP arguments of a comparative to these two sides respectively, thus establishing two referential loci in the signing space. The function of these referential loci is to facilitate subsequent constituents referring to their intended arguments in a comparative. The findings of this study reveal interesting facts about comparative constructions in TSL, and provide insightful typological implications.

## References

- Beck, Sigrid, Krasikova, Sveta, Fleischer, Daniel, Gergel, Remus, Hofstetter, Stefan, Savelsberg, Christiane, Vanderelst, John & Villalta, Elisabeth. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook*, 9(1): 1-66.
- Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge: MIT Press.

- Chang, Jung-hsing, Su, Shiou-fen & Tai, James H-Y. 2005. Classifier predicates reanalyzed, with special reference to Taiwan Sign Language. *Language and Linguistics* 6:247-278.
- Chen, Yijun. 2012. Questioning in Taiwan Sign Language conversation. Chiayi, Taiwan: National Chung Cheng University. (Doctoral dissertation.)
- Dixon, R. M. W. 2008. Comparative constructions: A cross-linguistic typology. *Studies in Language* 32:787-817.
- Dixon, R. M. W. 2012. *Basic linguistic theory: Further grammatical topics*. New York: Oxford University Press.
- Engberg-Pedersen, Elisabeth. 1993. *Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg: Signum Verlag.
- Fischer, Susan D. 1975. Influences on word-order change in American Sign Language. In Li, Charles N. (ed.), *Word order and word order change*, 1-25. Austin: University of Texas press.
- Friedman, Lynn A. 1975. Space, time, and person reference in American Sign Language. *Language* 51(4): 940-961.
- Heine, Bernd. 1997. *Cognitive foundations of grammar*. New York: Oxford University Press.
- Johnston, Trevor & Schembri, Adam. 2007. *Australian Sign Language (Auslan): An introduction to sign language linguistics*. Cambridge: Cambridge University Press.
- Kennedy, C. 1997. *Projecting the adjective: the syntax and semantics of gradability and comparison*. Santa Cruz: University of California. (Doctoral dissertation.)
- Klein, Ewan. 1982. The interpretation of adjectival comparatives. *Journal of Linguistics* 18:113-136.
- Klima, Edward S. & Bellugi, Ursula. 1979. *The signs of language*. Harvard: Harvard University Press.
- Liddell, Scott K. 1980. *American sign language syntax*. The Hague: Mouton
- Liddell, Scott K. 2003. *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Lin, Chien-hung. 2011. Comparative constructions in Taiwan Sign Language. In Chang, Jung-hsing (ed.), *Language and cognition: Festschrift in honor of James H-Y. Tai on his 70th Birthday*, 235-58. Chia-Yi, Taiwan: The Taiwan Institute for the Humanities, National Chung Cheng University.
- Lin, Chien-hung. 2014. The syntax and semantics of comparatives in Taiwan Sign Language. Chiayi, Taiwan: National Chung Cheng

- University. (Doctoral dissertation.)
- Lin, Jo-wang. 2009. Chinese comparatives and their implicational parameters. *Natural Language Semantics* 17:1-27.
- Liu, Chen-Sheng Luther. 2011. The Chinese *bi* comparative. *Lingua* 121:1767-1795.
- Meir, Irit & Sandler, Wendy. 2008. *A language in space: The story of Israeli Sign Language*. New York: Lawrence Erlbaum Associates.
- Padden, Carol. 1990. The relation between space and grammar in ASL verb morphology. In Lucas, Ceil (ed.), *Sign language research: Theoretical issues*, 118-132. Washington, DC.: Gallaudet university press.
- Sandler, Wendy & Lillo-Martin, Diane. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press.
- Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford: Basil Blackwell.
- Stassen, Leon. 2005. Comparative constructions. In Haspelmath, Martin, Dryer, Matthew S., Gil, David, & Comrie, Bernard (eds.), *The world atlas of language structures*, 490-493. Oxford: Oxford University Press.
- Su, Shiou-fen. 2011. *Modality effects in grammar and discourse in Taiwan Sign Language*. Chiayi, Taiwan: National Chung Cheng University. (Doctoral dissertation.)
- Tai, James H-Y. 2008. The nature of Chinese grammar: Perspectives from sign language. (Paper presented to the Proceedings of The 20th North American Conference on Chinese Linguistics, Ohio State University, 25-27 April 2008)
- Tai, James H-Y. & Shiou-fen Su. 2006. Agreement in Taiwan Sign Language. In Chang, Henry Y., Huang, Lilian M. & Ho, Dah-an (eds.), *Streams converging into an ocean: Festschrift in honor of professor Paul Jen-kuei Li on his 70th Birthday*, 341-363. Taipei, Taiwan: Academic sinica.
- Tsao, Feng-fu. 1989. Comparison in Chinese: A topic-comment approach. *Tsing Hua Journal of Chinese Studies* 19:151-189.
- Ulan, Russell. 1972. Some features of basic comparative constructions. *Working Papers on Language Universals* 9:117-162.
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
- Winston, Elizabeth A. 1991. Spatial referencing and cohesion in an American Sign Language text. *Sign Language Studies* 73:397-410.

- Winston, Elizabeth A. 1993. *Spatial mapping in comparative discourse frames in an American Sign Language lecture*. Washington, D.C.: Georgetown University. (Doctoral dissertation.)
- Winston, Elizabeth A. 1995. Spatial mapping in comparative discourse frames. In Emmorey, Karen & Reilly, Judy S. (eds), *Language, gesture, and space*, 87-114. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Xiang, Ming. 2006. *Some topics in comparative constructions*. Michigan: Michigan State University. (Doctoral dissertation.)





**PART 2.**

**ARGUMENT MARKING**



# DATIVE CASE AND THREE-PLACE PREDICATES IN JAPANESE<sup>1</sup>

HIDEKI KISHIMOTO  
KOBE UNIVERSITY

## Abstract

The present paper argues that the distribution of *ni* marking assigned to the indirect internal arguments of Japanese three-place verbs can be optimally characterized on the basis of their inherent lexical meanings. In Japanese, dative case *ni* is assigned to an indirect internal argument by the default case assignment rule, so that the indirect internal arguments of change-of-possession verbs are marked with dative case regardless of whether they count as ‘source’ or ‘goal’. With change-of-location verbs, on the other hand, indirect internal arguments counting as ‘locative-goal’ are obliquely marked. It is argued that the indirect internal arguments of three-place predicates are marked with *kara* ‘from’ if they are identified as ‘locative-source’. It is shown that the dative-marked source arguments of change-of-possession verbs are constrained by the animacy constraint, but that when they are assigned *kara* marking, they are exempt from the animacy constraint.

## Keywords

Dative case, three-place predicate, change of location, change of possession, ablative

## 1. Introduction

Cross-linguistically, it is fairly common to mark goal arguments with da-

---

<sup>1</sup> This is a revised version of the paper presented at Role and Reference Grammar 2015 (held on July 31-August 2, 2015 at the Heinrich Heine University Düsseldorf, Germany), which is a significantly shortened version of the ms. entitled “The Janus-faced nature of dative marking in Japanese”. I am grateful to Robert Van Valin Jr., Sebastian Löbner, Kiyoko Toratani, Wataru Nakamura, Toshio Otori, and Mitsuaki Shimojo for their various comments, suggestions and encouragements.

tive case, but at the same time, it is sometimes observed that the same dative case marking can appear on either the source or goal argument of three-place predicates. The same morphological behavior is observed in three-place predicate constructions in Japanese; a three-place predicate of change of possession like *ageru* 'give' allows its goal argument to be marked with dative case *ni*, and *morau* 'get' allows its source argument to receive dative *ni* marking, even though these arguments are taken to be the participants of events representing different orientations of transfer.<sup>2</sup> The ambivalent behavior of *ni* marking appearing on indirect internal arguments is not always observed, however, since the pairs of predicates denoting a change of location, such as *okuru* 'send' and *uketoru* 'receive', display different distributions with regard to the marking of their indirect internal arguments, i.e. the verb *okuru* permits *ni* marking for its goal argument, but the source argument of *uketoru* cannot be marked with *ni*. This fact raises the question of why three-place predicates do not behave in a unitary manner with regard to the case marking of their indirect internal arguments.

The present paper provides an RRG account for how the dative *ni* marking is assigned in Japanese. I suggest that the prima facie puzzling behavior of dative *ni* marking receives a principled account if dative case is defined as the default case, as proposed by Van Valin (1991). It is first argued that in Japanese, dative case marking is allocated to the non-macrorole core argument of a change-of-possession verb (counting as a possessor), and that the assignment of dative case is made without reference to the negative operator NOT, which defines a thematic difference between 'source' and 'goal'. Then, I show that the dual behavior of *ni* marking is not observed for change-of-location verbs because their indirect internal arguments need to be obliquely marked. I argue that the indirect internal argument of a change-of-location verb receives locative *ni* marking when identified as 'goal', whereas it is marked with ablative *kara* 'from' when identified as 'source'. I also suggest that change-of-possession verbs invoke a change in the LS when their source argument is marked with ablative *kara* 'from'.

The discussion in this paper proceeds as follows. In section 2, I will review some notable properties of three-place predicates in Japanese. It is shown that they are divided into two classes according to whether they denote a change of possession or a change of location. Section 3 provides an RRG account for how case marking is assigned to the indirect internal

---

<sup>2</sup> The terms 'direct internal argument' and 'indirect internal argument' are used in informal, pre-theoretical descriptions for the purpose of identifying the types of arguments selected by three-place predicates (Williams 1981; Levin and Rappaport Hovav 1995).

arguments of change-of-possession and change-of-location verbs. Section 4 shows how ablative *kara* marking is assigned to the source arguments of three-place predicates. Section 5 is the conclusion.

## 2. Three-place predicates in Japanese

As often discussed (see e.g. Blake 1994; Blansit 1988; Maling 2001; Haspelmath 2003; Malchukov et al. 2010; Van Valin 2018), dative arguments can represent recipient, experiencer, or beneficiary, all of which fall under the general cover term of ‘goal’. Even though marking a goal argument with dative case is a common cross-linguistic option, it is sometimes observed that the same dative case marking can be assigned to the source and goal arguments of three-place predicates (see e.g. Janda 1993; Van Velle & Van Langendonck 1996).

The same type of morphological behavior is found in Japanese. In point of fact, in descriptive studies of Japanese grammar, it is often mentioned (e.g. Martin 1975; Shimizu 1977; Sugimoto 1988; Matsumoto 2000) that certain three-place verbs, which include the pairs of verbs such as *ataeru* ‘give’/*morau* ‘get, receive’, *ageru* ‘give’/*morau* ‘get, receive’, *sazukeru* ‘furnish’/*sazukaru* ‘receive’, *osieru* ‘teach’/*osowaru* ‘be taught’, *osieru* ‘teach’/*nararu* ‘learn’, etc., allow their source or goal arguments to bear dative *ni* marking, as exemplified in (1).

- (1) a. *Eri-ga tomodati-ni manga-o atae-ta.*  
 Eri-NOM friend-DAT comic-ACC give-PST  
 ‘Eri gave her friend the comics.’  
 b. *Tomodati-ga Eri{-ni/-kara} manga-o morat-ta.*  
 friend-NOM Eri{-DAT/-ABL} comic-ACC get-PST  
 ‘Her friends got the comics from Eri.’

In (1a), the verb *ataeru* ‘give’ expresses transfer of possession from the subject to the *ni*-marked indirect internal argument, and the *ni*-marked argument *tomodati* ‘friend’ is construed as a goal. In (1b), the verb *morau* ‘get’ describes reverse transfer of possession, and the *ni*-marked indirect internal argument counts as a source. The two indirect internal arguments in (1) receive the identical dative *ni* marking even though they are identified as the participants of events representing different orientations of transfer. Note that with a predicate like *morau* ‘get’, the source argument can alternatively be marked with ablative *kara* ‘from’.

The dual behavior of *ni* marking is not always observed, however, because the source arguments of some three-place predicates cannot receive

*ni* marking, as seen in the pair of the verbs *okuru* ‘send’ and *uketoru* ‘receive’ in (2).

- (2) a. *Eri-ga sensei-ni tegami-o okut-ta.*  
 Eri-NOM teacher-LOC letter-ACC send-PST  
 ‘Eri sent the letter to the teacher.’  
 b. *Sensei-ga Eri{\**-ni/-kara*} tegami-o uketot-ta.*  
 teacher-NOM Eri{-DAT/-ABL} letter-ACC receive-PST  
 ‘The teacher received the letter from Eri.’

The change-of-location verb *okuru* ‘send’ permits *ni* marking for its goal argument, whereas the source argument of the change-of-location verb *uketoru* ‘receive’ cannot be marked with *ni* (and can only be marked with ablative *kara*).

The three-place predicate constructions in (1a) and (2a) have identical surface case frames, but they represent different meanings. The verb *ataeru* ‘give’ in (1a) describes abstract transfer of ownership, so the act of ‘giving the comics to the friend’ can be successful even if no spatial change takes place. By contrast, the verb *okuru* ‘send’ in (2a) describes a physical change of location with no change of possession implied. The action described by *okuru*, unlike *ataeru*, involves spatial movement of an entity. The two classes of three-place predicates, i.e. change-of-possession and change-of-location verbs, convey distinct meanings, and thus can be distinguished in semantic terms.

The status of *ni* marking on the goal arguments differs between (1a) and (2a), as seen from the fact that direct passivization is possible with the *ni*-marked goal argument of *ataeru* ‘give’ but not the goal argument of *okuru* ‘send’.

- (3) a. *Tomodati-ga kinoo (Eri-ni) manga-o atae-rare-ta.*  
 friend-NOM yesterday (Eri-by) comic-ACC give-PASS-PST  
 ‘The friend was given the comics (by Eri) yesterday.’  
 b. *#Sensei-ga kinoo (Eri-ni) tegami-o okur-are-ta.*  
 teacher-NOM yesterday (Eri-by) letter-ACC send-PASS-PST  
 ‘The teacher was adversely affected by (Eri’s) sending the letter yesterday.’

The passive sentence in (3a) formed on *ataeru* ‘give’ can have a neutral direct passive interpretation, but the passive clause in (3b) with *okuru* ‘send’ cannot.

The *ni*-marked argument of a change-of-location verb like *okuru* ‘send’

patterns with locative adjuncts. The passive clauses in (4b) and (4c), formed from the clause in (4a) with the verb *hikidasu* ‘withdraw’ taking a locative-source PP as well as a theme RP (accusative object), illustrate the point.

- (4) a. *Dareka-ga kono ginkoo-kara okane-o hikidai-ta.*  
 someone-NOM this bank-ABL money-ACC withdraw-PST  
 ‘Someone withdrew money from this bank.’
- b. ● *Okane-ga kono ginkoo-kara hikidas-are-ta.*  
 money-NOM this bank-ABL withdraw-PASS-PST  
 ‘Money was withdrawn from this bank.’
- c. # *Kono ginkoo-ga okane-o hikidas-are-ta.*  
 this bank-NOM money-ACC withdraw-PASS-PST  
 ‘The bank was adversely affected by withdrawing money from it.’

(4b) is a direct passive clause where the RP *okane* ‘money’ in (4a) has been rendered as the passive subject, and has a neutral direct passive interpretation. By contrast, (4c) is construed as an adversity (or indirect) passive, where the subject is interpreted as emotionally affected—most typically ‘adversely’, and lacks the neutral interpretation that should be available for ordinary direct passive clauses.

The passive clause in (3b), just like (4c), is construed as an adversity (or indirect) passive, where the subject counts as an ‘affectee’. As often discussed in the Japanese literature (e.g. Kageyama 2006; Takezawa & Whitman 1998), the ‘affectee’ argument of an adversity passive is conceived of as being added to the clause with no promotion of an argument. Furthermore, it is often claimed (see Kishimoto 2001; Sadakane & Koizumi 1995; Miyagawa 1989; among others) that the *ni* marking assigned to nominals could be construed as a case marker or a postposition. Provided that RPs, but not PPs, can be promoted by direct passivization, it is reasonable to postulate that the indirect internal argument of *ataeru* ‘give’ is realized as an RP, where *ni* is construed as a dative case marker, while *okuru* ‘send’ selects an indirect internal argument realized as a PP, where *ni* represents a postposition. The lack of neutral passive interpretation in (3b) suggests that the direct internal argument of *okuru* ‘send’ cannot be promoted to a subject by direct passivization.

The data suggest that the three-place predicate construction headed by a change-of-possession verb like *ataeru* ‘give’ is construed as a double object construction (equivalent to *John gave Mary a book*), and the one headed by a change-of-location verb like *okuru* ‘send’ is a postpositional construction (comparable to the prepositional *to*-dative construction *John*



*gave a book to Mary*). This fact suggests that the distribution of dative marking is rather limited in Japanese: Dative case marking appears on the indirect internal arguments of three-place predicates denoting a change of possession, but not a change of location. It is worth noting that this conforms to a cross-linguistic tendency for the indirect internal arguments of change-of-possession verbs to receive dative case, while change-of-location verbs more often take indirect arguments marked with adpositions (see Van Velle & Van Langendonck 1996).

### 3. RRG account for case marking in three-place verbs

In this section, I suggest that the case-marking patterns of three-place predicates can be accounted for on the basis of Logical Structures (LSs) proposed by Van Valin (1993) (see also Van Valin & LaPolla 1997; Van Valin 2005). Specifically, I suggest that the case marking of indirect internal arguments can be determined by way of lexical representations which include a predicate specifying a source alongside a predicate specifying a goal, and that the prima facie dual behavior of dative marking follows naturally from the assumption that dative case is the default case assigned to a non-macrorole core argument, as proposed by Van Valin (1991) and Van Valin & LaPolla (2007).

#### 3.1. Change-of-possession verbs

To begin, change-of-possession verbs (as well as change-of-location verbs) are divided into two sub-classes depending on the directionality of described transfer—one in which the subject serves as a source (i.e. the ‘source-subject’ verb) and the other in which the subject counts as a goal (i.e. the ‘goal-subject’ verb). This difference in the orientation of transfer is easily grasped, given the proposal on the LS representation of three-place predicates advanced by Van Valin (2007).

To be concrete, Van Valin (2007) suggests that three-place predicates (verbs of change of possession) can be analyzed as having a causative representation like (5), which essentially means that the ownership of *y* is changed from *x* to *z*.

- (5) [D● (*w*, [do' (*w*, ∅)])] CAUSE [BECOME NOT have' (*x*, *y*) & BECOME have' (*z*, *y*)]

The LS in (5) is a representation for a verb denoting a change of posses-

sion, where the predicate **have'** ( $x, y$ ) specifies a possessive relation, and the connective & is used to refer to the states of affairs that take place successively (Van Valin & La Polla 1997: 109). **BECOME** is an operator associated with a predicate formula (Foley & Van Valin 1984: 38). The **D** operator indicates that the verb describes an agentive action, and **do'** ( $w, \emptyset$ ) is a generalized activity predicate, where the second argument is left unspecified as ' $\emptyset$ ' (see Dowty 1979; Foley & Van Valin 1984: 38-39; Van Valin 1993: 34-37; Van Valin & La Polla 1997: 103-104).<sup>3</sup> Since change-of-possession verbs like *ageru* 'give' denote transfer of possession from a prior possessor to a recipient (i.e. the succeeding possessor), the LS must include two instances of **have'**—**BECOME NOT have'** ( $x, y$ ), which specifies an event pertinent to a prior possessor, and **BECOME have'** ( $z, y$ ), which describes an event pertinent to a recipient.

Change-of-possession verbs are often analyzed as denoting only the change of state pertinent to the recipient with a representation like [x cause [y to have z]] (see Pinker 1989; Goldberg 1995; Harley 2003; etc.). In this representation, the argument  $x$ , which is realized as a syntactic subject, counts as a causer (and if the operator **D** is included, it is an agent). To be precise, however, this representation does not describe the exact meaning of a three-place verb like *ataeru* 'give'. Since the causer/agent subject of *ataeru* also counts as a prior possessor, there is a sense in which the LS should include a specification for a prior possessor as well as for a recipient, i.e. the LS should consist of predicates specifying two subevents of a change of possession.<sup>4</sup>

The LS in (5) serves as the basis for determining the distinction between goal-subject and source-subject verbs. In (5), the variable  $w$  for the causer is represented in Italics, since it is equated with either  $x$  or  $z$  when specifying the directionality of transfer. The different orientations of transfer can be fixed by identifying  $w$  as either  $x$  or  $z$ . When  $w$  is identified as  $x$ , the meaning of the source-subject verb *ataeru* 'give' is expressed.

- (6) a. *Eri-ga tomodati-ni manga-o atae-ta.*  
 Eri-NOM friend-DAT comic-ACC give-PST  
 'Eri gave her friend the comics.'
- b. [**D**( $x$  [**do'** ( $x, \emptyset$ )])] CAUSE [**BECOME NOT have'** ( $x, y$ ) & **BECOME have'** ( $z, y$ )]

<sup>3</sup> If a verb simply has a causer (or an effector) as its subject, the causing event in the LS in (5) is represented by [**do'** ( $w, \emptyset$ )] in lieu of [**D**( $w, [\mathbf{do}' (w, \emptyset)]$ )], as in [**do'** ( $w, \emptyset$ )] CAUSE [**BECOME NOT have'** ( $x, y$ ) & **BECOME have'** ( $z, y$ )].

<sup>4</sup> The  $x$  in **have'** ( $x, y$ ) is constrained by the animacy constraint, so it can only be inanimate in a 'part-whole' relation (Pinker 1989; Kishimoto 2005; see section 3).

The LS in (6b) (posited for (6a)) means that  $x$  (=Eri) causes the ownership of  $y$  (=the comics) to change from  $x$  to  $z$  (=her friend), where  $y$  is the theme and  $z$  is the goal. The source argument  $x$  in **BECOME NOT have'** ( $x, y$ ), which counts as the previous owner of the book, is associated with the **DO** operator, and thus identified as the willful doer of the action—the agent—as well (see Ross 1972; Dowty 1979).

On the other hand, if  $w$  is identified as  $z$ , the LS in (5) means that  $z$  causes the ownership of  $y$  to change from  $x$  to  $z$ , i.e. the LS names the act of getting, which is described by a goal-subject verb like *morau* 'get' in (7a).

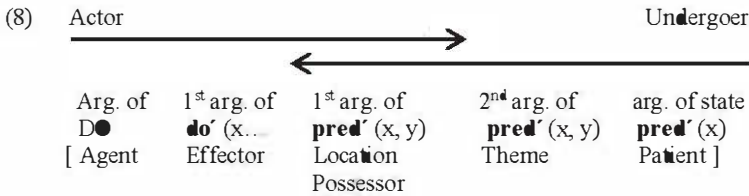
- (7) a. *Tomodati-ga Eri-ni manga-o morat-ta.*  
 friend-NOM Eri-DAT comic-ACC get-PST  
 'Her friends got the comics from Eri.'
- b. [**DO** ( $z, [\text{do}' (z, \text{O})]$ )] CAUSE [**BECOME NOT have'** ( $x, y$ )] &  
**BECOMEhave'** ( $z, y$ )]

In (7b), the recipient argument  $z$  in **BECOME have'** ( $z, y$ ), which counts as the new owner of the book, is associated with the **DO** operator, and thus is identified as the willful doer of the action. The LS in (7b) differs minimally from the one in (6b), in that in the former LS, the causer argument is identified as a goal (the recipient) rather than a source (the prior possessor).

With these LS representations for three-place predicates in mind, let us proceed to illustrate how morphological markings are assigned to their arguments. First, a three-place verb like *ageru* 'give' takes three participants: agent(source), goal(recipient), and theme. RRG postulates two generalized semantic roles (i.e. macroroles) called 'Actor' and 'Undergoer'. These two macroroles represent 'prototypical agent' and 'prototypical patient', respectively (cf. Dowty 1991).<sup>5</sup> The Actor/Undergoer status is determined in accordance with the hierarchy in (8) (Van Valin & LaPolla 1997: 127).

---

<sup>5</sup> For our purposes, it is sufficient to note that the RRG terms 'Actor' and 'Undergoer' broadly correspond to 'external argument' and 'internal argument', which are the terms commonly used in the generative literature.



The thematic labels added at the bottom of the hierarchy indicate some representative semantic relations determined by the position of arguments in LSs. Thematic relations may differ depending on the nature of *pred'*, so the label 'location', for instance, includes various 'locative-related' notions such as source, goal, recipient, possessor, and the like (see Pinker 1989; Van Valin 2005).

Two of the three arguments of a three-place predicate bear macroroles (Actor or Undergoer). Since Japanese is a nominative-accusative language, the macrorole arguments are assigned either nominative or accusative case in accordance with (9a-b).

- (9) a. The highest ranking macrorole argument is assigned nominative case.  
 b. The other macrorole argument is assigned accusative case.  
 c. Non-macrorole core arguments are assigned dative case as their default case.

Note that (9c) is the dative case assignment rule, which states that dative case is the default case assigned to non-macrorole core arguments (see Van Valin 1991; Van Valin & LaPolla 1997; for Japanese, see Imai 1998).

The macrorole ranking is 'Actor > Undergoer', which means that the Actor counts as the higher of the two macroroles, and therefore the agent/prior-possessor of the source-subject verb *ataeru* 'give' is identified as Actor, and the theme as Undergoer. Then, according to (9a-b), the 'agent/prior-possessor' argument is assigned nominative case, and the theme argument, accusative case. The remaining 'recipient' argument is a non-macrorole core argument, and thus receives dative case marking by the dative case assignment rule in (9c).

Essentially the same account applies to the case marking on the arguments of a goal-subject verb like *morau* 'get'. With *morau*, the recipient is realized as subject. This is because the recipient falls under the scope of the D● operator, and counts as the agent instigating an action (as well as the recipient of the entity transferred). The agent/recipient and theme arguments, which count as macrorole arguments, are construed as Actor and

Undergoer, respectively, in accordance with the Actor-Undergoer hierarchy. Further, by way of (9a-b), the former is assigned nominative case, and the latter, accusative case. The remaining source argument, which is a non-macrorole core argument, is assigned dative case in conformity with the dative case assignment rule in (9c).

Note that the range of thematic relations that the dative-marked arguments of three-place predicates can bear is restricted, in the sense that the dative arguments are anchored to a particular argument position in LSs, i.e. the first argument position of (NOT) **have'** (x, y), and are construed as either source or goal arguments. Nevertheless, the labels 'source' and 'goal' are distinct in thematic terms, and thus, no one-to-one correspondence can be posited between dative case marking and the thematic relations it can represent. In this respect, dative arguments are distinguished from oblique expressions with fixed thematic relations. This facet of dative case marking follows naturally from its characterization as the default case assigned to a non-macrorole core argument.

### 3.2. Change-of-location verbs

Let us now turn to the discussion of change-of-location verbs. In RRG, a locative relation is expressed by the predicate **be-LOC'** (x, y)—most typically, **be-at'** (x, y)—where x and y represent 'location' and 'theme', respectively.<sup>6</sup> Three-place predicates expressing a change of location need to encode two different locative relations—namely, the locative-source, which corresponds to the locative argument x in **BECOME NOT be-at'** (x, y), and the locative-goal, which corresponds to the locative argument x in **BECOME be-at'** (x, y) (or **BECOME be-toward'** (x, y)) (see Gruber 1965).<sup>7</sup> Thus, it can be postulated that change-of-location verbs have the LS in (10).

- (10) [D $\bullet$  (w, [do' (w,  $\emptyset$ )])] CAUSE [BECOME NOT **be-at'** (x, y) & BECOME **be-at'/be-toward'** (z, y)]

<sup>6</sup> In RRG, 'locative' is specified as the first argument x of **be-LOC'** (x, y).

<sup>7</sup> For change-of-location verbs, the causer does not have to be identified with a location. Thus, this type of verb can take two indirect internal arguments representing a source and a goal, as in (i).

- (i) *John-wa zitaku-kara daigaku-ni tegami-o okut-ta.*  
 John-TOP home-ABL university-LOC letter-ACC send-PST  
 'John sent a letter from his home to the university.'

For change-of-location verbs, which carry the core meaning of ‘y moves from x to z’, the directionality of movement can be determined in essentially the same way as verbs of change of possession, i.e. the direction of transfer can be fixed by identifying *w* as either x or z in the LS in (10).

If  $w=x$ , then the LS in (10) signifies the act of sending. Thus, a source-subject predicate like *okuru* ‘send’ in (11a) can be regarded as possessing the LS in (11b).

- (11) a. *Eri-ga sensei-ni tegami-o okut-ta.*  
 Eri-NOM teacher-LOC letter-ACC send-PST  
 ‘Eri sent a letter to the teacher.’  
 b. [D●(x, [do'(x, O))]] CAUSE [BECOME NOT be-at'(x, y) & BECOME be-toward'(z, y)]

The LS in (11b) specifies the meaning of ‘x (=Eri) causes y (=the letter) to move from x toward z (=the teacher)’.<sup>8</sup> Since the argument x occurs as the first argument lying under the operator D●, it serves as an agent. The argument x is also the first argument of BECOME NOT be-at'(x, y), so it counts as a locative-source as well.

If  $w=z$ , then the LS in (10) signifies the act of receiving. The goal-subject verb *uketoru* ‘receive’ in (12a) has the lexical representation in (12b).

- (12) a. *Sensei-ga Eri-kara hon-o uketot-ta.*  
 teacher-NOM Eri-ABL book-ACC receive-PST  
 ‘The teacher received a book from Eri.’  
 b. [D●(z, [do'(z, O))]] CAUSE [BECOME NOT be-at'(x, y) & BECOME be-at'(z, y)]

The LS in (12b) conveys the meaning of ‘z (=the teacher) causes y (=the book) to go from x (=Eri) to z’. With *uketoru* ‘receive’, a completed change of location is described, which suggests that the LS for *uketoru* should include be-at'(z, y), where the first argument z specifies a locative-goal, as in (12b).

In (11a), the goal argument *sensei* ‘the teacher’ does not count as a recipient, since no ownership is conferred to the teacher through the act of receiving. Rather, the argument serves as a locative-goal, which needs to be obliquely marked. I suggest that the locative-goal argument is assigned

---

<sup>8</sup> The verb *okuru* ‘send’ does not entail that the theme *tegami* ‘letter’ reaches the endpoint *sensei* ‘teacher’.

the postposition *ni* in accordance with the postposition assignment rule given in (13).

- (13) The postposition *ni* is assigned to the non-macrorole core argument *x* in the LS segment... **BECOME be-at'/toward'** (*x, y*).

In (13), the goal specified by **BECOME be-at'** (*x, y*) or **BECOME be-toward'** (*x, y*) is realized as an argument with the postposition *ni* (i.e. **BECOME be-at'/toward'** (*x, y*) → *x-ni*) when it serves as a non-macrorole core argument. (Note that the locative marker *ni* can specify a direction, as in *eki-ni mukau* [station-for head] 'head for the station' as well as the endpoint, as in *eki-ni tuku* [station-at arrive] 'arrive at the station'.) Needless to say, the default dative case assignment rule does not apply to a non-macrorole core argument which receives a postposition.

In regard to the question of how case marking is assigned to the arguments of change-of-location verbs, observe that the source-subject verb *okuru* 'send' in (11a) has three core arguments, and that the LS in (11b) (posited for (11a)) includes two instances of **be-LOC'**—one which occurs with **NOT** and the other which does not. Since the argument *x* is an agent, it is realized as the higher macrorole argument (Actor) and the theme serves as the lower macrorole argument (Undergoer) in accordance with (8). The former is assigned nominative case, and the latter accusative case by the rules in (9a-b). The locative argument of *okuru* 'send' appears as the first argument of **BECOME be-toward'** (*x, y*), and is not a macrorole argument. Thus, the locative argument occurs with the postposition *ni*, since the rule in (13) applies to it (i.e. **BECOME be-at'/toward'** (*x, y*) → *x-ni*).

The goal-subject verb *uketoru* 'receive' in (12a) takes three arguments of agent/goal, source, and theme. Among them, the agent/goal and the theme are construed as Actor and Undergoer, respectively, according to the Actor-Undergoer hierarchy in (8). The agent/goal argument is assigned nominative case, and the theme argument is assigned accusative case by the rules in (9a-b). The indirect internal argument of the goal-subject verb *uketoru* 'receive' in (12a) is a locative-source, which needs to be obliquely marked. I suggest that the locative-source argument is assigned the postposition *kara* via the postposition assignment rule in (14).

- (14) The postposition *kara* is assigned to the non-macrorole core argument *x* in the LS segment ... **BECOME NOT be-at'** (*x, y*).

The non-macrorole source argument of the goal-subject verb *uketoru* is a locative-source, so it is assigned oblique *kara* marking in accordance with

(14) (i.e. BECOME NOT **be-at'** ( $x, y$ )  $\rightarrow$   $x$ -*kara*).

The postposition-assignment rules in (13) and (14) are sensitive to the negative operator NOT in the LS. If **be-at'** ( $x, y$ ) does not accompany NOT, a non-macrorole locative argument is given the postposition *ni*. If NOT occurs with **be-at'** ( $x, y$ ), the postposition *kara* 'from' is allocated to a non-macrorole locative argument. Consequently, with change-of-location verbs, different postpositions are assigned to their indirect internal arguments depending on the directionality of movement expressed by the verbs.

#### 4. The dual behavior of source arguments

As discussed previously, the source argument of a change-of-possession verb can be marked with *kara* 'from' alongside the dative *ni*. The postposition *kara* designates a starting point (which could be either a location or a point of time), but does not specify a goal, indicating that *kara* does not function in the same manner as the dative case *ni*. The question now arises as to why *kara* 'from' can be replaced with the dative case *ni* on the indirect internal argument of the goal-subject type of change-of-possession verb. In this section, I suggest that a *kara*-marked argument appears in the three-place verb construction when the verb includes a lexical specification for a locative-source in the LS.

##### 4.1. The animacy constraint

It is instructive to begin by noting a semantic constraint imposed on possessors. As often discussed (see Jackendoff 1983, 1990; Pinker 1989; Van Valin & LaPolla 1997; among others), the spatial notion of 'location' (represented by **be-LOC'** ( $x, y$ )) is distinct from, but closely related to, the notion of 'possession' (expressed by **have'** ( $x, y$ )). Therefore, these two semantic relations display a number of distinct properties even though they often show certain parallelisms. To take a few examples, while a spatial relation can often invoke an incremental transition from a source to a goal, transfer of possession lacks internal structure, and hence cannot be incremental. Spatial relations come in many different varieties, but possessive relations do not. Thus, different locative relations are expressed by different prepositions/postpositions, but the prepositions/postpositions that can express possessive relations are highly restricted.

Furthermore, possessors and locatives show a difference in susceptibility to the animacy constraint (see Pinker 1989). Note that the argument filling the variable  $x$  of **have'** ( $x, y$ )—but not **be-LOC'** ( $x, y$ )—must qual-



fy as a possessor, and hence must be animate. Thus, it is not possible for *ataeru* ‘give’ to have an inanimate noun as its *ni*-marked argument, while *okuru* ‘send’ allows for an inanimate *ni*-marked argument.

- (15) a. *Hahaoya-wa {kodomo-ni/\*Tokyo-ni} okane-o atae-ta.*  
 mother-TOP {child-DAT/Tokyo-DAT} money-ACC give-PST  
 ‘Mother gave {the child/\*Tokyo} money.’  
 b. *Hahaoya-wa {kodomo-ni/Tokyo-ni} nimotu-o okut-ta.*  
 mother-TOP {child-LOC/Tokyo-LOC} package-ACC send-PST  
 ‘Mother sent the package to {the child/Tokyo}.’

Only an animate noun qualifies as the indirect internal goal argument of *ageru* ‘give’. Accordingly, in (15a), the animate *kodomo* ‘child’ is allowed as the goal argument, but not *Tokyo*, which is only interpreted as a location. In (15b), by contrast, *kodomo*, as well as *Tokyo*, can occur as the goal argument of the verb *okuru*. In (15b), *kodomo* ‘child’ happens to refer to a human, but is thematically construed as a location—the goal for the letter to reach via spatial movement (with no transfer of possession denoted). Since the animate noun *kodomo* can be identified as a locative-goal, the locative-goal argument of *okuru* ‘send’ can be either inanimate (e.g. *Tokyo*) or animate (e.g. *kodomo*).

As I will discuss in the next section, a similar difference in susceptibility to the animacy constraint emerges on the indirect internal argument of change-of-possession verbs according to whether it is marked with dative case or ablative *kara*.

## 4.2. The status of *kara*-marked arguments

As noted earlier, the source argument of change-of-possession verbs may receive either dative *ni* marking or ablative *kara* marking. Interestingly, with goal-subject predicates, the difference in the source marking gives rise to a difference in susceptibility to the animacy constraint. In the first place, when the source argument of *morau* receives dative *ni* marking, it cannot refer to an inanimate entity, but the animacy constraint is voided if *kara* marking substitutes for *ni* marking, as shown in (16).

- (16) a. *Eri-wa {otoosan-ni/\*otoosan-no tyokinbako-ni} okane-o*  
 Eri-TOP {father-DAT/father-GEN piggybank-DAT} money-ACC  
*morat-ta.*  
 get-PST  
 ‘Eri got money from {her father/her father’s piggybank}.’

- b. *Eri-wa {otoosan-kara/otoosan-no tyokinbako-kara} okane-o*  
 Eri-TOP {father-ABL/father-GEN piggybank-from} money-ACC  
*morat-ta.*  
 get-PST  
 ‘Eri got money from {her father/her father’s piggybank}.’

The absence of an animacy effect on the *kara*-marked indirect internal argument in (16b) suggests that the *kara*-marked source argument does not count as a possessor. In view of this fact, I propose that the source argument of a goal-subject verb is marked with *kara* ‘from’ if the possessive predicate **have**’ (x, y) is changed to the locative predicate **be-at**’ (x, y) via the lexical rule in (17) (for discussion on lexical rules changing LSs, see Van Valin & LaPolla 1997: 178-184).

- (17) **BECOME NOT have**’ (x, y) → **BECOME NOT be-at**’ (x, y)

For a goal-subject verb like *morau* ‘get’, the LS in (18b) can be derived from the LS in (18a) by effecting a change on the underlined predicate sequence.

- (18) a. [D● (z, [do’ (z, ∅)])] CAUSE [**BECOME NOT have**’ (x, y) & **BECOME have**’ (z, y)]  
 b. [D● (z, [do’ (z, ∅)])] CAUSE [**BECOME NOT be-at**’ (x, y) & **BECOME have**’ (z, y)]

●nce (18b) is derived, the indirect internal argument, which is the non-macrorole ‘source’ argument, can be assigned ablative *kara*, following the *kara*-assignment rule in (14) (i.e. **BECOME NOT be-at**’ (x, y) → *x-kara*). If the assignment of the postposition *kara* ‘from’ to the indirect internal argument of a goal-subject verb like *morau* ‘get’ is based on the LS in (18b), the *kara*-marked argument is construed as a locative-source rather than a prior possessor. This analysis can provide a ready account for the facts regarding the animacy constraint. The dative-marked source argument in (16a) is subject to the animacy constraint because it is the first argument of **have**’ (x, y), but the *kara*-marked source argument in (16b) is not constrained by the animacy constraint because it is a first argument of **be-at**’ (x, y).

The proposed analysis gains further support from the variable behavior verb *toru* ‘take’, which can take three arguments. Because of the verb’s elasticity in meaning, a close look at its behavior provides us with additional support for the analysis taking change-of-possession verbs to invoke

a change in their LS when they take *kara*-marked rather than dative-marked indirect internal arguments.

To be concrete, observe that the possibility of dative marking on the indirect internal argument of a verb like *toru* 'take' varies depending on whether or not transfer of possession is conceptualized.

- (19) a. *Ano hito-wa {sensei-ni/sensei-kara} kakunin-o*  
 that man-TOP {teacher-DAT/teacher-ABL} confirmation-ACC  
*tot-ta.*  
 take-PST  
 'The man got confirmation from the teacher.'  
 b. *Ano hito-wa {\*sensei-ni/sensei-kara} okane-o tot-ta.*  
 that man-TOP {teacher-DAT/teacher-ABL} money-ACC take-PST  
 'The man got money from the teacher.'

The source-subject verb *toru* 'take' carries different meanings depending on the choice of its object. When the object is *kakunin* 'confirmation', transfer of authorization (which can be conceived of as a kind of transfer of possession) is expressed, so its source argument can be marked with dative case. It is also possible for the source argument to be marked with *kara*, since it can be a locative. When the object is *okane* 'money', the act of stealing is described. In this case, the owner does not give an authorization for a change of ownership, i.e. no transfer of ownership is denoted. Accordingly, the source argument, which is a non-macrorole core argument, cannot count as a (previous) possessor. In this case, since the source argument can only be a locative-source, it can only be marked with ablative *kara*.

In this connection, observe that if the source argument of the verb *toru* 'take' describing a change of possession is marked with dative case, it can only be animate, but that if the source argument is marked with *kara*, it can be either animate or inanimate, as shown in (20).

- (20) a. *Ano hito-wa {sono sensei-ni/\*sono syorui-ni}*  
 that man-TOP {that teacher-DAT/that document-DAT}  
*kakunin-o tot-ta*  
 confirmation-ACC take-PST  
 'That man got a confirmation from {that teacher/that document}.'  
 b. *Ano hito-wa {sono sensei-kara/sono syorui-kara}*  
 that man-TOP {that teacher-ABL/that document-from}  
*kakunin-o tot-ta.*  
 confirmation-ACC take-PST  
 'That man got a confirmation from {that teacher/that document}.'

The data illustrate that with the verb *toru* expressing the meaning of the transfer of knowledge, its source argument is subject to the animacy constraint when marked with dative case, but not when marked with *kara*. The facts fall out naturally, given that the verb *toru* ‘take’ taking three arguments can have the two distinct LSs in (21).

- (21) a. [D●(z, [do'(z, ∅))]] CAUSE [BECOME NOT have'(x, y) & BECOME have'(z, y)]  
 b. [D●(z, [do'(z, ∅))]] CAUSE [BECOME NOT be-at'(x, y) & BECOME have'(z, y)]

The difference in the marking of the source argument is determined by the specification of the source in the LS. In (20a), the source argument, i.e. the non-macrorole core argument of the verb *toru*, counts as the first argument of *have'* (x, y) in the LS in (21a). Thus, this argument receives dative marking by the default dative case assignment rule in (9c), and the dative argument must refer to an animate entity by virtue of the animacy constraint. On the other hand, (20b) is acceptable regardless of whether the indirect internal argument is animate or inanimate. This fact follows straightforwardly if the verb has the LS in (21b) (derived from (21a) via the lexical rule in (17)). In the LS in (21b), the source is construed as a locative argument thematically, as it appears as the first argument of NOT *be-at'* (x, y). The *kara*-marked direct internal argument in (20b) can be animate, since an animate entity, as well as an inanimate entity, can be construed as a locative, as discussed in section 4.1.

Moreover, when the verb *toru* takes *okane* ‘money’ as its object, the source argument, which must be marked with ablative *kara*, can be animate or inanimate, as shown in (22).

- (22) *Ano hito-wa {sono sensei-kara/saihu-kara} okane-o tot-ta.*  
 that man-TOP {that teacher-ABL/purse-ABL} money-ACC take-PST  
 ‘That man got money from {that teacher/the purse}.’

The sentence describes an act of stealing. In this case, the *kara*-marked argument is unambiguously construed as a locative argument, since no transfer of ownership is conceived by the act of stealing. (22) also illustrates that the argument to which *kara* marking is assigned is not identified as a possessor.

To give one more concrete example, observe that the verb (*hanasi o*) *kiku* ‘hear (a story)’ can express the meaning of transfer of knowledge. This verb can take an inanimate noun as its source argument when it is marked with ablative *kara*. But when the source argument is marked with

dative *ni*, it cannot be inanimate.

- (23) a. *Watasi-wa {kare-kara/kare-no kuti-kara} zikani sono*  
 I-TOP {he-ABL/he-GEN mouth-ABL} directly that  
*hanasi-o kii-ta.*  
 story-ACC hear-PST  
 'I heard that story directly from {him/his mouth}.'
- b. *Watasi-wa {kare-ni/\*kare-no kuti-ni} zikani sono*  
 I-TOP {he-DAT/he-GEN mouth-DAT} directly that  
*hanasi-o kii-ta.*  
 story-ACC hear-PST  
 'I heard that story directly from {him/\*his mouth}.'

The fact shows, again, that since dative case is assigned to a possessor argument, the dative-marked source argument cannot be inanimate. By contrast, the *kara*-marked indirect internal argument is a locative-source, so that it can take either an animate or an inanimate noun. The fact that source arguments show distinct behaviors depending on their marking follows naturally, given that a non-macrorole core argument counting as the first argument of **have'** (x, y) is assigned default dative case *ni* (by the rule in (9c)), while ablative *kara* is assigned to the source argument counting as the argument x in the LS segment: **BECOME NOT be-at'** (x, y) (by the rule in (14)).

In a nutshell, with goal-subject verbs expressing transfer of possession, the difference in the source marking gives rise to a difference in susceptibility to the animacy constraint. When the source argument receives dative *ni* marking, it cannot refer to an inanimate entity. The animacy effect is not imposed on the *kara*-marked source argument because it does not count as a possessor. This shows that the indirect internal source argument of a three-place goal-subject verb is marked with dative case when the verb has the LS: [D $\bullet$  (z, [do' (z, O)])] CAUSE [BECOME NOT **have'** (x, y) & BECOME **have'** (z, y)]. *Kara* marking is assigned to the indirect internal source argument if it represents a location rather than a possessor, i.e. if the three-place goal-subject verb has the LS: [D $\bullet$  (z, [do' (z, O)])] CAUSE [BECOME NOT **be-at'** (x, y) & BECOME **have'** (z, y)].

## 5. Conclusion

In this article, it has been shown that the distribution of *ni* marking assigned to the indirect internal arguments of Japanese three-place predicates

can be optimally characterized on the basis of the verb's inherent lexical meaning. Dative case is assigned to a non-macrorole core argument by the rule for default dative case assignment. Since this case assignment is implemented on change-of-possession verbs, the dative case appears on their possessor argument regardless of whether it counts as a source or a goal, i.e. dative case *ni* is assigned to the non-macrorole core argument *x* appearing in the predicate *have'* (*x*, *y*) regardless of whether it occurs with the negative operator NOT, which contributes to differentiating between the two thematic labels of 'source' and 'goal'. This analysis can provide a ready account for why both source and goal arguments of change-of-possession verbs can be assigned the dative case *ni* even though they represent opposite directionality of transfer.

It has also been proposed that the postposition *ni* is assigned to the non-macrorole locative argument *x* in BECOME *be-at'* (*x*, *y*). Thus, the indirect internal goal argument of a change-of-location verb like *okuru* 'send' is marked with the postposition *ni*. This postposition cannot be used for marking the non-macrorole source argument of a change-of-location verb like *uketoru* 'receive'. The source argument is instead assigned the postposition *kara* 'from' via the *kara* assignment rule which applies to the non-macrorole locative argument specified by BECOME NOT *be-at'* (*x*, *y*). With three-place change-of-location verbs, dative case is not assigned to their indirect internal arguments because locative arguments are obliquely marked.

The present analysis making use of decomposed lexical representations of verbs allow us to capture some prima facie puzzling case-marking properties of the indirect internal arguments of three-place predicates. In particular, the proposed RRG analysis can readily capture the case-marking patterns of three-place predicates that are found in many languages—the fact that change-of-possession verbs allow both goal (i.e. the recipient) and source (i.e. the prior possessor) arguments to be assigned dative case marking when they are realized as indirect internal arguments.

## References

- Blake, Barry. 1994. *Case*. Cambridge: Cambridge University Press.
- Blansitt, Edward. 1988. Datives and allatives. In Hammond, Michael & Moravcsik, Edith & Wirth, Jessica (eds.), *Studies in syntactic typology*, 173–191. Amsterdam: John Benjamins.
- Dowty, David. 1979. *Word meaning and montague grammar*. Dordrecht: Reidel.

- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.
- Foley, William & Van Valin, Robert. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Gruber, Jeffrey. 1965. *Studies in lexical relations*. Cambridge: MIT. (Doctoral dissertation.)
- Harley, Heidi. 2003. Possession and the double object construction. *Linguistic variation yearbook* 2: 31–70.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Tomasello, Michael (ed.), *The new psychology of language Vol. 2*, 211–243. Mahwah, NJ: Laurence Erlbaum.
- Imai, Shingo. 1998. *Logical structure and case marking in Japanese*. New York: State University of New York at Buffalo. (MA thesis.)
- Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.
- Janda, Laura. 1993. The shape of the indirect object in central and eastern Europe. *The Slavic and East European Journal* 37: 533–563.
- Kageyama, Taro. 2006. Property description as a voice phenomenon. In *Voice and grammatical relations*, Tsunoda, Tasaku & Kageyama, Taro (eds.), 85–114. Philadelphia: John Benjamins.
- Kishimoto, Hideki. 2001. The role of lexical meanings in argument encoding: Double object verbs in Japanese. *Gengo Kenkyu* 120: 35–65.
- Kishimoto, Hideki. 2005. *Toogo koozoo-to bunpoo kankei* [Syntactic structure and grammatical relations]. Tokyo: Kurosio Publishers.
- Levin, Beth & Rappaport Hovav, Malka. 1995. *Unaccusativity*. Cambridge, MA: MIT Press.
- Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard. 2010. *Ditransitive constructions: A typological overview*. In Malchukov, Andrej, Haspelmath, Martin & Comrie, Bernard (eds.), *Studies in ditransitive constructions: A comparative handbook*, 1–64. Berlin: De Gruyter Mouton.
- Maling, Joan. 2001. Dative: The heterogeneity of the mapping among morphological case, grammatical functions and thematic roles. *Lingua* 111: 419–464.
- Martin, Samuel. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.
- Matsumoto, Yo. 2000. *Osowaru/osieru nado no tadoosi* [Transitive verbs

- like *osieru/osowaru*]. In Yamada, Susumu & Kikuti, Yasuto & Momi-yama, Yosuke (eds.), *Nihongo: imi-to bunpoo-no huukei* [Japanese: A landscape from meaning and grammar], 79–95. Tokyo: Hituzi Syobo.
- Miyagawa, Shigeru. 1989. *Syntax and semantics 22: Structure and case marking in Japanese*. San Diego, CA: Academic Press.
- Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Ross, John. 1972. Act. In Davidson, Donald & Herman, Gilbert (eds.), *Semantics of natural languages*, 70–126. Dordrecht: D. Reidel.
- Sadakane, Kumi & Koizumi, Masatoshi. 1995. On the nature of the “dative” particle *ni* in Japanese. *Linguistics* 33: 5–33.
- Shimizu, Yasuyuki. 1977. *Syoyuu-no idoo-o arawasu doosi-no kaku-to wakugumi* [The case and frame of verbs that express transfer of possession]. In *Matumura Akira kanreki kinen: Kokugogaku-to kokugosi* [A festschrift for Professor Akira Matsumura: Japanese linguistics and Japanese linguistic history], 39–35 (1184–1170). Tokyo: Meiji Shoin.
- Sugimoto, Takeshi. 1986. *Kakuzyoshi ga, o, ni to bunpookankei* [Case particles *ga*, *o* and *ni*, and their grammatical relations]. In Okutsu, Keiichiro (ed.), *Iwayuru nihongo-zyosi-no kenkyu* [A study of so-called particles in Japanese], 227–379. Tokyo: Bonjinsha.
- Takezawa, Koichi & Whitman, John. 1998. *Kaku-to gozyun-to toogokoo-zoo* [Case, word order and syntactic structure]. Tokyo: Kenkyusha.
- Van Belle, William & Van Langendonck, Willy. 1996. *The dative: Volume 1: Descriptive studies*. Amsterdam: John Benjamins.
- Van Valin, Robert. 1991. Another look at Icelandic case marking and grammatical relations. *Natural Language & Linguistic Theory* 9: 145–194.
- Van Valin, Robert. 1993. A synopsis of Role and Reference Grammar. In Van Valin, Robert (ed.), *Advances in Role and Reference Grammar*, 1–164. Amsterdam: John Benjamins.
- Van Valin, Robert. 2004. Semantic macroroles in Role and Reference Grammar. In Kailuweit, Rolf & Hummel, Martin (eds.), *Semantische Rollen*, 62–82. Tübingen: Gunter Narr Verlag.
- Van Valin, Robert. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert. 2007. The Role and Reference Grammar analysis of three-place predicates. *Suvremena Lingvistika* 33.1 (63): 31–64.
- Van Valin, Robert. 2018. Dative case and oblique subjects. In Jóhanna Barðdal, Pat-El, Na’ama & Carey, Stephen Mark (eds.), *Non-canonically case-marked subjects. The Reykjavík-Eyjafjallajökull papers*, 115–131. Amsterdam: John Benjamins.



- Van Valin, Robert. & LaPolla, Randy. 1997. *Syntax: Structure, meaning and syntax*. Cambridge: Cambridge University Press.
- Williams, Edwin. 1981. Argument structure and morphology. *The Linguistic Review* 1: 81–118.

# PURPOSIVE CASE AND SEMANTIC ARGUMENTS IN AUSTRALIAN WESTERN DESERT DIALECTS

CONOR PYLE

TRINITY COLLEGE DUBLIN

## Abstract

This chapter examines the use of the case marker *-la:* in Pitjantjatjara and Yankunytjatjara, two dialects of the Western Desert group of Australia, using the Role and Reference Grammar (RRG) framework. This case ending is prototypically used in purpose and non-body part possession (both alienable and inalienable). Case polysemy is not unexpected however (Goddard 1991a) and *-la:* has other functions including stimulus in emotion verbs, goal and beneficiary. We find it is used frequently with apparently syntactically intransitive but semantically transitive verbs. A common feature is that there are unaffected participants in the scene. We extend the analysis to include sub-clauses based on nominalised verbs, where *-la:* indicates different subject purposive.

## Keywords

Western Desert, purpose, dative, emotion verbs, case syncretism

## 1. Introduction

Pitjantjatjara and Yankunytjatjara (P/Y) are closely related dialects of Australia's Western Desert language group, which covers a large area of Central Australia. They are part of the Pama-Nyungan family containing most languages outside the Top End of Australia, and they are classified as suffixing languages (Dixon 2011: 469). The two dialects are similar enough to share dictionaries and grammars (Goddard 1993, 1996). They are dependent-marking; there is no marking on the verb for person, gender or number. Verbs are generally classed as syntactically transitive or intransitive (with a handful of ambitransitives) and they are divided into four conjugations with only a few minor irregular forms (Goddard 1993: 10-

11). Case marking is on the reference phrase (RP) rather than individual elements.

The choice of RRG in this discussion is motivated by its universality, its emphasis on communication, and by the fact that it can be used to analyse the semantic structure of predicates (see Nolan 2012: 2ff, Van Valin 2001: 205) and thus the nature of arguments and non-arguments. In this chapter we are concerned with the nature of the *-ku* case ending, described within this framework. Van Valin & LaPolla (1997: 25) draw two universal semantic distinctions in language: predicating versus non-predicating elements; and phrases that are arguments of the predicate versus those that are not. These are shown in the syntactic categories of nucleus, core and periphery in the layered structure of the clause (LSC) (Nolan 2012: 5) shown in Figure 1.

Arguments themselves may be of three types: direct core arguments (DCAs), oblique core arguments (do not contribute to the semantic representation) and argument-adjuncts (contribute to the semantic representation) (VV p.c.). The latter two are distinguished by non-DCA case marking. Placing constituents marked *-ku* within the LSC forms a central element of the current investigation.

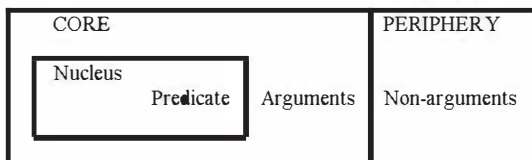


Figure 1: Layered Structure of the Clause (LSC)

Arguments of predicates fill semantic roles such as agent, patient, experiencer, theme, benefactive and others (Kim & Sells 2008: 43-45). These roles may be generalised into two “macroroles”, the actor and undergoer, respectively prototypical “doers” and those that are “affected” by the action (Van Valin & LaPolla 1997: 140). Semantics in RRG is linked to syntax for the speaker’s perspective and syntax to semantics for the addressee’s processing (Pavey 2010: 298), summed up in Figure 2.

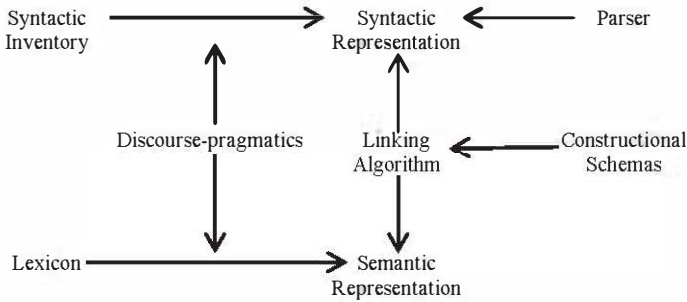


Figure 2: Linking algorithm

In semantic analysis RRG lexically decomposes propositions into the logical structure (LS) of predicates and arguments. In order to look at this it is useful to group predicates into classes based on their characteristics. This has a bearing on the argument structure and nature of the participants, which is central to this chapter. These classes are based on *Aktionsarten* and are valid cross-linguistically; they are summarised in Table 1. Tests can be done to categorise any particular predicate.

Table 1: Characteristics of predicates

	Static	Change of state	Dynamic	Telic	Punctual
State	+	-	-	-	-
Activity	-	-	+	-	-
Achievement	-	+	-	+	+
Semelfactive	-	-	±	-	+
Accomplishment	-	+	-	+	-
Active Accomplishment	-	+	+	+	-

In the LS the predicate is placed in bold with a prime (Van Valin 2007) and the arguments in brackets afterwards. These forms are a semantic metalanguage and do not represent words in any particular language (Chang 2007). Aspect, iteration and duration can all be marked on predicates by operators; however, these can also be indicated by the lexical semantics of the predicate itself (Pavey 2010: 93-94). The actor-undergoer hierarchy assigns macroroles depending on the nature of the predicate and the position of the argument in the LS. Importantly, RRG distinguishes S-transitivity (the number of syntactic DCAs), M-transitivity (the number of

macroroles) and semantic valence (the total number of arguments of the predicate in the nucleus).

Grammatical relations have two distinct properties (Van Valin & LaPolla 1997: 250):

- morphological coding such as case and verb agreement
- behavioural properties defining the role of NP<sup>1</sup> in grammatical constructions

Grammatical relations in P/Y are indicated by case marking but not verb agreement. Sentences in P/Y have a basic SOV order (Bowe 1990: viii, Goddard 1983: 20-21), but there is some freedom of word order, enabled by the case marking. Core RPs in P/Y generally precede the verb with other RPs after (Goddard & Harkins 2002: 214). Thus, adverbs such as locations tend to be clause-final (D. Rose p.c.), giving AOVX word order.

## 2. P/Y case system

Three basic roles are associated with syntactically intransitive and transitive verbs: S, the single argument of an intransitive verb; A<sub>T</sub>, the actor of a transitive verb; and U<sub>T</sub>, the undergoer of a transitive verb (Van Valin 2005: 96-97). These are marked with cases of different kinds: for example, P/Y has an *ergative-absolutive* nominal case system, grouping S and U<sub>T</sub> together in absolutive case. By contrast it has a *nominative-accusative* pronominal system. Case assignment rules are based on macrorole status (King 2010): in ergative systems the lowest ranking macrorole argument receives absolutive case and the highest ranked one receives ergative (Van Valin 2005: 108-110).

Cases in Australian languages have been studied under theories apart from RRG. Dixon (2011: 293-294) groups Australian language case functions generally into core and peripheral. Core case markings are on the arguments of a predicate, such as ergative, nominative, accusative and absolutive. The peripheral cases may be local, describing the location or movement of the action, or syntactic, adding further information such as the indirect object or goal. Dixon claims that there are twelve case functions in Australian languages, but that no language has them all as there is overlap and sharing of structures. In discussing case syncretism, Blake (1987: 52), groups dative, purposive, benefactive and genitive as “indirect object”, a term usually associated with the dative. This fits in well with the

---

<sup>1</sup> More recently Reference Phrase (RP).

situation in P/Y with the case marked by *-ku*: if we adopted Dixon's scheme, P/Y noun cases could be presented as in Table 2.

It might appear that Dixon's functions correspond to the RRG elements: the syntactic core with arguments and the periphery with non-arguments. As noted though, RRG distinguishes the core with direct core arguments, oblique core arguments and argument-adjuncts from the periphery with adjuncts. Notably the indirect object (marked "dative") is in the RRG core rather than the syntactic periphery. In this chapter we seek to characterise the *-ku* case, which appears in several places in the table, in RRG terms.

The *-ku* form is used with nouns of all types as well as with demonstratives (Goddard 1996: 42-43), but not with most pronouns. As seen in Table 3, the equivalent form is *-mpa* for pronouns, the exception being the *-ku* ending in *ngayuku* 'my, mine, for me'.

Table 2: P/Y nominal case endings according to Dixon's scheme

		Common	Proper
Core	ABSolutive (S, UT)	-Ø (- <i>pa</i> )	- <i>nya</i>
	ERGative (Ar)	- <i>ngku</i> (- <i>tju</i> , - <i>tu</i> )	- <i>lu</i>
Local peripheral	LOCative	- <i>ngka</i> (- <i>tja</i> , - <i>ta</i> )	- <i>la</i>
	ALLative	PURP + - <i>tu</i>	L●C+ PURP+ <i>-tu</i>
	ABLative	- <i>nguru</i>	L●C + - <i>nguru</i>
	TRANSverse	- <i>wanu</i>	L●C + - <i>wanu</i>
Syntactic peripheral	PURPositive/ POS-essive	- <i>ku</i>	
	DATive	L●C/PURP	
	CAUSal	- <i>tjara</i> , L●C, ABL	
	INSTRumental	L●C/ERG	
	AVERSive	L●C + - <i>tawara</i>	

Table 3: P/Y pronouns

Singular	Subject (S/AT)	Object (UT)	Purposive/Possessive
1 <sup>st</sup> Person	<i>ngayulu</i>	<i>ngayunya</i>	<i>ngayuku</i>
2 <sup>nd</sup> Person	<i>nyuntu/nyuntulu</i>	<i>nyuntunya</i>	<i>nyuntumpa</i>
3 <sup>rd</sup> Person	<i>paluru</i>	<i>palunya</i>	<i>palumpa</i>
Dual			
1 <sup>st</sup> Person	<i>ngali</i>	<i>ngalinya</i>	<i>ngalimpa</i>
2 <sup>nd</sup> Person	<i>nyupali</i>	<i>nyupalinya</i>	<i>nyupalimpa</i>
3 <sup>rd</sup> Person	<i>pula</i>	<i>pulanya</i>	<i>pulampa</i>
Plural			
1 <sup>st</sup> Person	<i>nganana</i>	<i>nganananya</i>	<i>nganampa</i>
2 <sup>nd</sup> Person	<i>nyura</i>	<i>nyuranya</i>	<i>nyurampa</i>
3 <sup>rd</sup> Person	<i>tjana</i>	<i>tjananya</i>	<i>tjanampa</i>

There are no allomorphs of *-ku* with place of articulation assimilation (Goddard 1993: 17) as occurs with ergative *-ngku* or locative *-ngka*. In the following sections we look at the uses and functions of *-ku*, and attempt to relate the various uses.

### 3. Possession: *-ku* within a reference phrase

In contrast to relatively flexible (though commonly S●V) order in clauses, there is a constraint on the ordering of elements within the RP in P/Y. The unmarked constituent order is Possessive-Noun-Determiner-Adjective-Number (Bowe 1990: 148). In a RP such as (1), *-ku* marks a possessor. The internal structure of a RP may be analysed similarly to a clause in RRG. Pavey (2010: 183-185) says possessive case marked RPs can be analysed as outside the core or as dependents, arguments of the head noun. P/Y lacks the definite alternatives of English ('my book'/'a book of mine') and there is no evidence for a "Reference Phrase Initial Position" grouping possessives and demonstratives (which occur after the head noun), suggesting that in P/Y the possessive-marked noun is a RP core argument. In (1), possessor *watiku* is an argument of possessum *katji*. The constituent projection is shown in Figure 3.

- (1) *Wati-ku katji*  
 Man-P●SS spear  
 'The man's spear' (Goddard 1996: 42-43)  
 have' (wati, katji)

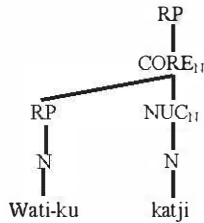


Figure 3: Reference phrase and possession

In (2), *wati wara* ‘tall man’ is marked *-ku* as an argument of the head noun *minyama* ‘woman’; the RP itself is marked ergative.

- (2) *Wati wara-ku minyama ninti pulka-ngku nyuma palya-nu*  
 Man tall-POSS woman clever big-ERG bread.ABS make-PST  
 ‘The tall man’s very clever wife<sub>ACT</sub> made the traditional bread<sub>UND</sub>.’  
 (Bowe 1990: 40)  
 do’ ([have’ (wati, minyama)],  
 [make’ ([have’ (wati, minyama)], nyuma)]) [simplified]

The *-ku* case is used to indicate the owner or rightful user of something, the custodian or caretaker of animates such as younger people or dogs and in reference to social relationships (Goddard 1996: 42-43). Note that with the different word order in (5), *ngayuku* is outside the RP and is predicative.

- (3) *Tjilpi-ku ngura*  
 old.man-POSS place  
 ‘The old man’s place.’  
 have’ (tjilpi, ngura)
- (4) *Kungka-ku katja*  
 Woman-POSS son  
 ‘The woman’s son.’  
 have’ (kungka, katja)
- (5) *Tjau ngayu-ku!*  
 Resin 1SG-POSS  
 ‘It’s my edible resin!’  
 be’ (tjau, [mine’])



For inalienable possession like body parts (but not social relationships), *-ku* is ungrammatical:

- (6) \**Wati-ku mara-ngka*  
 Man-P<sub>POSS</sub> hand-L<sub>LOC</sub>  
 Intended meaning: ‘on the man’s hand’

Instead the case of the ‘owner’ matches that of the body part, as in (7) and (8).

- (7) *Wati-ngka mara-ngka*  
 Man-L<sub>LOC</sub> hand-L<sub>LOC</sub>  
 ‘on the man’s hand’ (Goddard 1993: 17)

- (8) *Wati-ngku mara-ngku papa pu-ngu*  
 Man-ERG hand-ERG dog.ABS hit-PST  
 ‘The man<sub>ACT</sub> hit the dog<sub>UND</sub> with his hand.’ (Blake 1987: 95)  
 do’ (wati, [hit’ (wati, papa)  
 ∧ use’ (wati, [have.as.part’ (wati, maru)]))

#### 4. Purpose and goal

● rather than within the RP, *-ku* may indicate the semantic role of goal in a clause (Bowe 1990: 16). In (9)-(11) with M- and S-intransitive *ananyi* ‘go’, the sole DCA is absolutive. *-ku* here marks argument-adjuncts, because the case marking is predicative and contributes semantically to the LS with ‘purpose’. Figure 4 shows the absolutive and purposive marked arguments in the constituent representation of (11).

- (9) *Wati malu-ku a-nkupai*  
 Man.ABS kangaroo-PURP go-CHAR<sup>2</sup>  
 ‘Men<sub>ACT</sub> would go for kangaroo.’ (Bowe 1990: 16)  
 want’ (wati, [INGR have’ (wati, malu)])  
 ∧ D● (wati, [[do’ (wati, [go’ (wati)])  
 CAUSE [INGR have’ (wati, malu)])])

This can be abbreviated:

<sup>2</sup> *-pai* (glossed CHAR) is a common verb ending indicating an action typically or characteristically done.

[do' (wati, [go' (wati)])] PURP [INGR have' (wati, malu)]

- (10) *Tjitji tjuta kapi uru-ku a-nu.*  
 Kid PL.ABS water hole-PURP go-PST  
 'The kids<sub>ACT</sub> went off for a swim (lit. for a water hole).' (Goddard 1996: 42-43)

[do' (tjitji, [go' (tjitji)])] PURP [INGR be-at' (tjitji, kapi uru)]

- (11) *Kungka ngura-ku a-nu.*  
 Woman.ABS camp-PURP go-PST  
 'The woman<sub>ACT</sub> went home (lit. for camp).'

[do' (kungka, [go' (kungka)])] PURP [INGR be-at' (ngura, kungka)]

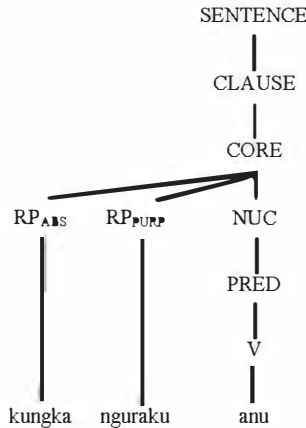


Figure 4: Purposive with movement

Goddard (1996: 42-43) notes *-ku* is also used with more intrinsically goal-directed verbs. Intransitive *patani* 'wait' is exemplified in (12) and (13). The person or thing waited for takes the purposive ending, indicating it is non-DCA. Again, *-ku* here marks an argument-adjunct.

- (12) *Mani-ku nganana pata-ni.*  
 Money-PURP 1PL.NOM wait-PRS  
 'We<sub>ACT</sub>'re waiting for (our) money.' (Goddard 1996: 42-43)  
 [do' (1PL, [wait' (1PL)])] PURP [have' (1PL, mani)]

- (13) *Paluru ngana-mpa paṭa-ni.*  
3SG.NOM IPL-PURP wait-PRS  
'He/she<sub>ACT</sub>'s waiting for us.' (Goddard 1996: 128)  
[do' (3SG, [wait' (3SG)]] PURP [be.with' (3SG, 1PL)]

In (14), nothing is specified as being waited for, *kuta panya* 'the older brother' is absolutive, confirming it as the sole DCA of the S-intransitive verb.

- (14) *Kuta panya rawa putu paṭani-ŋi.*  
Older.brother DEM.ABS long.time in.vain wait-PST.CONT  
'The older brother<sub>ACT</sub> waited in vain for a long time.'

The verb *ngurini* may be translated as 'find' with a direct object and 'search for' with a *-ku* marked object. Example (15) shows the searching has a purpose of finding, with *nyaa* being an argument-adjunct.

- (15) *Nyaa-ku=n nguri-ni?*  
what-PURP=2SG.NOM look-PRS  
'What are you<sub>ACT</sub> looking for?' (Goddard 1996: 102)  
[do' (2SG, [search' (2SG, nyaa)]] PURP [have' (2SG, nyaa)]

## 5. Emotion

The suffix *-ku* is used with verbs of emotion (Bowe 1990: 16, Goddard 1991b), such as loving, hating, fearing and knowing (Myers 1978: 22). For example, *mukuringanyi* 'want, like' takes *-ku* on the thing wanted or liked (Goddard 1996: 80); the stimulus is not optional. This verb is M- and S-intransitive: the sole DCA *kulunypa* 'toddler' is absolutive in (16). By contrast with the purposive situation in the previous section, *-ku* here is not predicative so *ngampu* is marked oblique core (rather than argument-adjunct) as in Figure 5. The verb is semantically divalent, reflected in the semantic representation: logically, there needs to be a stimulus. There is one macrorole: as a state predicate this is the undergoer, *kulunypa*.

- (16) *Kuluny-pa ngampu-ku mukuri-nganyi*  
Toddler-ABS egg-PURP like-PRS  
'The toddler<sub>UND</sub> likes eggs.'  
feel' (kuluny, [desire.for' (ngampu)])

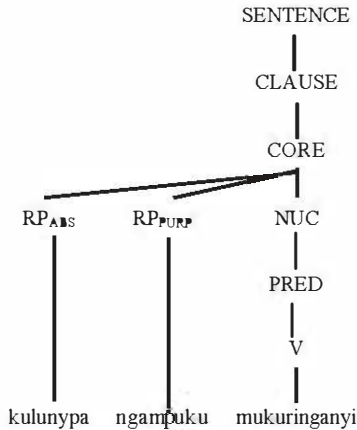


Figure 5: *Mukuringanyi* with stimulus in oblique case *-ku*

Example (17) shows the pronominal purposive form *-mpa*. Again, this is oblique core for the stimulus of emotion.

- (17) *Kurunpa unngu=na nyuntu-mpa mukuri-nganyi.*  
 Heart.ABS deep=1SG.NOM 2SG-PURP love-PRS.  
 'I<sub>UND</sub> love you in my heart.' (Goddard 1996: 54)

This situation superficially resembles the antipassive (Blake 1987: 57-58), which is found in several Australian languages. In an antipassive construction the  $A_T$  becomes S and the  $U_T$  has an oblique case as a complement, ceasing to be a macrorole. This produces a sentence with lower semantic transitivity (Pavey 2010: 160-161). However, there is no S-transitive alternation in P/Y, and there is no varying semantic transitivity, so we do not term this an antipassive.

Morphologically, *mukuringanyi* has the appearance of being an inchoative verb: such M- and S-intransitive verbs are derived from state predicates (adjectives or nouns) in P/Y by suffixing *-ri* (Bowe 1990: 26)<sup>3</sup>. Inchoative verbs are normally accomplishments in RRG terms; a process and achievement with change of state. This is consistent: the 'target' or object of an emotion or attitude is marked with *-ku* with inchoative verbs

<sup>3</sup> *Muku(lyu)* 'love' is the root. Glass & Hackett (2003: 172) show how this emotion verb may have arisen in the closely related Western Desert dialect Ngaanyatjarra. The reduplicated root *muku-muku* means 'gentle, kind'. The verb *mukuringku* derives from this, literally 'become kind', with the Ngaanyatjarra inchoative suffix *-rri*. This is then lexicalised as 'like, want'.

generally, such as in (18) *pikaringu* ‘become angry’ and (19) *kuntaringu* ‘become shy’. Examples (18)-(21) are again oblique core, marking internal experience. As accomplishments, these are episodic: an emotion arises in a situation.

- (18) *Wati paluru katja-ku pika-ri-ngu.*  
 Man 3SG.NOM son-PURP angry-INCH-PST  
 ‘The man<sub>UND</sub> felt/got angry toward his son.’  
**BECOME feel’** (wati, [**angry.towards’** (katja)])
- (19) *Ngayu-lu kungka-ku kunta-ri-ngu.*  
 1SG-NOM woman-PURP shy-INCH-PST  
 ‘I<sub>UND</sub> got shy toward the woman.’ (Goddard 1996: 42-43)  
**BECOME feel’** (1SG, [**shy.towards’** (kungka)])

Stative adjectival predicates use *-ku* too. As non-verbs, there is no marking for aspect and tense and no sense of their being accomplishments. In (20) and (21), an object is the target for an emotional state. As an unexpressed argument in P/Y is interpreted as 3<sup>rd</sup> person singular by default, and we include this in the assignment of macroroles. The only available argument for macrorole assignment is 3SG; *nganytju* is marked *-ku* as NMR and non-DCA in (20).

- (20) *Nganytju-ku ngulu*  
 Horse-PURP afraid  
 ‘(He/she<sub>UND</sub> is) afraid of horses.’ (Goddard 1993: 18)  
**feel’** (3SG, [**frightened.of’** (nganytju)])
- (21) *Mumu ya-nu=lta wati-ku ngulli*  
 And.SS go-PST=INTP man-PURP wary  
 ‘And (she<sub>ACT</sub>) left, (she<sub>UND</sub> was) wary of the man.’ (Goddard 1996: 42-43)

The relation between experiencer and stimulus is expressed differently cross-linguistically. In English ‘I like’ is an emotion from an experiencer towards a target. P/Y *mukuringanyi* ‘like’ might be translated as ‘be/feel gentle to’. Greek turns it the other way around: there is an emotion happening to an experiencer in dative case, for example in (22) with *mou aresei* ‘is pleasing to me’. This pattern is common in European languages (Van Valin 2018).

- (22) *Afti ti stigmi mou ares-ei para poli*  
 DEMDET moment me.DAT please-3SG.PRS very much  
*i Ellada*  
 DET Greece  
 ‘Right now I like Greece very much.’ (Archakis 2014)

## 6. Ability or familiarity

The *-ku* case indicates the object of ability or familiarity with predicates of cognition. (23) and (24) show the state predicates *ninti* ‘able’ and *ngurpa* ‘not knowing’. The object is oblique core; the predicate is S-intransitive stative with an undergoer.

- (23) *Wangka-ku ninti*  
 Word-PURP able  
 ‘(He/she)<sub>UND</sub> is able to talk.’
- (24) *Mutuka-ku ngurpa*  
 Car-PURP not.knowing  
 ‘(He/she)<sub>UND</sub> does not know about cars.’ (Goddard 1996: 42-43)

In (25), *ngurpa* is used on its own as a state predicate. The object of familiarity is presumed from previous dialogue.

- (25) *Wiya, ngayu-lu ngurpa*  
 NEG 1SG-NOM not.knowing  
 ‘No, I<sub>UND</sub> don’t know (about it).’ (Goddard 1996: 103)

## 7. Beneficiary

The suffix *-ku* also marks the semantic role of beneficiary. In (26), *ngayu-lu* ‘I’ is nominative with the active M- and S-transitive verb *tjumu* ‘make’; *yuu* ‘windbreak’ is in absolutive case as the U<sub>T</sub> direct core argument. The beneficiary *tjamu* ‘grandfather’ is optional and marked *-ku*. This is an argument-adjunct, *-ku* is predicative, contributing to the LS as PURP.

- (26) *Ngayu-lu yuu tjumu, tjamu-ku*  
 1SG-NOM windbreak.ABS make-PST grandfather-PURP  
 ‘I<sub>ACT</sub> made a windbreak<sub>UND</sub> for grandfather<sub>NMR</sub>.’ (Goddard 1996: 42-43)  
 [do’ (1SG, [make’ (1SG, yuu)])]  
 PURP [BECOME have’ (tjamu, yuu)]

## 8. Trivalent verbs

In other cases, *-ku* is not used where it might be expected in P/Y. Blake (1987: 44) has described *-ku* as the dative, with a “beneficiary”. However, where there is a transfer of an object in P/Y, the semantically trivalent verb (*y)unganyi* ‘give’ uses the strategy in (27). Both objects are in absolutive case (Bowe 1990: 24), the receiver coming first. This is a S-ditransitive verb with three DCAs: the noun nearest the verb is undergoer. So *-ku* is not used in this prototypical dative marking construction.<sup>4</sup> The constituent representation is in Figure 6.

- (27) *Minyma-ngku tjitji mai u-ngu*  
 Woman-ERG child.ABS bread.ABS give-PST  
 ‘The woman<sub>ACT</sub> gave the child<sub>NMR</sub> some bread<sub>UND</sub>.’  
 [do’ (minyma, Ø)] CAUSE [BECOME have’ (tjitji, mai)]

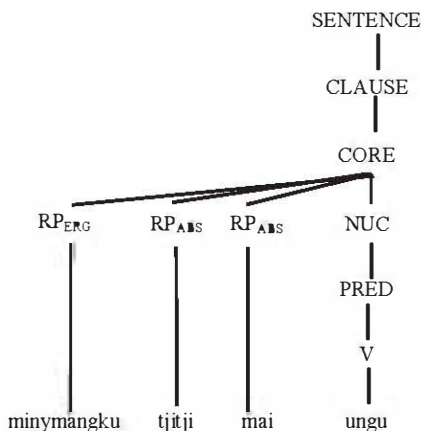


Figure 6: Semantically trivalent structure, with three DCAs

The P/Y situation contrasts with another potential three-argument verb, *wangkanyi* ‘to tell/speak’. Bowe (1990: 25) shows that there are two structures that can be used with it. Both the addressee and the theme can be in the absolutive case, with the obligatory direct object or theme being closest to the verb as in (28):

<sup>4</sup> The neighbouring related dialect Ngaanyatjarra has *-ku* on the recipient of verbs of giving (Glass & Hackett 1979: 123-124). This is further evidence that the theme rather than recipient is undergoer.

- (28) *Minyma-ngku tjitji tjukur-pa wangka-ngu*  
 Woman-ERG child.ABS story-ABS tell-PST  
 ‘The woman<sub>ACT</sub> told the child a story<sub>UND</sub>.’  
 [**do**’ (minyma, [**express**.( $\alpha$ ).**to**.( $\beta$ ).**in.language**.( $\gamma$ )’ (minyma, tjitji)]] CAUSE [**BECOME know**’ (tjitji, tjukurpa)],  
 where  $\alpha$  = *tjukurpa*,  $\beta$  = *tjitji*

Alternatively, the addressee can get the locative case as in (29), making it a non-DCA, in a syntactically (but not semantically) valence reducing operation.

- (29) *Minyma-ngku tjitji-ngka tjukur-pa wangka-ngu*  
 Woman-ERG child-LOC story-ABS tell-PST  
 ‘The woman<sub>ACT</sub> told the child a story<sub>UND</sub>.’

## 9. Purposive different-subject clauses

We can extend this “purposive” notion of *-ku* to certain dependent clauses. P/Y has a dependent clause form, known as “purposive different subject” in Bowe (1990: 74-75) or just “purpose” in Eckert & Hudson (1988: 307-309). It is formed by the nominalised verb (root with *-nytja*) and *-ku*. In (30), the person makes the artefact, for the purpose of someone else giving her money (Goddard 1993: 33).

- (30) *Punu palya-ni mani u-ngkunytja-ku*  
 Artefact.ABS make-PRS money.ABS give-NMLZ-PURP  
 ‘(She<sub>ACT</sub>’s) making an artefact<sub>UND</sub>, so (they) will give (her) money.’  
 [**do**’ (3SG, [**make**’ (3SG, punu)]]  
 PURP [[**do**’ (3PL, O)] CAUSE [**BECOME have**’ (3SG, mani)]]

Bowe (1990: 71) suggests that these “dependent clauses” are subordinate to the main clause as they are embedded rather than adjoined; as evidence for this proposal, we see in (31) that the form can be moved within the main clause. Wilkins (1988) states that purposive clauses in Mparntwe Arrernte (Pama-Nyungan, Northern Territory) are embedded rather than adjoined; and sub-clauses can be arguments of the core. From our analysis previously in this chapter with argument-adjunct *-ku*, such “sub-clauses” behave like non-direct core arguments, indicating the purpose of an action (here Trevor spoke with the purpose of Mary going). This is consistent with *-ku* being a nominal non-DCA case form; in this case on a nominalised verb. In our RRG analysis, this is two coordinated cores within a clause, sharing an argument *Mary*. *Marynya* is the U<sub>T</sub> argument of S-transitive *watjanu* and the S argument of S-intransitive *ankunytjaku*.



- (31) *Trevor-lu Mary-nya a-nkuntja-ku watja-nu*  
 Trevor-ERG Mary-ABS go-NMLZ-PURP say-PST  
 ‘Trevor<sub>ACT</sub> told Mary<sub>UND</sub> to go.’  
 [**do**’ (Trevor, [**express**.( $\alpha$ ).to.( $\beta$ ).in.language.( $\gamma$ )’ (Trevor, Mary)))] CAUSE [**BECOME aware.of**’ (O, Mary)],  
 where  $\beta$  = 1SG  
 PURP [**do**’ (Mary, [**move.away.from.ref.point**’ (Mary)))] & INGR **be-at**’ (O, Mary)]

The form is “different subject” so the listener is interpreted as someone other than Trevor, and, in context, Mary. The linking of (31) is shown in Figure 7: the solid line represents obligatory control whereby for different subject constructions the undergoer of the first core verb is the actor of the second (Van Valin 2005: 249).

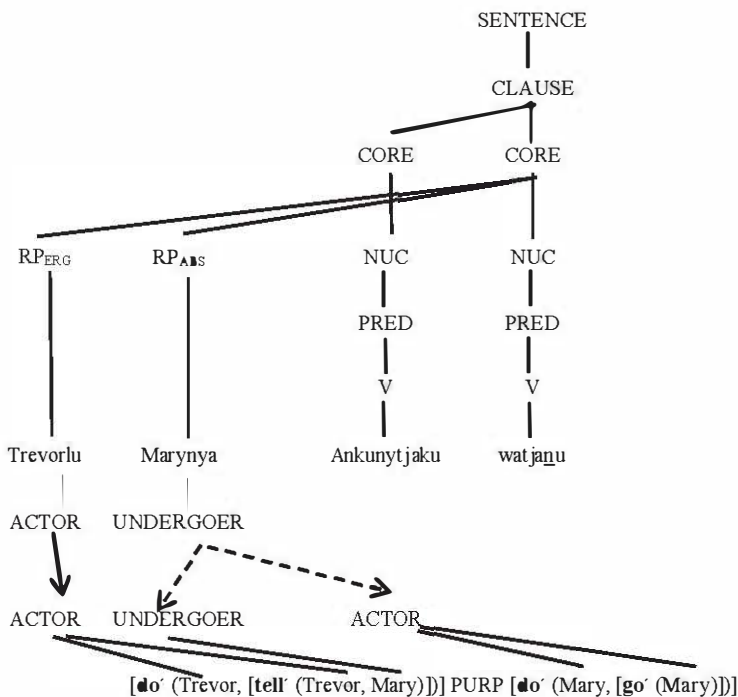


Figure 7: Purposive different-subject core coordination

We saw M- and S-intransitive verb *paṭani* ‘wait’ in (12) and (13) having a purpose marked object for the thing waited for. In (32), both the pronominal *nyuntumpa* and dependent *nyakunytjaku* are marked with purpose: the first an argument-adjunct, the second a coordinated core with shared arguments.

- (32) *Paluru kuwari paṭa-ni nyuntu-mpa, .*  
 3SG.NOM now wait-PRS 2SG-PURP,  
*a-nkula nyaku-nytja-ku*  
 go-SER see-NMLZ-PURP  
 ‘He’s waiting for you now, for (you) to go see (him).’ (Goddard 1996: 128)  
 [[[do’ (3SG, [wait’ (3SG))]]] PURP [be.with’ (3SG, 2SG)]]  
 PURP [[do’ (2SG, [go’ (2SG)]]] PURP [see’ (2SG, 3SG)]]]

## 10. Periphery

Up to now we have seen *-ku* mark oblique arguments and argument-adjuncts, which are in the core. Can it also mark peripheral elements? Peripheral elements take the LS as an argument, typically locating the scene in time or location and modifying the core (Van Valin 2005: 19). In P/Y, locative *-ngka* is used for these rather than *-ku* (Goddard 1996).

Peripheral elements may also mark an activity being done in place of someone (VV p.c.). The question lies in separating out peripheral elements from non-DCA recipient beneficiaries. Examples (33) and (34) show activities being carried out for someone. The example in (34) has the pronoun form *-mpa*. As we saw in (26), something can be made (S-transitive verb) for someone (beneficiary, argument-adjunct).

- (33) *Nyuutji ngayu-ku ura-la!*  
 News 1SG-PURP gather-IMP  
 ‘Find out the ‘news’ for me.’ (Goddard 1996: 116)
- (34) *Munu=na nyuntu-mpa inma kantu-ni.*  
 And.SS=1SG.NOM 2SG-PURP dance dance-PRS  
 ‘I’m doing this dance for you.’ (Goddard 1996: 141)

## 11. Discussion

In this chapter, we have discussed the different situations where the suffix *-ku* is used as a case marker in Pitjantjatjara and Yankunytjatjara. These situations include non-body part possession, target of emotion, ability, purposive goal and beneficiary. Possession is of a different nature to the others as it is phrase-internal.

What do the non-possessive uses have in common? We found two main patterns, with arguments being oblique core arguments or argument-adjuncts. Oblique *-ku* is a marker of emotion or knowledge. The LS is based on a state predicate, without *do'*. Verbs of emotion are low on Hopper & Thomson's (1980) transitivity scale: the stimulus is unaffected and there is low volition from the experiencer. The verb may be atelic without a change of state; or as in (18) and (19), the verb is based on a static predicate but is an accomplishment, developing over time.

Argument-adjunct *-ku* is used to show purpose with prototypical action verbs such as *ananyi* 'come, go, travel'. In this case the *-ku* marked nominal shows the reason for travel or the location to which it is directed. The LS in such cases has D● in the long form, or PURP in the abbreviated form, with *do'*. The suffix can also mark a beneficiary with a S-transitive action verb such as *tjunanyi* 'make, put'. In other situations the verb is not an action one, but there is volition such as waiting: there is still a purpose in mind.

What bearing does this have on the syntactic valence of the predicate? RRG distinguishes syntactic and semantic valence. While a semantic argument may be syntactically expressed outside the core (such as an actor in the periphery with a passive, or in the pre-core slot), the reverse does not apply: arguments in the core are reflected in the LS. However syntactic transitivity is determined by the number of DCAs, not total number of core arguments. If the predicate is S-intransitive, this is confirmed by absolutive case marking on the sole DCA, as in the case of *ngurpa* or *ananyi*. This is syntactic evidence for *-ku* being an oblique case marker or argument-adjunct rather than DCA.

This commonality between stimulus and beneficiary has been remarked upon by other authors. Blake (1987: 44) describes *-ku* on the stimulus as the "dative", which ties in to the sense of "beneficiary". Dixon & Aikhenvald (2000: 3) posit "extended intransitive" clauses with S and E (extension to core), typically used for seeing, hearing, liking and wanting. The E argument frequently receives dative case and belongs with the similar "extended transitive" or ditransitive where the recipient or beneficiary is marked dative. "Extension to core" is in line with the RRG analysis: rather than putting this case in the syntactic periphery: the *-ku* marked ar-

guments are in the core but not DCAs. “Extended intransitives” are oblique and “extended transitives” are argument-adjuncts in RRG terms. In P/Y there is however a distinction between recipient and beneficiary “datives”; recipients in P/Y are absolutive marked DCAs. Beneficiaries for whom a transitive action was performed are marked with *-ku*, as argument-adjuncts. Douglas (1958: 94) describes *-ku* as marking the indirect object or the reason for the action in Western Desert: this is only partially true for the P/Y dialects.

Syntactic transitivity is the number of DCAs, macrorole transitivity the number of macroroles. So *-ku* marked nouns are not part of this determination. However, the argument-adjunct use of *-ku* has certain similarities to applicative constructions which bring in another argument to the core, increasing valence. Benefactives are a subset of applicative: in (26) a beneficiary is introduced and marked *-ku*. Applicatives may vary undergoer assignment (Van Valin & LaPolla 1997: 337-338) but here the original undergoer remains as DCA. The net result is that the semantic valence of the predicate is raised but not the M-or S-transitivity.

We have extended the analysis to include the purposive different subject dependent clause, which has a nominalised verb in its core; this nominalised verb takes *-ku* to indicate the (different subject) purpose of the action depicted in the main verb. We analyse these as coordinated cores within one clause. This is in line with the purpose marked RPs discussed earlier in the chapter.

We summarise our findings concerning *-ku* in Table 4:

Table 4: Summary of uses of *-ku*

Type	Found	Notes
Adnominal	Yes	Alienable and kin possession. Unlike the other uses, it is RP-internal
Direct core dative	No	Recipient is absolutive and not marked with <i>-ku</i>
Oblique core argument	Yes	Target or static goal of internal states, such as emotion or attitude
Argument-adjunct	Yes	Goal of action, such as beneficiary (but not recipient)
Core coordination	Yes	Purposive different subject constructions, may be argument-like
Adjuncts	Marginal	Purposive, goal-directed constructions with <i>-ku</i> can usually be analysed as having a beneficiary

## Abbreviations

ABL ‘ablative’, ABS ‘absolutive’, CHAR ‘characteristic’, CONT ‘continuous’, DCA ‘direct core argument’, DEM ‘demonstrative’, ERG ‘ergative’, IMP ‘imperative’, INCH ‘inchoative’, INTP ‘interest particle’, LOC ‘locative’, NEG ‘negative’, NMLZ ‘nominaliser’, NMR ‘non-macrorole’, NOM ‘nominative’, NP ‘noun phrase’, OBL ‘oblique’, PL ‘plural’, POSS ‘possessive’, PRS ‘present’, PST ‘past’, PURP ‘purposive’, RP ‘reference phrase’, SER ‘serial participle’, SG ‘singular’.

## References

- Archakis, Argiris. 2014. Immigrant voices in students’ essay texts: Between assimilation and pride. *Discourse & Society*, 25(3): 297–314.
- Blake, Barry. 1987. *Australian Aboriginal grammar*. London: Croom Helm.
- Bowe, Heather J. 1990. *Categories, constituents and constituent order in Pitjantjatjara*. London: Routledge.
- Chang, Jung-hsing. 2007. Linking semantics and syntax in Mandarin serial verbs: A Role and Reference grammar account. *Language and Linguistics*, 8(1): 235–266.
- Dixon, R. M. W. 2011. *The languages of Australia*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. & Aikhenvald, Alexandra. Y. 2000. Introduction. In Dixon, R. M. W. & Aikhenvald, Alexandra. Y. (eds.), *Changing valency: Case studies in transitivity*, 1–29. Cambridge: Cambridge University Press.
- Douglas, W. H. 1958. *An introduction to the Western Desert language of Australia*. Sydney: University of Sydney.
- Eckert, Paul & Hudson, Joyce. 1988. *Wangka Wiru: A handbook for the Pitjantjatjara language learner*. Underdale: ASTEC.
- Glass, Ameer & Hackett, Dorothy. 1979. *Ngaanyatjarra texts*. Canberra: Australian Institute of Aboriginal Studies.
- Glass, Ameer & Hackett, Dorothy. 2003. *Ngaanyatjarra & Ngaatjatjarra to English dictionary*. Alice Springs: IAD Press.
- Goddard, Cliff. 1983. *A semantically-orientated grammar of the Yankunytjatjara dialect of the Western Desert language*. Canberra: Australian National University. (Doctoral thesis.)
- Goddard, Cliff. 1991a. Testing the translatability of semantic primitives into an Australian Aboriginal language. *Anthropological Linguistics*, 33(1): 31–56.

- Goddard, Cliff. 1991b. Anger in the Western Desert: A case study in the cross-cultural semantics of emotion. *Man, New Series*, 26(2): 265–279.
- Goddard, Cliff. 1993. *Pitjantjatjara/Yankunytjatjara learner's guide*. Alice Springs: IAD Press.
- Goddard, Cliff. 1996. *Pitjantjatjara/Yankunytjatjara to English dictionary* (2nd Edition). Alice Springs: IAD Press.
- Goddard, Cliff & Harkins, Jean. 2002. Posture, location, existence, and states in Central Australia. In Newman, John (ed.), *The linguistics of sitting, standing and lying*, 213–238. Amsterdam: John Benjamins.
- Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language*, 56(2): 251–299.
- Kim, Jong-bok & Sells, Peter. 2008. *English syntax: An introduction*. Stanford: CSLI Publications.
- King, Deborah. 2010. *Voice and valence-altering operations in Falam Chin: A Role and Reference Grammar approach*. Arlington: The University of Texas. (Doctoral dissertation.)
- Myers, Deborah. 1978. *A simple explanation of Pitjantjatjara*. Adelaide: Education Department of South Australia.
- Nolan, Brian. 2012. *The structure of Modern Irish*. Sheffield: Equinox.
- Pavey, Emma L. 2010. *The structure of language: an introduction to grammatical analysis*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2001. *An introduction to syntax*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2007. The Role and Reference Grammar analysis of three-place predicates. *Suvremena lingvistika*, 63, 31–63.
- Van Valin, Robert D., Jr. 2018. Dative case and oblique subjects. In Barðdal, Jóhanna, Pat-El, Na'ama & Carey, Stephen Mark (eds.), *Non-canonically case-marked subjects. The Reykjavik Eyjafjallajökull papers*, 115–131. Amsterdam: John Benjamins.
- Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
- Wilkins, David P. 1988. Switch-reference in Mparntwe Arremte (Aranda): Form, function, and problems of identity. In Austin, Peter (ed.), *Complex sentence constructions in Australian languages*, 141–176. Amsterdam: John Benjamins.



# A NEO-JAKOBSONIAN ACCOUNT OF DEFAULT OBLIQUE CASES: INSTRUMENTAL VS. DATIVE

WATARU NAKAMURA  
TOHOKU UNIVERSITY

## Abstract

The aim of this paper is twofold: to recast the RRG theory of case assignment (Van Valin & LaPolla 1997; Van Valin 2005) in terms of Jakobson's (1936/1984) featural definitions of cases and to explore a few empirical and theoretical consequences of their integration. Adopting the OT formulation of case assignment rules in RRG (Nakamura 1999), this paper redefines dative and instrumental case in such a way that dative-marked nouns (i.e. non-macrorole core arguments) constitute a subset of those nouns that are qualified to receive instrumental case (i.e. all nouns other than macrorole arguments) and spells out how to derive the wide range of semantic interpretations of instrumental case (e.g. instrument, implement, locatum, unit, domain restriction, path, manner, duration) that have been analyzed in terms of polysemy networks (Janda 1993; Narrog & Ito 2007; Malchukov & Narrog 2009; Narrog 2009) from the interaction of its abstract invariant meaning with the clausal context.

## Keywords

Case, dative, instrumental, Qualia Structure, Optimality Theory, Russian

## 1. Introduction

The issue of oblique case assignment has played a prominent role in the RRG linking theory since Van Valin (1991), who defined dative as the default case for non-macrorole core arguments. This macrorole-based definition of dative case captures a wide range of its uses crosslinguistically (e.g. Narasimhan 1998; Nakamura 2000, 2008). Foley & Van Valin (1984:



81-87) is the first RRG account of the English instrumental preposition *with*, but Foley & Van Valin and their successors (Jolly 1991, 1993; Van Valin & LaPolla 1997: 377-382; Van Valin 2005: 109-115) are not successful in deriving the whole range of contextual functions of *with* (or rather its prepositional phrases as a whole) (e.g. instrument, co-actor, locatum, path, manner) from its unified meaning. The fact that the instrumental case in Russian marks a wide range of clausal adjuncts suggests that it functions as a default marker of some kind, but this, in turn, raises the question of how to describe its relation with dative case, which has been analyzed as the least marked oblique case morpheme in RRG (Van Valin 1991; cf. Silverstein 1980/1993) <sup>1</sup>

Jakobson (1936/1984) stands in contrast to Van Valin & LaPolla (1997) and Van Valin (2005), in that he defines instrumental as being the least marked oblique case within his case theory. Jakobson analyzes the Russian case system in terms of three privative semantic features, [peripheral], [directional], and [quantified] (as shown in Table 1 below) and defines instrumental case as having [peripheral] alone, while defining dative case as having [directional] and [peripheral]:

Table 1: Jakobson's (1936/1984) definitions of cases in Russian<sup>2</sup>

	Peripheral	Directional	Quantified
Nominative			
Accusative		+	
Genitive			+
Dative	+	+	
Instrumental	+		
Locative	+		+

<sup>1</sup> The term *oblique* used in this paper needs clarification. The term has been treated as ambiguous between all non-nominative cases and all cases other than syntactic cases such as nominative/absolute, accusative, and ergative. What is innovative about Van Valin (1991) (and his later work) is that he includes dative under the heading of *direct* (as opposed to *oblique*) cases. The present paper stays neutral with respect to whether or not dative should be classified as a direct or oblique case and adopts the traditional term *oblique* as a cover term to refer to both dative and instrumental for lack of any better term.

<sup>2</sup> The second genitive ([directional], [quantified]) and second locative ([peripheral], [directional], [quantified]) are omitted in Table 1. Their treatment is beyond the scope of this paper, but I would refer the reader to Brown (2007) and Corbett (2008) for how to analyze the second locative case.

The featural definitions of dative and instrumental case, highlighted by shading in Table 1, view dative as being more marked than instrumental, since dative involves two of the three features ([peripheral] and [directional]), while instrumental involves only one ([peripheral]).

These two accounts of oblique case assignment do not seem to be compatible at first blush: Jakobson (1936/1984) defines instrumental as being less marked than dative, while Van Valin (1991, 2005) analyzes dative as the least marked oblique case (cf. Silverstein 1980/1993). Van Valin (2005) views the instrumental preposition *with* as non-predicative and distinguishes it from predicative prepositions (i.e. those prepositions that come with their own logical structure [LS] and contribute substantive semantic information to the clause in which they occur), but Van Valin doesn't assign a default status to the preposition *with*. Against this backdrop, I will propose to integrate these two accounts into a coherent framework that defines instrumental and dative as default oblique cases with varying degrees of specificity and will derive the diverse uses of the instrumental case in Russian from its underdetermined invariant meaning and contextual information.

The rest of this paper is organized as follows. Section 2 surveys the major uses of instrumental case in Russian. Section 3 critically reviews the RRG account of the instrumental preposition *with* (Van Valin & LaPolla 1997; Van Valin 2005). Section 4 proposes an underdetermined definition of instrumental case that accords well with Jakobson's (1936/1984) featural classification of the cases in Russian. Section 5 demonstrates how the highly abstract meaning of instrumental case is enriched through its interaction with the clausal context. Section 6 extends the present account to the two instrumental prepositions in English. Section 7 concludes the paper.

## 2. Instrumental case in Russian

(1a-n) and (2a-e) are the core Russian data analyzed in this paper. (1a-n) are non-predication uses of the instrumental case, while (2a-e) illustrate its uses for representing primary and secondary predication. The instrumental-marked nouns and adjectives in (1a-n) and (2a-e) and their English glosses are highlighted in bold for clarification.<sup>3</sup>

---

<sup>3</sup> Abbreviations: ACC 'accusative', AUH 'Actor-Undergoer Hierarchy', CONST 'constitutive', DAT 'dative', DCR 'dative case rule', GEN 'genitive', GL 'Generative Lexicon', ICR 'instrumental case rule', INSTR 'instrumental', LOC 'locative', LS 'logical structure', MR 'macrorole', NOM 'nominative', OPT 'Optimality Theory',

(1) Instrument

- a. *Ivan rezal xleb nožom.*  
 Ivan.NOM cut bread.ACC knife.INSTR  
 'Ivan cut bread **with a knife**.' (Janda 1993)

Implement

- b. *On el ložkoj ikru.*  
 he.NOM ate spoon.INSTR caviar.ACC  
 'He ate caviar **with a spoon**.' (Demjjanow & Strigin 2003)

Transportation

- c. *On exal poezdom.*  
 he.NOM drove train.INSTR  
 'He was going **by train**.' (Demjjanow & Strigin 2003)

Unit

- d. *On pil vino litrami.*  
 he.NOM drank wine.ACC liters.INSTR  
 'He drank wine **by the liter**.' (Demjjanow & Strigin 2001)

Quantity

- e. *Fašisty uničtožali ljudej millionami.*  
 fascists.NOM destroyed people.ACC millions.INSTR  
 'The fascists killed people **by the millions**.' (Janda 1993)

Manner

- f. *Monax dolžen svjazat' usta svoi molčaniem.*  
 monk.NOM must tie.up lips own.ACC silence.INSTR  
 'A monk must seal his lips **with silence**.' (Janda 1993)

Point in time

- g. *Oni vstretilis' osen'ju v Pariže.*  
 they.NOM met autumn.INSTR in Paris.LOC  
 'They met **in autumn**, in Paris.' (Wierzbicka 1980)

Duration

- h. *On (celymi) časami vozilsja s radiopriemnikom.*  
 he.NOM (whole.INSTR) hours.INSTR fiddled  
 with radio.INSTR  
 'For hours on end he would fiddle with the radio.' (Wierzbicka 1980)

Path

- i. *●ni šli pjat' kilometrov lesom.*  
 they.NOM went five.ACC kilometers.GEN forest.INSTR  
 ‘They walked 5 kilometers **through a forest.**’ (Sofia Tumanova  
 personal communication)

Domain restriction

- j. *Lošad' sosedu lučše moej*  
 horse.NOM neighbor.GEN better mine.GEN  
*i siloj i krasotoj.*  
 and strength.INSTR and beauty.INSTR  
 ‘The neighbor’s horse is better than mine both **in strength and in appearance.**’ (Janda 1993)

Part-whole relation

- k. *Petr kivnul golovoj.*  
 Peter.NOM nodded head.INSTR  
 ‘Peter **noded his head.**’ (lit. ‘Peter **noded with the head.**’)  
 (Wierzbicka 1980)

Passive actor

- l. *●kno bylo razbito mal'čikami.*  
 window.NOM was broken boys.INSTR  
 ‘The window was broken **by the boys.**’ (Wierzbicka 1980)

Antipassive undergoer

- m. *Ivan švyrjal(-sja) kamnjami.*  
 Ivan.NOM was.throwing(-REFL) stones.INSTR  
 ‘Ivan was throwing **stones.**’ (Wierzbicka 1980)

Locatum

- n. *●ni gruzili baržu drovami.*  
 they.NOM loaded barge.ACC firewood.INSTR  
 ‘They **loaded the barge with firewood.**’ (Kilby 1977)

(2) Primary predication

- a. *●n byl šaxterom(šaxter).*  
 he.NOM was miner.INSTR(NOM)  
 ‘He was **a miner.**’ (Wierzbicka 1980)
- b. *●n byl veselym(veselyj).*  
 he.NOM was cheerful.INSTR(NOM)  
 ‘He was **cheerful.**’ (Israeli 2007)

Secondary predication

- c. *My tancevali golymi.*  
 we.NOM danced nude.INSTR  
 ‘We **danced nude.**’ (Bailyn 2012)

- d. *On*            *rabotaet*            *vracom.*  
 he.NOM        works                doctor.INSTR  
 ‘He works **as a doctor.**’ (Bailyn 2012)
- e. *Anja*         *poet*                 *solov'em.*  
 Anja.NOM     sings                 nightingale.INSTR  
 ‘Anja sings **like a nightingale.**’ (Janda 1993)

Six remarks are in order about (1) and (2). First, (1a) is distinguished from (1b), in that (1a) contains a noun for the instrument that participates in a causal chain starting from the effector and ends at the patient, while (1b) does not. Second, the instrumental-marked nouns in (1a-b) and (11-n) are semantic arguments of verbs, while all the other instrumental-marked nouns in (1) are syntactic adjuncts.<sup>4</sup> Third, (11-m) involve passivization and antipassivization, respectively. Both the actor and undergoer functioning as a syntactic adjunct in (11-m) receive their semantic interpretation from the verb. Another important point to note regarding (1m) is that it exhibits an alternation between accusative and instrumental on the undergoer argument of the verb, as shown in (3a-b). This case alternation is allowed by some transitive verbs including verbs of throwing, as illustrated in (1m). (3c) shows that when Ivan wanted to affect Peter rather than the stones, the undergoer argument (*kamni* ‘stones’) has to receive instrumental case:

- (3)a. *Ivan*            *švyrjal(-sja)*            *kamnjami.*  
 Ivan.NOM    was.throwing(-REFL)    stones.INSTR
- b. *Ivan*            *švyrjal*                 *kamni.*  
 Ivan.NOM    was.throwing                stones.ACC  
 ‘Ivan was throwing stones.’ (Wierzbicka 1980)
- c. *Ivan*            *švyrjal*                 *kamnjami/\*kamni*    *v Petra.*  
 Ivan.NOM    was.throwing                stones.INSTR/ACC    atPeter.ACC  
 ‘Ivan was throwing stones at Peter.’ (Wierzbicka 1980)

<sup>4</sup> Koenig et al. (2003, 2007) show that instrument and some implement nouns in English are semantic arguments of verbs in terms of semantic obligatoriness and semantic specificity, but they leave it unanswered whether these semantic arguments are core arguments or not. Rissman (2010) compares English counterparts to (1a-b) and (1n) and states that only the locatum noun behaves as a syntactic argument (cf. Ono 1992) (contrary to Van Valin (2012), who argues that all of the three instrumental-marked nouns are core arguments). I assume that the same argument holds for (1a,b,n).

The instrumental marking of the undergoer in (3a) reflects its peripheral and subordinate role in the action. Specifically, (3a) describes a situation in which Ivan was not interested in the stones he was throwing: Ivan may have a target other than the stones. Another point to note with respect to (1m) is that transitive verbs as illustrated in (1m) are often accompanied by the reflexive suffix *-sja*. If we may follow Wierzbicka (1980) in analyzing *-sja* as having an antipassive function, then it follows that these verbs optionally involve antipassivization (which makes the aspect of the sentence atelic and causes the undergoer to be realized as an adjunct). Fourth, I follow Schwartz (1993) in assuming that the instrumental-marked noun and adjective in (2a-b), which occupy the second argument slot of the copula verb (see the upper part of (5)), have become part of the nucleus as a consequence of *argument incorporation* in (4). This operation deprives the theme argument of its status as an LS argument and makes it part of the nucleus, which is not eligible for macrorole assignment. (5) shows how the macrorole assignment proceeds in (2a):<sup>5</sup>

(4) Argument incorporation (Schwartz 1993: 447):

**be' + attribute/identificand** → **predicate (nucleus)**

(5) LS: [be' (he, miner)] [be' (he, [miner'])]

Locative Theme → Locative (Theme)

MR:

Undergoer

The locative argument is associated with undergoer in (5), since it is the only LS argument available for macrorole assignment. Fifth, we may extend the argument incorporation account of (2a-b) to (2c-e), in which the depictive noun or adjective serves as a predicate about the subject argument. Finally, I build on (4) and propose that (2c-e) involve a non-subordinate core juncture (Schultze-Berndt & Himmelman 2004), in which the depictive noun or adjective forms a core that shares its PSA with the one formed by the matrix verb.<sup>6</sup>

(6) is a summary of the above overview:

<sup>5</sup> Schwartz (1993) argues that the crucial piece of evidence for the operation of *argument incorporation* in Russian comes from the fact that the genitive-of-negation applies only to undergoer arguments that involve no violation of the AUH and that it doesn't apply to subjects of copular verbs despite their undergoer status.

<sup>6</sup> See Ullrich (2018) for an analogous proposal.

(6) Classification of the instrumental-marked nouns and adjectives in (1) and (2)

- |                        |   |
|------------------------|---|
| a. Periphery (adjunct) | Semantic argument: (1a), (1b), (1l), (1m)<br>Non-argument: (1c-k) |
| b. Core argument       | Semantic argument: (1n)   |
| c. Part of the nucleus | (2a-e) ((2c-e) involve a non-subordinate core juncture)           |

No matter how the instrumental case/adposition may be defined, we will have to define it in such a way as to accommodate the semantic and syntactic diversity of instrumental-marked nouns (summarized in (6)).

Before proposing new definitions of the dative and instrumental case in Russian in Section 4, let us take a brief look at Van Valin & LaPolla (1997) and Van Valin (2005) as representative of the previous RRG account of the instrumental case/adposition assignment.

### 3. The previous RRG account

RRG distinguishes three classes of adpositional phrases, depending on whether they mark core arguments or clausal adjuncts and whether they are predicative or non-predicative. (7) represents the three-way classification of adpositions (Van Valin & LaPolla 1997; Van Valin 2005):

(7) Three-way classification of adpositions in RRG

- |   |  |
|---|--|
| a. Adjunct adpositions (predicative adpositions that mark clausal adjuncts):            |  |
|   | e.g. John was running <i>in the park</i> .<br>LS: <b>be-in'</b> (park, [ <b>do'</b> (John, [ <b>run'</b> (John)])])    |
| b. Argument-marking adpositions (non-predicative adpositions that mark core arguments): |  |
|   | e.g. John gave the magazine <i>to Tom</i> .<br>LS: [ <b>do'</b> (John, O)] CAUSE [INGR <b>have'</b> (Tom, magazine)]   |
| c. Argument-adjunct adpositions (predicative adpositions that mark core arguments):     |  |
|   | e.g. John ran <i>to the store</i> .<br>LS: <b>do'</b> (John, [ <b>run'</b> (John)]) & INGR <b>be-at'</b> (store, John) |

The first class of adpositions function as the head and nucleus of the phrase in which they occur. They add substantive semantic information to the clause and introduce a modifier of the clause. The second class of adpositions don't behave as the head of the adpositional phrase and don't

license the argument they mark. These adpositions introduce a core argument of the predicate and serve as a free-morphemic case marker. Finally, the third class of adpositions are predicative (i.e. function as the head of the adpositional phrase), but introduce a core argument rather than a modifier. For example, *to* in the example of (7c) has its own LS (“INGR **be-at**’ (store, John)”) and shares one of its arguments (*John*) with the LS of the verb. This argument sharing is the hallmark of argument-adjunct adpositions and distinguishes them from adjunct adpositions and argument-marking adpositions.<sup>7</sup>

Given the above classification, let us have a look at how RRG analyzes English counterparts to (1a,n) and comitative uses of *with* as illustrated in (8a-b):

- (8) a. Pat loaded hay on the truck *with Kim*. (Pat and Kim loaded hay on the truck)  
 LS: [**do**’ (Pat  $\wedge$  Kim, O)] CAUSE [**BECOME be-on**’ (truck, hay)]
- b. Pat served wine *with cheese* to the guest. (Pat served wine and cheese to the guest)  
 LS: [**do**’ (Pat, O)] CAUSE [**BECOME have**’ (guest, wine  $\wedge$  cheese)]

Van Valin (2005: 110) formulates the instrumental case rule [ICR] in (9) (cf. Foley & Van Valin 1984: 87, Van Valin & LaPolla 1997: 378-382):<sup>8</sup>

- (9) Instrumental case rule  
 Assign instrumental case to non-MR *b* argument if, given two arguments, *a* and *b*, in a logical structure, with (i) both as possible candidates for a particular macrorole and (ii) *a* is equal or higher (to the left of *b*) on the AUH, *b* is not selected as that macrorole.

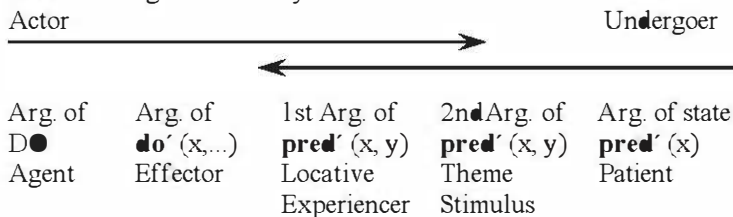
The essential idea behind (9) is that those effector or theme arguments that fail to receive an actor or undergoer status according to the Actor-Undergoer Hierarchy [AUH] receive instrumental case:

<sup>7</sup> There is one more logically possible class of adpositions missing in the three-way classification in (7): non-predicative adpositions that function as a marker of clausal adjuncts. My proposal (see Sections 4 and 6 below) is that the instrumental prepositions in English (*with*, *by*) belong to this class.

<sup>8</sup> See Farrell (2009) for discussion of an exception to the ICR in (9).



## (10) Actor-Undergoer Hierarchy



## (11)a. (1a) Instrument

[do' (Ivan, [use' (Ivan, knife)])] CAUSE [[do' (knife, [cut' (knife, bread)])] CAUSE [BECOME cut' (bread)]]

## b. (1n) Locatum

[do' (they, Ø)] CAUSE [BECOME be-on' (barge, firewood)]  
[Undergoer=Locative]

The ICR applies to (1a,n), whose LSs are given in (11a-b), respectively. (11a), respectively, assigns an actor and undergoer status to *Ivan* and  *хлеб* 'bread.' The instrumental argument *nož* 'knife' cannot be an actor (despite the fact that both *Ivan* and *nož* 'knife' are an effector), since it is subordinated to its user (*Ivan*). It receives non-macrorole status and receives instrumental case according to the ICR. Likewise, *drova* 'firewood' in (1n) fails to receive an undergoer status, despite the fact that it is the highest-ranking argument with respect to the undergoer end of the AUH and is more qualified for receiving the undergoer status than *baržu* 'barge.' This is the reason the ICR applies to the theme argument *drova* 'firewood' under the assumption that the ICR has priority over the dative case rule [DCR] ("Non-macrorole core arguments receive dative case") when they are in conflict.<sup>9</sup>

The ICR also applies to the comitative uses of *with* in (8a-b) as follows. *Pat* and *Kim* in (8a) are equally qualified to receive an actor status, but only *Pat* receives it, while *wine* and *cheese* in (8b) are equally qualified to receive an undergoer status, but only *wine* receives it. This leaves *Kim* in (8a) and *cheese* in (8b) with no macrorole status and makes them conform to the ICR. This is the reason both of them receive instrumental marking.

Van Valin & LaPolla (1997: 376-382) further try to extend the ICR to English counterparts to (1f), but this move raises the question of how to

<sup>9</sup> Van Valin (2005: 110) states that instrumental is not a default case and has specific conditions on its application (as described in (9)). See Van Valin (2005: 109-115) for discussion of how the ICR interacts with the DCR in Croatian, Dyirbal, and English.

classify the instrumental preposition *with* in terms of (7). It is a non-predicative preposition, since it has no logical structure of its own. However, it can occur not only with a core argument (as illustrated in (1n)), but also with a clausal adjunct (as illustrated in (1a-m)). The fact that *with* spans the syntactic distinction between core arguments and clausal adjuncts suggests that it is a default case marker of some sort, but Van Valin & LaPolla (1997) and Van Valin (2005) deny it the default status and don't address the question of how to spell out the default property of *with* (or, more generally, instrumental cases/adpositions).

I will propose in the next section that both the instrumental and dative case in Russian are default oblique cases with varying degrees of specificity.

#### 4. Proposal: An elsewhere definition of instrumental case

This section shows how to incorporate Jakobson's (1936/1984) featural definitions of nominative, accusative, dative, and instrumental case into RRG. (12) is a verbal paraphrase of his featural definition of instrumental case (Jakobson 1936/1984: 82-83, bold in the original):

Table 1: Jakobson's (1936/1984) definitions of cases in Russian

	Peripheral	Directional	Quantified
Nominative			
Accusative		+	
Genitive			+
Dative	+	+	
Instrumental	+		
Locative	+		+

- (12) ..., the I [=instrumental] itself denotes nothing more than **peripheral status**; it occupies the same position among the peripheral cases that the N [=nominative] does among the full cases: that of the unmarked category. Correspondingly the I, like the N, tends toward the role of pure "lexical form". Insofar as this tendency is realized, the peripherally marked I becomes an **adverb**.... Everything other than peripheral status is given in individual uses of the I by the actual meaning of its referent and by the context, but not by the case form.

The above definitions presuppose that each case has its *general meaning* (from which its *particular meanings* are derived) and that the former is

defined independently of its environment, while the latter are derived from the interaction of the former with the clausal and situational context. I will adopt the Jakobsonian distinction between the *general* and *particular meanings* of case and will propose a set of rankable and violable constraints that defines each case with reference to its *general meaning* alone (see Nakamura (1999) for the first OT-RRG account of case systems).

Three remarks about Jakobson's (1936/1984) definitions of cases are in order. First, we may take the feature [directional] as referring to non-actor core arguments, since it is defined as referring to those verbal arguments that the action of the verb is directed to. This characterization allows us to draw a parallelism between Jakobson's featural definitions of dative and accusative case and the RRG distinction between non-macrorole and undergoer arguments. Specifically, the combination of [directional] and [peripheral] (which defines dative case) refers to non-macrorole core arguments, while [directional] alone (which defines accusative case) refers to undergoer arguments. Second, nominative is defined as the least marked case with no semantic content at all. This definition accords well with the fact that nominative nouns alone have the pure naming function and explains why nominative case may mark actor or undergoer arguments, but it isn't enough to determine its syntagmatic distribution in the clause (Kilby 1977). For example, it fails to explain why undergoer arguments of passivized transitive verbs receive nominative case, as illustrated in (11):

- (1)1. *okno*                    *bylo*                    *razbito*                    *mal'čikami*.  
 window.NOM            was                    broken                    boys.INSTR  
 'The window was broken by the boys.' (Wierzbicka 1980)

Third, instrumental is unmarked with respect to both directionality and quantification and is thus defined as the default peripheral case. The fact that the featural definition of instrumental case refers to a proper superset of what the definition of dative case refers to suggests that instrumental case should be defined in such a way as to reflect the superset relationship between instrumental and dative. The above discussion leads us to define dative and instrumental case in terms of violable constraints, as shown in (13a-b), under the assumption that (13b) always outranks (13a):

- (13)The instrumental and dative case in Russian [Preliminary]  
 a. All nouns other than macrorole arguments receive instrumental case.  
 b. Non-macrorole direct core arguments receive dative case.

(13a-b) and their ranking ensure that non-macrorole direct core arguments receive dative case and that all nouns other than macrorole arguments that (13b) fails to apply to are qualified to receive instrumental case and explain the complementary distribution of dative and instrumental.

However, (13) suffers from two problems. First, (13b) makes an incorrect prediction that the locatum noun in (1n) receives dative case, since it is a non-macrorole core argument. The most parsimonious way to solve this problem is to revise (13b) in such a way that it cannot apply to non-macrorole core arguments that involve a violation of the AUH (as in (1a,n)). Second, (13a) cannot handle the instrumental marking of predicative adjectives as illustrated in (2b). These predicative uses require us to revise (13a) in such a way that it may apply to both nouns and adjectives.

The above consideration leads to modification of (13a-b) as shown in (14a-b), under the crucial assumption that (14b) always outranks (14a):

(14) The instrumental and dative case in Russian (Revised)

- a. All case-bearing elements other than macrorole arguments receive instrumental case.
- b. Non-macrorole direct core arguments that do not involve a violation of the AUH receive dative case.

Given that (14b) assigns dative case only to a subset of non-macrorole direct core arguments that involves no violation of the AUH, (14a) assigns instrumental case to the locatum argument in (1n).<sup>10</sup> The above revision allows us to provide a unified account of the uses of the instrumental case in (1a-n). Furthermore, allowing predicate nouns and adjectives in predicate constructions as illustrated in (2a-b) to receive case licenses the instrumental marking of the predicate nouns and adjectives and allows us to maintain the Jakobsonian tenet that each case has its invariant meaning that is independent of the syntactic and discourse-pragmatic context.<sup>11</sup>

---

<sup>10</sup> It is important to note that (14b) applies to direct core arguments alone. This needs emphasis, since some languages (e.g. Japanese) have two types of the dative form: postpositions and case markers (Sadakane & Koizumi 1995; Kishimoto, this volume). The former uses of *ni* (the dative form in Japanese) are predicative; they involve a layered structure with the postposition in the nucleus that takes a core argument as its complement. In contrast, the latter uses of *ni* are non-predicative; they serve as a case marker of core arguments and involve no layered structure. (14b) is designed to apply only to the latter (non-predicative) uses.

<sup>11</sup> One point of contrast between the nominative-marked and instrumental-marked predicates in (2a-b) is the one between permanence and temporariness (see Kagan (2020: 223-236) for a succinct summary of what governs the case alternation), but

To summarize, I have shown that (14a) covers all the uses of instrumental case in (1a-n) and (2a-e) by restricting the application of (14b) to those non-macrorole direct core arguments that do not involve a violation of the AUH and ranking (14b) above (14a).

## 5. Contextualizing Instrumental Case

One may wonder how the wide variety of contextual meanings of instrumental case are derived from the highly abstract definition in (14a). Jakobson (1936/1984) analyzes the invariant meaning of instrumental case as denoting peripheral status, but he leaves it open how its particular meanings are derived from the semantics of an instrumental-marked noun and its syntactic and/or discourse-pragmatic context. The problem of how to contextualize the invariant meaning of instrumental case merits serious consideration, since (14a) only indicates that all nouns it marks are not macrorole arguments and provides no clue as to the role played by the instrumental-marked noun in the clause.

An important clue to the contextualization problem comes from Generative Lexicon [GL] (Pustejovsky 1995; cf. Jackendoff 2002: 369-373), which proposes that a noun, verb, and adjective have an associated Qualia Structure [QS] with four different qualia, constitutive, formal, telic, and agentive. (15a-d) describe these four qualia roles of nouns with a few illustrations:<sup>12</sup>

- (15) a. Constitutive quale: the relation between an object and its structural attributes:
- (i) its part structure or what larger structure it is a part of
  - (ii) the material it is composed of (e.g. *house* → bricks, wood)
- b. Formal quale: what distinguishes it within a larger domain (e.g. orientation, magnitude, shape, dimensionality, color, shape)
- c. Telic quale: information about activities in which an object takes part:
- (i) its purpose (e.g. *pencil* → writing with, *book* → reading, *fork* → eating with)
  - (ii) its proper function (e.g. *heart* → pumping blood)

---

it is outside the scope of the present paper to provide a full account of the alternation illustrated in (2a-b). I only assume that (14a) does not apply where the case-agreement option is available (cf. Corbett 1988: 41; Krasovitsky et al. 2008).

<sup>12</sup> The notion of QS was incorporated into RRG by Van Valin & LaPolla (1997: 184-186).

- d. Agentive quale: the origin of an object or what brings it about  
(e.g. *house*→building, *book*→writing, *rice*→growing)

The constitutive quale is mainly concerned with a part-whole structure of the object denoted by the noun and its material, the formal quale presupposes a taxonomic structure in which the object is located, the telic quale refers to an activity in which the object participates (in particular, its purpose and proper function), while the agentive quale informs about how the object comes into being.

Another GL notion that is directly relevant to the discussion to follow is co-composition. This operation applies when there is an explicit reference to the predicate or the predicate's semantic type within the argument's lexical entry including its QS. A typical example is an alternation between the *change-of-state* and *creation* senses of creation verbs. (16a) describes a change of state of the fence (the fence became covered with paint), while (16b) describes a process of creation (the picture came into being as a result of painting) (Pustejovsky & Batiukova 2019: 329-330):

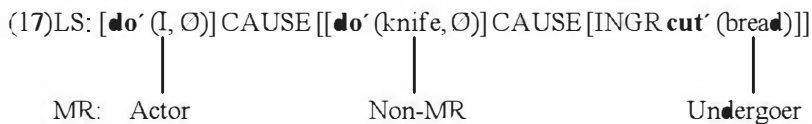
- (16)a. John painted a fence. [change-of-state]  
b. John painted a portrait. [creation]

The notion of co-composition allows us to derive the semantic contrast between (16a) and (16b) from which quale of their undergoer argument is invoked: (16a) refers to its formal quale (i.e. a physical object), while (16b) refers to its agentive quale (i.e. paint turns into a portrait). This means that the *creation* sense in (16b) doesn't exist on a priori basis, but emerges as a consequence of a bottom-up composition in which the undergoer argument affects the meaning of the clause beyond its role as an argument of the predicate.

The GL notions of QS and co-composition make it possible to derive the variety of contextual meanings of the instrumental-marked nouns in (1) from (14a) with no appeal to polysemy. First, the crucial difference between (1a) and (1b) is that the former involves a causal chain, while the latter does not:

- (1)a. *Ivan rezal xleb nožom.*  
Ivan.NOM cut bread.ACC knife.INSTR  
'Ivan cut bread with a knife.' (Janda 1993)
- b. *On el ložkoj ikru.*  
he.NOM ate spoon.INSTR caviar.ACC  
'He ate caviar with a spoon.' (Demjjanow & Strigin 2003)

In (1a), *nož* ‘knife’ does something to the bread that causes the bread to be cut, while *ložka* ‘spoon’ in (1b) plays no such active role in the event and is only used for some purpose (in this case, eating caviar).<sup>13</sup> This is the reason *nož* in (1a) (but not *ložka* in (1b)) is available for actor selection. (17) shows the macrorole assignment in (1a):<sup>14</sup>



The instrumental argument *nož* ‘knife’ is a non-macrorole argument that involves a violation of the AUH and receives instrumental case from (14a).

The crucial question here is how the LS of (1a) given in (17) is derived. First, the starting point for our consideration is the definition of instrumental case in (14a). Its radically underdetermined meaning gives no clue as to which argument position (in the semantic representation of the clause) *nož* is linked to. In order to avoid a violation of the Completeness Constraint (see Van Valin (2005: 129-130) for its formulation), we are forced to interpret *nož* as a participant of an event involved by its QS (so that an additional LS component may be derived which allows *nož* ‘knife’ to be linked) and to identify the relation between the event denoted by the verb and the one in its QS. (18a) is part of the telic quale of *nož* ‘knife’ that is relevant here (cf. Van Valin 2012), while (18b) is the LS of the verb:

(18)a. QS of knife (y)

TELIC: [**do'** (x, [**use'** (x, y)))] (CAUSE [**do'** (y, [...])])  
 b. LS of *rezat'*: [**do'** (x, O)] CAUSE [INGR **cut'** (z)]

(18a) describes a situation in which one uses a knife in some way to do something and makes it optional whether or not the knife plays an active role in the causal chain. That is, if the knife is construed as exerting the force (coming from its human user) onto some object (as in (1a)), its telic quale is “[**do'** (x, [**use'** (x, y)))] CAUSE [**do'** (y, [...])],” while if the knife is construed as a tool that helps its user to do something, its telic quale is “[**do'** (x, [**use'** (x, y))].” Since the knife plays an active role in (1a), we may say that merging the LS in the telic quale of *nož* ‘knife’ in (18a) with the LS of the verb *rezat'* ‘cut’ in (18b) yields the LS of (1a) in (17).

<sup>13</sup> See Van Hooste (2018: Ch.4, and this volume) for related discussion.

<sup>14</sup> ‘I’ in (17) is an abbreviation for *Ivan*.

An analogous account holds for the instrumental-marked nouns in (1b) and (1c-e):

- (1) c. *On* *exal* *poezdom.*  
 he.NOM **drove** train.INSTR  
 ‘He was going **by train**.’ (Demjjanow & Strigin 2003)
- d. *On* *pil* *vino* *litrami.*  
 he.NOM **drank** wine.ACC liters.INSTR  
 ‘He **drank wine by the liter**.’ (Demjjanow & Strigin 2001)
- e. *Fašisty* *uničtožali* *ljudelj* *millionami.*  
 fascists.NOM **destroyed** people.ACC millions.INSTR  
 ‘The fascists killed **people by the millions**.’ (Janda 1993)

First, (1b) differs from (1a), in that its implement argument *ložka* ‘spoon’ (unlike *nož* ‘knife’ in (1a)) doesn’t participate in the causal chain. The noun *ložka* has a telic quale in (19a), the shaded part of which specifies the default purpose of using a spoon (i.e. eating something).<sup>15</sup>

- (19) a. **Q**S of **spoon** (y)  
 TELIC: [**do**’ (x, [**use**’ (x, y)))] PURP [**do**’ (x, [**eat**’ (x, z)))]  
 b. LS of *est*: **do**’ (x, [**eat**’ (x, z)])  
 c. LS of (1b): [**do**’ (he, [**use**’ (he, spoon)))]  
 PURP [**do**’ (he, [**eat**’ (he, caviar)))]

The Completeness Constraint dictates that the implement noun is linked to an argument slot in the semantic representation of the clause and requires us to construe *ložka* ‘spoon’ as a participant of an event involved by its telic quale. Merging the LS of the verb *est*’ in (19b) with the LS in the telic quale of *ložka* in (19a) yields the LS of (1b) in (19c).<sup>16</sup> What is notable about (1b) is that its meaning comes mainly from the telic quale of *ložka*

<sup>15</sup> The use of the PURP connective in (19a) is motivated by the observation that instrumental constructions such as (1b) can be paraphrased by constructions such as *He used a spoon to eat caviar* (Lakoff 1968). Van Valin (2012) alternatively paraphrases constructions (such as (1b)) as *He ate caviar, using a spoon* and proposes that the telic quale of a spoon is something like “[**do**’ (x, [**eat**’ (x, z)  $\wedge$  **use**’ (x, y)))]”. I will leave it for further research to decide which paraphrase is (more) appropriate. See also Koenig et al. (2007) and Rissman (2011, 2013) for relevant discussion.

<sup>16</sup> It is important to note that (1b) combines two events, one denoted by the verb and the other involved by the telic quale of *ložka* ‘spoon,’ albeit not symmetrically. This means that the former is subordinated to the latter.



and that the verb *est'* and its arguments elaborates the event schema provided by *ložka*.

Second, (1c) contains the instrumental-marked noun (*poezdom* 'train') that indicates the means of transportation. This noun has a set of telic qualia, two of which are given in (20a):

- (20) a. QS of **train** (y)  
 TELIC: **do'** (x, [**ride.on'** (x, y)]), **do'** (y, [**carry'** (y, x)])  
 b. LS of *exat'*: **do'** (x, [**go.by.vehicle'** (x)])  
 c. LS of (1c): **do'** (he, [**go.by.vehicle'** (he)  $\wedge$  **ride.on'** (he, train)])

The first telic quale states that people **ride** a **train**, while the second one states that a train carries people. The first one is relevant here. Merging the LS of the verb *exat'* 'go by vehicle, drive' in (20b) with the LS in the telic qualia of *poezd* in (20a) yields the LS of (1c) given in (20c). This merge makes the manner of motion more specific (by identifying the means of transportation) and introduces an additional LS component ("**do'** (x, [**ride.on'** (x, y)])") that comes from the telic quale of *poezd* 'train' and allows *poezd* to be linked.

Third, both (1d) and (1e) involve the use of a unit of measurement. These instrumental-marked nouns are termed *measure instrumental* and denote a measure unit by which to quantify some object, while engaging in some activity involving it (e.g. drinking wine, killing people). (21a) is the telic quale of measure units (including *litr* 'liter' and *million* 'million'), while (21b) is the LS of the verb *pit'* 'drink.' Merging the LS of the verb *pit'* 'drink' in (21b) with the LS in the telic quale of *litr* 'liter' in (21a) yields the LS of (1d) given in (21c):

- (21) a. QS of **liter** (y)  
 TELIC: **do'** (x, [**use'** (x, y)]) PURP **do'** (x, [**measure'** (x, z)])  
 b. LS of *pit'*: **do'** (x, [**drink'** (x, z)])  
 c. LS of (1d): **do'** (he, [**drink'** (he, wine)])  $\wedge$  [**use'** (he, liter)  
 PURP **measure'** (he, wine)]

(21c) combines two events, drinking wine (main event) and using the unit of liter to measure the amount of wine (subordinate event). We may extend

the above account of (1d) to (1e), which has another measure unit (*millio-  
na* ‘million’) that is used to count the number of *ljudi* ‘people.’<sup>17</sup>

The contrast between (1d-e) and (1b) is that the purpose of using the unit of liter/million is fixed, while the purpose of using a spoon includes stirring during cooking and measuring the volume of liquid in addition to being used when eating food. The observation that the spoon has more than one common function suggests that the shaded part of the telic quale in (17a) (“do’ (x, [use’ (x, y)])”) is not lexical, but arises pragmatically.

In contrast to (1a-e), the instrumental-marked nouns in (1f-h) (*molčanie* ‘silence,’ *osen’* ‘autumn,’ and *časny* ‘hours’) have in common that their referents have no (default) value for their constitutive, telic, or agentive quale:

- (1) f. *Monax dolžen svjazat’ usta svoi molčaniem.*  
 monk.NOM must tie.up lips own.ACC silence.INSTR  
 ‘A monk must seal his lips **with silence.**’ (Janda 1993)
- g. *Oni vstretilis’ osen’ju v Pariže.*  
 they.NOM met autumn.INSTR in Paris.LOC  
 ‘They met **in autumn,** in Paris.’ (Wierzbicka 1980)
- h. *On (celymi) časami vozilsja*  
 he.NOM (whole.INSTR) hours.INSTR fiddled  
*s radiopriemnikom.*  
 with radio.INSTR  
 ‘**For hours on end** he would fiddle with the radio.’ (Wierzbicka 1980)

In other words, they are not conventionally associated with any constituent material, purpose, function, or origin. This yields the consequence that we need to interpret these nouns with no reference to any event involved by their QS. Since they are neither an argument of the predicate nor part of the predicate, we have to conclude (by a process of elimination) that these instrumental-marked nouns function as modifiers: the instrumental-marked noun in (1f) modifies the manner of the action denoted by the verb *svjazat’* ‘tie up’ and serves as a nuclear modifier, while the instrumental nouns in (1g-h) are temporal modifiers that narrow down the time at which the event denoted by the verb took place and function as a core modifier.<sup>18</sup>

<sup>17</sup> (1e) receives an iterative interpretation, which makes the sentence imperfective and licenses *millionami* ‘by the millions’ (which presupposes imperfectivity of the whole sentence) to appear in the sentence.

<sup>18</sup> I propose to classify manner adverbs as in (1f) as nuclear modifiers, since they neither take the whole core within their scope nor refer to any core argument (un-

Next, let us consider (1i), in which the instrumental-marked noun (*lesom* ‘forest’) is construed as referring to the trajectory of motion denoted by the verb. Russian has two strategies of encoding the trajectory of a motion event: a bare NP in the instrumental case (as in (1i), repeated below as (22a)) and a prepositional phrase headed by *po* ‘along’ (as in (22b)).<sup>19</sup>

- (22)a. *Oni šli pjat' kilometrov lesom.*  
 they.NOM went five.ACC kilometers.GEN forest.INSTR  
 ‘They walked 5 kilometers through a forest.’ (= (1i))
- b. *Oni šli pjat' kilometrov po lesu.*  
 they.NOM went five.ACC kilometers.GEN along forest.DAT  
 ‘They walked 5 kilometers through a forest.’

The motion verb in (22a-b) is *idti* ‘go, walk,’ which is the most generic motion verb in Russian, denotes a unidirectional motion, and occurs in almost half of all clauses that contain an instrumental-marked path noun (Philippova 2017: 362).<sup>20</sup>

The path preposition *po* in (22b) is an argument-adjunct preposition, since it is a preposition associated with its own LS, but (unlike an adjunct preposition) introduces an argument rather than a modifier. It also differs from an argument-marking preposition, in that the meaning of its argument is not derived from the verb, but it shares its argument with the LS of the verb. In contrast, the path noun in (22a) is an adjunct, simply because it is not subcategorized by the verb.<sup>21</sup> The underdetermined meaning of instrumental case on the path noun requires the path to be linked to an argument position of an event involved by its  $\Theta$ S.

I propose that the path interpretation of *lesom* ‘forest’ in (22a) is derived from one of its telic roles in (23a): *les* ‘forest’ is a place that people travel in by whatever means and for whatever purpose.<sup>22</sup>

like subject-oriented adverbs). Van Valin (personal communication) suggests that the subtypes of adverbs (including manner adverbs) do not map neatly onto the layered structure of the clause and require their own distinctions.

<sup>19</sup> Another preposition *čerez* ‘through’ may be used instead of *po* ‘along.’

<sup>20</sup> See Philippova (2017) for a multifactorial account of the availability of bare instrumental constructions as illustrated by (22a) (= (1i)).

<sup>21</sup> The contrast between (22a) and (22b) comes down to whether or not there is a preposition that shares its argument with the verb and allows the prepositional phrase that contains the path noun to function as a core argument.

<sup>22</sup> The motion verb in (23a) is underspecified with respect to the manner of motion. (23b) is the LS of activity (or active accomplishment) verbs of motion represented by *idti* (Van Valin 2018). This is a departure from Van Valin & LaPolla (1997) and

(23)a. ●S of forest (y)<sup>23</sup>

TELIC: [do' (x, [go' (x)]) &amp; be-in' (y, x)]

b. LS of *iditi*: [do' (x, [go' (x)]) ∧ PROC cover.path.distance' (x, (z))]

c. LS of (1i): do' (they, [go' (they)]) ∧ PROC cover.path.distance' (they, 5 km) &amp; be-in' (forest, they)

The Completeness Constraint requires that the path noun in (22a) is linked to an argument position in the semantic representation of the clause and requires us to interpret *les* 'forest' as a participant of an event involved by its telic quale. Merging the LS of the verb in (23b) with the LS in the telic quale of *les* in (23a) yields the LS of (1i) given in (23c).

Next, let us consider how the meaning of domain restriction is derived in (1j):

- (1)j. *Lošad' sosedā lučše moej*  
 horse.NOM neighbor.GEN better mine.GEN  
*i siloj i krasotoj.*  
 and strength.INSTR and beauty.INSTR  
 'The neighbor's horse is better than mine both in strength and in appearance.' (Janda 1993)

The first point to note is that both *sila* 'strength' and *krasota* 'beauty, appearance' are part of the formal qualia of *lošad* 'horse.' The fact that these attributes constitute part of the attributes of a horse explains why the two instrumental-marked nouns relativize the evaluation of the neighbor's horse by the speaker and receive the interpretation of domain restriction. The fact that the two nouns are not subcategorized by *lučše* suggests that they function as a modifier of the predicate.<sup>24</sup>

We are now in a position to consider the instrumental-marked noun in (1k):

---

Van Valin (2005) with respect to how to represent the semantics of activity/active accomplishment verbs. See Van Valin (2018) for detailed discussion of how to capture the incremental changes of states denoted by these motion verbs.

<sup>23</sup> (23a) is different from the telic role of *ložka* 'spoon' in (19a), in that the forest is not an artifact and has no built-in purpose, but the fact that (19a) and (23a) describe how the artifact and natural objects interact with people allows us to treat (23a) on a par with (19a) despite that (23a) doesn't show the inherent purpose of *les* 'forest.'

<sup>24</sup> Both *sila* 'strength' and *krasota* 'beauty, appearance' are abstract nouns and don't serve as secondary predicates to *lošad* 'horse.'

- (1)k. *Petr kivnul golovoj.*  
 Peter.NOM nodded head.INSTR  
 ‘Peter nodded his head.’ (lit. ‘Peter nodded with the head’)  
 (Wierzbicka 1980)

What is peculiar about (1k) and analogous examples in (24a-c) is that the body-parts are not construed as a target of manipulation. For example, (1k) doesn’t mean that Peter moved his head in order to have it in a new position. (1k) denotes a semelfactive situation, in which a punctual event occurs without causing any change of state.<sup>25</sup> This observation is corroborated by the fact that (1k) has no agentless reflexive passive counterpart, as shown by (24d), in contrast to a transitive construction with an accusative-marked body-part noun that has a reflexive passive counterpart, as shown by (25b):

- (24)a. *Ivan požal plečami.*  
 Ivan.NOM shrugged shoulders.INSTR  
 ‘Ivan shrugged his shoulders.’ (lit. ‘Ivan shrugged with the shoulders’) (Wierzbicka 1980)
- b. *Maša topnula nogoj.*  
 Mary.NOM stamped foot.INSTR  
 ‘Mary stamped her foot.’ (lit. ‘Mary stamped with the foot’)  
 (Wierzbicka 1980)
- c. *Ona byla blečna licom.*  
 she.NOM was pale.SF.NOM face.INSTR  
 ‘She was pale in the face.’ (Janda 1993)
- d. *\*Golova Petra kivnula-s’.*  
 head.NOM Peter.GEN nodded-REFL  
 ‘Peter’s head nodded.’ (Sofia Tumanova personal communication)
- (25)a. *Bol’noj otkryl glaza.*  
 sick.NOM opened eyes.ACC  
 ‘The sick man opened his eyes.’ (Wierzbicka 1980)
- b. *Glaza bol’nogo otkryli-s’.*  
 eyes.NOM sick.GEN opened-REFL  
 ‘The eyes of the sick man opened.’ (Wierzbicka 1980)

<sup>25</sup> Wierzbicka (1980: 25) notes that (24a) presents Ivan’s action as a whole as an ‘institutionalized’ sign that represents a dismissive behavior. An analogous observation can be made about nodding one’s head and stamping one’s foot, which indicate consent and anger, respectively.

The contrast between (24d) and (25b) indicates that (25a) describes a change of state in the sick man's eyes, while (1k) construes Ivan's head as part of his action rather than a target his action is directed against. That is, the body-part noun *golova* 'head' doesn't saturate the verb's argument slot in its LS, but restricts the verb's meaning by specifying the body-part directly involved in the event.<sup>26</sup>

The following consideration gives a crucial clue to how the body-part noun in (1k) restricts the verb's meaning. First, the meaning of a part-whole relation is derived from the constitutive quale of *golova* 'head,' which encodes information about a larger structure (in this case, the human body) the head is part of as well as its components.<sup>27</sup> (26b) is a partial description of the QS of *golova* 'head':

- (26)a. SEML do' (Peter, [nod' (Peter)])  
 b. FORMAL: head (x)  
 C ONST: have.as.part' (y:human\_body, x)  
 c. SEML do' (Peter, [nod' (Peter)])  $\wedge$  have.as.part' (Peter, head)

Combining the LS in the constitutive quale of *golova* in (26b) with the LS of the verb *kivmul* 'nod' in (26a) yields (26c), which represents the sentential meaning of (1k). What is peculiar about (1k) (and (24a-c)) is that it involves two verbs, the main verb *kivmul* and the null possessive verb that comes from the constitutive quale of *golova* 'head' and that each of them comes with its own core. The two cores share one core argument (*Petr*) and form a non-subordinate core juncture.<sup>28</sup> Thus, the body-part noun *golovoj* in (1k) doesn't belong to the matrix core and serves as a modifier

<sup>26</sup> Note that (1k) is semantically complete without the body-part noun. The same observation holds for (24b-c) ((24a) needs *plečami* for disambiguation, since the verb *požat'* may occur with another body-part noun *ruku* 'hand' as illustrated in the following example:

*Ivan*      *požal*      *ruku*      *Petru/Petra*.  
 Ivan.NOM shook hand.ACC Peter.DAT/ Peter.GEN

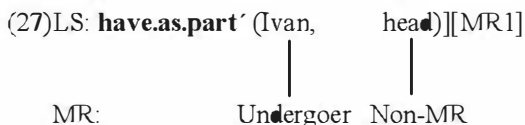
'Ivan shook Peter's hand.' (Sofia Tumanova personal communication)

<sup>27</sup> The Completeness Constraint requires that the body-part noun is linked to the semantic representation of the clause. This is the reason its constitutive quale is invoked when one is interpreting (1k) (body-part nouns have no default value for their telic or agentive quale).

<sup>28</sup> Schultze-Berndt & Himmehnam (2004) propose that secondary-predicate constructions embody a non-subordinate core juncture (see also Ullrich (2018) for an RRG account). (1k) has the same nexus-juncture type as secondary-predicate constructions, but (1k) departs from them, in that no predicative relation holds between the possessor and possessum noun.

of the main verb that specifies which body-part was directly involved in the event, but it is distinguished from clausal adverbs including the instrumental-marked nouns in (1f-h), since it belongs to the other core formed by the null possessive verb.

The next question to ask is how to account for the instrumental marking of the body-part noun in (1k). The point of departure is the observation that Russian is a prime example of *be*-languages (Isačenko 1974) and productively combines the *be*-form (*byť*) with an oblique noun (*u* + genitive-marked noun) that represents a possessor. This observation suggests that possessive verbs in Russian receive only one macrorole despite having two LS arguments. The only macrorole that the possessive verbs receive is undergoer, since they have a two-place state predicate (“**have.as.part**’ (x, y)” or “**have**’ (x, y)”) in their LS. The possessor argument is chosen as the undergoer when it is construed as more prominent than the possessum:



(27) describes the marked undergoer assignment to the possessor argument (under the assumption that the macrorole assignment proceeds in each core independently). The body-part noun *golova* receives instrumental case according to (14a), since the linking in (27) violates the AUH.<sup>29</sup>

Finally, (1l,m) involve voice alternation:

- (1) l. *okno bylo razbito mal’čikami.*  
 window.NOM was broken boys.INSTR  
 ‘The window was broken by the boys.’ (Wierzbicka 1980)
- m. *Ivan švyryjal(-sja) kamnjami.*  
 Ivan.NOM was.throwing(-REFL) stones.INSTR  
 ‘Ivan was throwing stones.’ (Wierzbicka 1980)

<sup>29</sup> The possessum argument is the default choice for undergoer. The default macrorole assignment leads the possessor to become a non-macrorole and to receive oblique marking. This is illustrated by the following example, in which the possessor argument receives dative case or is accompanied by the locative preposition *u* ‘at, by’ (the external possessor is highlighted in bold below):

*Mina otorvala mne’u menja nogu.*  
 mine.NOM tore.off 1SG.DAT/at 1SG.GEN leg.ACC  
 ‘The mine blew off my leg.’ (Sofia Tumanova personal communication)

(1l) and (1m), respectively, illustrate an adjunct actor in passive constructions and an adjunct undergoer in antipassive constructions, both of which receive their semantic interpretation from the verb and receive instrumental case from (14a).

To summarize, this section has shown that the instrumental-marked nouns in (1a-n) receive their semantic interpretation from (i) the predicate, (ii) their Qs-based semantics (in particular, the telic and constitutive quale), and (iii) their interaction and receive instrumental case from (14a).

## 6. Extension: The case of English

This section aims to conduct a preliminary investigation of how far (14a) extends to English, with a particular focus on the partially overlapping and partially complementary distribution of *with* and *by*. Their differential distribution requires us to go beyond (14a), since (14a) applies to both *with* and *by* and doesn't help to differentiate their distribution.

The first point to note about *with* is that it syncretizes with the comitative prepositions (see below in this section).<sup>30</sup> In this respect, English is in contrast to Russian, which has only one instrumental marker that is distinct from the comitative preposition (*s*) that takes an instrumental-marked noun as its complement, as illustrated by (28):

- (28) *Slepec s sobakoj perešel ulicu.*  
 blind.man.NOM with dog.INSTR walked.across street.ACC  
 'The blind man crossed the street with a dog.' (Janda 1993)

Analyzing *with* as involving the comitative-instrumental syncretism allows us to distinguish their instrumental uses from their comitative uses and to describe the two uses separately (in contrast to Van Valin & LaPolla (1997), Van Valin (2005), and Farrell (2009)).

(29a-p) and (30a-b) illustrate the non-comitative uses of *with* and *by* and the comitative uses of *with*, respectively (the *with/by*-phrases in (29) and (30) are highlighted in bold).<sup>31</sup>

- (29) Instrument  
 a. John cut the bread **with a knife**.

<sup>30</sup> The comitative-instrumental syncretism is one of the most common types of case syncretism in Indo-European languages (Stolz et al. 2006).

<sup>31</sup> Those uses of *with* and *by* that occur within an NP (e.g. *a house with a pool*, *a poem by Coleridge*) are beyond the scope of this paper. See Van Valin & LaPolla (1997: 189-192) for relevant discussion.



Implement/means

- b. 1. John ate pasta **with a fork**.  
 2. John wrote his notes **by hand** in class.  
 John assassinated the politician **by stabbing him over ten times**.

Transportation

- c. 1. We commuted **with our rental car** to the factory.  
 2. The workers commuted **by car** to their workplace.

Unit

- d. We sell our wine **by the glass, bottle, and/or case**.

Quantity

- e. We shot **down** British bombers **by the hundreds**.

Passive actor

- f. John was beaten **by the students**.

Domain restriction

- g. Ivan was a teacher **by profession**.

Path

- h. John **drove by the river**. [path construed as a line]  
 My friend slipped into the house **by the garage door**. [path construed as a point]

Manner

- i. The students performed the task **with so much difficulty**.

Part-whole relation

- j. John caught me **by the arm/hand/collar**.

Duration

- k. John sings **by night** and runs the dealership **by day**.

Cause

- l. The war **didn't break out by accident**.  
 John **died by falling from a scaffolding**.

Standard

- m. Most merchants in our country play **by the credit card rules**.

Extent

- n. John missed the target **by two inches**.

Locatum

- o. John **loaded** the truck **with hay**.

Terminative

- p. John will finish the task **by tomorrow**.

(30) Accompaniment

- a. John went to the library **with Tom**. (John and Tom went to the library.)  
 b. John served wine **with cheese** to the guest. (John served wine and cheese to the guest.)

A look at (29a-p) and (30a-b) suggests that the semantic interpretations of the two instrumental prepositional phrases fall into three classes, depending on whether they are represented by the *with*-phrase and/or *by*-phrase and that the *with*-phrase and *by*-phrase are in complementary distribution except where either of them is available as in (29b-c):

- (31)a. Those represented by both the *with*-phrase and the *by*-phrase:  
 (29b-c)  
 b. Those represented by the *with*-phrase alone: (29a,i,o), (30a-b)  
 c. Those represented by the *by*-phrase alone: (29d-h), (29j-n), (29p)

Furthermore, there are three observations to make regarding the distribution and productivity of *by*-phrases. First, *by* covers a wider range of semantic interpretations than *with* in the instrumental uses. Second, *by*-phrases as illustrated in (29b-c) allow no element to occur between *by* and its complement noun, while the corresponding *with*-phrases exhibit no such restriction. This contrast suggests that *by*-phrases such as *by hand* and *by car* are fixed lexical units and that one may use a *with*-phrase as an alternative to the corresponding *by*-phrase when one needs to be more specific about the tool or transportation. Finally, the terminative uses of *by* (illustrated in (29p)) indicate the time that ends an action and are very close to those of the so-called terminative case in Estonian (Moseley 1994). The terminative *by*-phrase is in contrast with the *until*-phrase, in that the *by*-phrase occurs with a telic predicate, while the *until*-phrase occurs with an atelic predicate. The use of *by* in (29p) also falls under the scope of (14a), under the assumption that *until* is the terminative preposition that can occur only with an atelic predicate.<sup>32</sup>

The fact that *by* covers a wider range of semantic types than *with* underscores the default status of the instrumental preposition *by*, but the fact that some semantic interpretations of *by* are restricted to a small number of nouns and form an adverbialized lexical unit with their complement noun (e.g. *by train*, *by phone*, *by profession*, *by day*, *by night*, *by mistake*) also suggests that *by* has been losing its ability to form instrumental prepositional phrases productively.

The above discussion suggests that in order to distinguish the distribution of *with* from that of *by*, it is best to begin with the distribution of *with* and then to proceed to the distribution of *by*. It is important at this stage to

---

<sup>32</sup> Both *with* and *by* may be used to express the meaning of spatial proximity (see Farrell (2009) for an RRG account of the spatial proximity and some other uses of *with*). I leave it to another occasion to explain the distribution of the spatial proximity uses of *with* and *by*.

recall from Section 5 that all the Russian counterparts to the instrument, implement/means, transportation, unit, and quantity nouns in (29a-e) are an argument of an activity predicate in their LSs and that the Russian counterpart to the manner noun in (29i) modifies its predicate.<sup>33</sup> Since the locatum noun in (29o) is a core argument, we may formulate the instrumental (as opposed to comitative) preposition assignment constraints in English, as shown in (32):

- (32) *With* assignment constraint
- a. An argument of an activity predicate in the LS of the clause (e.g. (29a-e))
  - b. Core-internal elements (core arguments or nuclear modifiers) (e.g. (29i,o))
- By* assignment constraint
- c. Neither (32a) nor (32b)

(32a-c) apply only to those clausal elements that have a non-macrorole status or no macrorole status. (32a) fails to apply when the adverbialized *by*-phrase is available: the *with*-phrase is not available in (29b-c) when the corresponding *by*-phrase is available (e.g. *by car* vs. \**with a car*), while the *with*-phrase is not available in (29d-e), since unit and quantity nouns as in (29d-e) constitute a closed class comprising mostly measure and container nouns and form lexicalized units together with *by* and *the* (e.g. *by the meter*, *by the kilo*, *by the dozen*, *by the box*). These lexicalized/adverbialized *by*-phrases preempt the use of the productive instrumental preposition *with* in (29d-e).

Finally, a word is in order about the comitative (accompaniment) uses of *with* in (30a-b), whose LSs are given in (33a-b), respectively:

- (33) a.  $\text{do}'$  (John  $\wedge$  Mary, [ $\text{go}'$  (John  $\wedge$  Mary)] & INGR **be-at'** (library, John  $\wedge$  Mary)
- b. [ $\text{do}'$  (John,  $\emptyset$ )] CAUSE [**BECOME have'** (guest, wine  $\wedge$  cheese)]

As noted at the beginning of this section, English has syncretized the instrumental and comitative preposition and has chosen *with* as the syncretized form. This choice reflects the fact that *with* was historically derived from the Old English *wip*, which used to have the meaning of 'toward,

<sup>33</sup> Nuclear modifiers as illustrated in (29i) occur in the periphery of the nucleus and belong to the core, while core modifiers occur in the periphery of the core and belong to the clause.

against,' but took on the meaning of accompaniment from the now-obsolete *mid* and developed some of the instrumental meanings (including the instrument, implement, and manner).<sup>34</sup>

To summarize this section, I have shown that *with* and *by* exhibit a complementary distribution except where both of them are available as in (29b-c) and that when both of them are available, *by* has only lexicalized uses (e.g. *by car*, *by machine*), while *with* can productively form an instrumental prepositional phrase that takes a full NP as its complement.

## 7. Conclusion

I have proposed the novel, underdetermined definition of the instrumental case in Russian together with the definition of dative case slightly revised from the one proposed by Van Valin (1991, 2005) and Van Valin & LaPolla (1997):

- (14)a. All case-bearing elements other than macrorole arguments receive instrumental case.  
 b. Non-macrorole direct core arguments that do not involve a violation of the AUH receive dative case.

I have adapted Jakobson's (1936/1984) decompositional analysis of the Russian cases to RRG and have redefined the instrumental and dative case as default oblique cases with varying degrees of schematicity as in (14a-b). Next, I have explored how to contextualize the underdetermined meaning of the instrumental case in (14a) in terms of the semantics of verbs, the semantics of instrumental-marked nouns (including their QS), and the three-layered structure of the clause. Furthermore, I have shown that the above account of the Russian instrumental case extends to the two instrumental prepositions in English, under the assumption that we may group the two prepositions under the heading of the instrumental case/adposition.

Finally, the results of the present paper have an important theoretical consequence: they obviate the need to postulate a large number of distinct meanings of instrumental case that form a polysemy network (Janda 1993; Narrog & Ito 2007; Malchukov & Narrog 2009; Narrog 2009).<sup>35</sup> The most

<sup>34</sup> The now-obsolete *mid* is found in words such as *midwife* (Skeat 2005).

<sup>35</sup> Narrog & Ito (2007) and Narrog (2009) use crosslinguistic data to build the structured network of distinct meanings/functions of instrumental case, but we may group them together with Janda (1993), since all of them commit themselves to the idea that the instrumental case is polysemous.

serious problem with the above-quoted polysemy-based accounts is that they leave it unclear or unanswered how semantic interpretations of an instrumental-marked noun arise in context.<sup>36</sup> To the extent that the present account is successful in deriving the wide range of semantic interpretations of instrumental case (as illustrated in (1a-n) and (2a-e)) from the interaction of its invariant abstract meaning (which comes from the interaction of (14a) and (14b)) with the clausal context, it lends support to the proposal made in Sections 4 and 5 to integrate with RRG the Jakobsonian case theory, which holds that each case has its invariant meaning independent of its syntactic and discourse-pragmatic environment and that its particular meanings (i.e. semantic interpretations of the case in context) arise through the interaction of its invariant meaning with the clausal and situational context.

### Acknowledgements

This is a substantially revised version of my presentation at the International Conference on Role and Reference Grammar held at Heinrich-Heine-Universität Düsseldorf in July 2015. I appreciate the helpful comments from Robert D. Van Valin, Jr. on the penultimate version of this paper. I am also grateful to Sofiia Tumanova for her help with the Russian data. Needless to say, all remaining errors are my sole responsibility.

### References

- Bailyn, John Frederick. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press.
- Brown, Dunstan. 2007. Peripheral functions and overdifferentiation: The Russian second locative. *Russian Linguistics* 31:61-76.
- Corbett, Greville G. 1988. Agreement: A partial specification based on Slavonic data. In Ferguson, Charles A. & Barlow, Michael (eds.), *Agreement in natural language: Approaches, theories, descriptions*, 25-53. Stanford, CA: CSLI Publications.
- Corbett, Greville G. 2008. Determining morphosyntactic feature values: The case of case. In Corbett, Greville G. & Noonan, Michael (eds.), *Case and grammatical relations: Studies in honor of Bernard Comrie*, 1-34. Amsterdam: John Benjamins.

---

<sup>36</sup> Another criticism that could be levelled against any polysemy-based account of case is that the methods used to discern its (allegedly) distinct meanings are too subjective.

- Demjjanow, Assinja & Strigin, Anatoli. 2001. Measure instrumental in Russian. *ZAS Papers in Linguistics* 22 (Papers on predicative constructions: Proceedings of the workshop on secondary predication, October 16-17, 2000, Berlin). 69-97.
- Demjjanow, Assinja & Strigin, Anatoli. 2003. Real adjuncts in the instrumental in Russian. Lang, Ewald, Maienborn, Claudia & Fabricius-Hansen, Cathrine (eds.), *Modifying adjuncts*, 591-625. Berlin and New York: Mouton de Gruyter.
- Farrell, Patrick. 2009. The preposition *with* in Role and Reference Grammar. In Guerrero, Lilián & Ibáñez, Sergio & Bello, Valeria A. (eds.), *Studies in Role and Reference Grammar*, 179-202. Mexico City: UNAM Press.
- Foley, William A. & Van Valin, Robert D., Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Isačenko, Alexander V. 1974. On 'have' and 'be' languages. In Flier, Michael S. (ed.), *Slavic forum: Essays in linguistics and literature*, 43-77. The Hague: Mouton.
- Israeli, Alina. 2007. Nominative and instrumental variation of adjectival predicates with the Russian copula *byť*: Reference time, limitation, and focalization. In Divjak, Dagmar & Kočańska, Agata (eds.), *Cognitive paths into the Slavic domain*, 21-53. Berlin and New York: Mouton de Gruyter.
- Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Jakobson, Roman. 1936. Beiträge zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus [Contributions to the general theory of case: General meanings of the Russian cases]. *Travaux du Cercle Linguistique de Prague* 6. 240-288. Reprinted in Waugh, Linda R. & Halle, Morris (eds.), 1984. *Russian and Slavic grammar: Studies 1931-1981*, 59-103. Berlin: Mouton.
- Janda, Laura A. 1993. *A geography of case semantics: The Czech dative and the Russian instrumental*. Berlin: Mouton de Gruyter.
- Jolly, Julia A. 1991. *Prepositional analysis within the framework of Role and Reference Grammar*. New York: Peter Lang.
- Jolly, Julia A. 1993. Preposition assignment in English. In Van Valin, Robert D., Jr. (ed.), *Advances in Role and Reference Grammar*, 275-310. Amsterdam: John Benjamins.
- Kagan, Olga. 2020. *The semantics of case*. Cambridge: Cambridge University Press.
- Kilby, David A. 1977. *Deep and superficial cases in Russian*. München: Kubon and Sagner.

- Koenig, Jean-Pierre, Mauner, Gail & Bienvenue, Breton. 2003. Arguments for adjuncts. *Cognition* 89: 67–103.
- Koenig, Jean-Pierre, Mauner, Gail, Bienvenue, Breton & Conklin, Kathy. 2007. What with?: The anatomy of a (proto)-role. *Journal of Semantics* 25(2): 175–220.
- Krasovitsky, Alexander. Long, Alison, Baerman, Matthew, Brown, Dunstan & Corbett, Greville G. 2008. Predicate nouns in Russian. *Russian Linguistics* 32: 99–113.
- Lakoff, George. 1968. Instrumental adverbs and the concept of deep structure. *Foundations of Language* 4: 4–29.
- Malchukov, Andrej & Narrog, Heiko. 2009. Case polysemy. In Malchukov, Andrej & Spencer, Andrew (eds.), *The Oxford handbook of case*, 518–534. Oxford: Oxford University Press.
- Moseley, Christopher. 1994. *Colloquial Estonian: The complete course for beginners*. London: Routledge.
- Nakamura, Wataru. 1999. Functional Optimality Theory: Evidence from split case systems. In Darnell, Michael, Moravcsik, Edith A., Noonan, Michael, Newmeyer, Frederick J. & Wheatley, Kathleen (eds.), *Functionalism and formalism in linguistics*, vol. 2, 253–276. Amsterdam: John Benjamins.
- Nakamura, Wataru. 2000. Psych verbs and a typology of event structure. In Sasaki, Kan & Kikusawa, Ritsuko (eds.), *Modern approaches to transitivity*, 193–219. Tokyo: Kurosio.
- Nakamura, Wataru. 2008. Fluid transitivity and generalized semantic roles. In Van Valin, Robert D., Jr. (ed.), *Investigations of the syntax–semantics–pragmatics interface*, 101–116. Amsterdam and Philadelphia: John Benjamins.
- Narasimhan, Bhuvana. 1998. A lexical semantic explanation for ‘quirky’ case marking in Hindi. *Studia Linguistica* 52(1): 48–76.
- Narrog, Heiko. 2009. Varieties of instrumental. In Malchukov, Andrej & Spencer, Andrew (eds.), *The Oxford handbook of case*, 775–785. Oxford: Oxford University Press.
- Narrog, Heiko & Ito, Shinya. 2007. Reconstructing semantic maps: The comitative-instrumental area. *Sprachtypologie und Universalienforschung* 60(4): 273–292.
- No, Naoyuki. 1992. Instruments: A case study of the interface between syntax and lexical semantics. *English Linguistics* 9: 196–222.
- Philippova, Natalia. 2017. Prepositional phrase vs. bare instrumental: The trajectory of motion in Russian. In Luraghi, Silvia & Nikitina, Tatiana & Zanchi, Chiara (eds.), *Space in diachrony*, 347–367. Amsterdam and Philadelphia: John Benjamins.

- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: The MIT Press.
- Pustejovsky, James & Batiukova, Olga. 2019. *The lexicon*. Cambridge: Cambridge University Press.
- Rissman, Lilia. 2010. Instrumental *with*, locatum *with* and the argument/adjunct distinction. *LSA Annual Meeting Extended Abstracts 2010*.
- Rissman, Lilia. 2011. Instrumental *with* and use: Modality and implicature. *Proceedings of Semantics and Linguistic Theory 21*: 532-551.
- Rissman, Lilia. 2013. Periphrastic use: A modal account of instrumentality. *Proceedings of the Annual Meeting of the North East Linguistics Society 41*: 137-150.
- Sadakane, Kumi & Koizumi, Masatoshi. 1995. On the nature of the “dative” particle *ni* in Japanese. *Linguistics 33*(1): 5-33.
- Schultze-Berndt, Eva & Himmelmann, Nikolaus P. 2004. Depictive secondary predicates in cross-linguistic perspective. *Linguistic Typology 8*(1): 59-131.
- Schwartz, Linda. 1993. On the syntactic and semantic alignment of attributive and identificational constructions. In Van Valin, Robert D., Jr. (ed.), *Advances in Role and Reference Grammar*, 433-463. Amsterdam: John Benjamins.
- Silverstein, Michael. 1980. Of nominatives and datives: Universal grammar from the bottom up. In Van Valin, Robert D., Jr. (ed.), 1993. *Advances in Role and Reference Grammar*, 465-498. Amsterdam: John Benjamins.
- Skeat, Walter W. 2005. *An etymological dictionary of the English language*. London and New York: Dover Publications.
- Stolz, Thomas & Stroh, Carmelia & Urde, Aina. 2006. *On comitatives and related categories: A typological study with special focus on the languages of Europe*. Berlin: Mouton de Gruyter.
- Ullrich, Jan. 2018. *Modification, secondary predication, and multi-verb constructions in Lakota*. Düsseldorf: Heinrich Heine University. (Doctoral dissertation.)
- Van Hooste, Koen. 2018. *Instruments and related concepts at the syntax-semantics interface*. Düsseldorf: Düsseldorf University Press.
- Van Valin, Robert D., Jr. 1991. Another look at Icelandic case marking and grammatical relations. *Natural Language and Linguistic Theory 9*(1): 145-194.
- Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.



- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2012. Lexical representation, co-composition, and linking syntax and semantics. In Pustejovsky, James, Bouillon, Pierrette, Isahara, Hitoshi, Kanzaki, Kyoko & Lee, Chungmin (eds.), *Advances in Generative Lexicon Theory*, 67-107. Dordrecht: Springer.
- Van Valin, Robert D., Jr. 2018. Some issues regarding (active) accomplishments. In Kailuweit, Rolf, Künkel, Lisann & Staudinger, Eva (eds.), *Applying and expanding Role and Reference Grammar*, 71-93. Freiburg: Albert-Ludwigs-Universität, Universitätsbibliothek.
- Wierzbicka, Anna. 1980. *The case for surface case*. Ann Arbor, MI: Karoma Publishers.

# A CROSS-LINGUISTIC SURVEY OF THE INSTRUMENT-SUBJECT ALTERNATION

KOEN VAN HOOSTE  
HEINRICH HEINE UNIVERSITY DÜSSELDORF

## Abstract

This paper aims to investigate to what extent the so-called INSTRUMENT-SUBJECT ALTERNATION (or: *ISA*) occurs in a sample of 14 languages. *ISA* is a phenomenon where a phrase that would usually be realized with instrumental marking appears as the subject, marked with the typical means associated with this function. In this paper, I briefly discuss instruments in ROLE AND REFERENCE GRAMMAR (or: *RRG*) and, more specifically, *ISA* from a more theoretical point of view after which I move to the cross-linguistic survey. By using several salient examples, I will illustrate that there is an areal dimension to *ISA*, with English as an important catalyst. This paper is to be understood as a preliminary inquiry, not a full-scale corpus study.

## Keywords

Argument marking, *RRG*, Instrument-Subject Alternation, language variation, typology

## 1. Introduction

In the literature on instruments, an alternation is sometimes described where the instrument-argument occurs as the subject of the sentence (e.g. Schlesinger 1989, Grimm 2005, 2013, Webb 2008, Van Hooste 2018). This alternation is known as the Instrument-Subject Alternation and comes with three complexities with respect to its acceptability: (1) predicate variation, (2) instrument variation and (3) language variation. Examples of *ISA* and of these three complexities are given in (1).

- (1) a. Alex cut **down** the tree with the chainsaw. *Unshifted*  
 b. The chainsaw cut **down** the tree. *ISA*  
 c. The chainsaw cut **down** the tree. *Predicate variation*  
 d. ?The chainsaw knocked **down** the tree. *Predicate variation*  
 e. The chainsaw cut **down** the tree. *Instrument variation*  
 f. \*The axe cut **down** the tree. *Instrument variation*  
 g. *Alex vel-de de boom met de kettingzaag.*  
 Alex cut **down**-PST.3SG DEF tree with DEF chainsaw  
 ‘Alex cut **down** the tree with the chainsaw.’ (Dutch) *Language variation*  
 h. \**De kettingzaag vel-de de boom.* (Dutch)  
 DEF chainsaw cut **down**-PST.3SG DEF tree  
 ‘The chainsaw cut **down** the tree.’ *Language variation*

The example in (1a) shows a standard occurrence of an instrument in English: The instrument *chainsaw* appears in a phrase marked by the preposition *with*, as this is the typical instrument marking in English. In (1b), the instrument appears in subject position. In (1c) and (1d), the instrument has been kept the same, but the predicate has been changed. These examples reveal that the acceptability of ISA depends on the predicate one wishes to apply it to. Using *knock down* in an ISA context is less acceptable than *cut down*. Varying the instrument but keeping the predicate the same reveals that the choice of referent also affects the acceptability of ISA: *chainsaw* is acceptable as the subject of an ISA-sentence, but *axe* is not. Finally, there is a language-specific component to ISA, as exemplified in (1g) and (1h). Dutch does not allow for this particular instance of ISA, contrary to English. That does not mean that Dutch prohibits ISA in general, as the example in (2) illustrates.

- (2) a. *Alex beschadig-de het wegdek met de boor.*  
 Alex **damage**-PST.3SG DEF road surface with DEF drill  
 ‘Alex **damaged** the road surface with the drill.’ (Dutch)  
 b. *De boor beschadig-de het wegdek.*  
 DEF drill **damage**-PST.3SG DEF road surface  
 ‘The drill **damaged** the road surface.’

In this paper, I will focus on cross-linguistic variation with respect to ISA. In Section 2, I will provide a concise introduction to RRG’s treatment of instruments and a schematic overview of what ISA is. In Sections 3 and 4, I will propose a solution to all three problems by relating the acceptability of ISA to the semantic properties of the referents that fill the argument slots in the LOGICAL STRUCTURES (or: *LS*). In particular, I will argue that

the arguments slots in every LS place restrictions on their fillers. These restrictions are tied to the semantic status of the referents that fill the slots. I will propose a principled way to capture the relevant semantic properties and relate these to the argument slots. In Section 5, I will present an overview of the language sample and of the examples that are at the core of this paper. In Section 6, the language material will be presented illustrating the acceptability of ISA in several languages, thereby exploring the issue of *cross-linguistic variation*. Finally, in Section 7, I will draw some conclusions with respect to the typological prevalence of ISA. The data in this paper were supplied and validated by native speakers of the respective languages.

## 2. Instruments and ISA in Role and Reference Grammar

ISA is an alternation where an argument that would typically appear as an instrumental phrase, occurs in subject position. In RRG-terms, the instrument occurs as the actor. This is summarized schematically in (3).

$$(3) \text{RP}_A \quad \text{V} \quad \text{RP}_U \quad \text{INS} \quad \rightarrow \quad \text{INS}_A \quad \text{V} \quad \text{RP}_U$$

The summary in (3) is only a rough surface-approximation of ISA. To understand the workings of ISA, one must look at the level of the logical structures. In RRG, instruments are treated as a reading of the *effector* thematic relation (Van Valin & LaPolla 1997 118-119, Van Valin & Wilkins 1996: 309ff., Van Valin 2005: 58-59), which can be defined as the dynamic participant in an event. In decompositional terms, an effector is the first argument of a *do'* (x, y) sequence. There are several potential readings of this x-argument, such as agent, force, plain effector and arguably also causee. For the present paper, only the instrument reading is of any relevance. The instrument reading sets itself apart from *agent* and *force* in that it does not occur as the first effector in a chain. That is to say, instruments only occur as intermediate effectors. This stands to reason: An instrument is conceptually manipulated by another entity to arrive at a certain result, which is reflected in the logical structures. The LS of (1a) is given in (4).

$$(4) [\text{do}' (\text{Alex}, \text{O})] \text{ CAUSE } [[\text{do}' (\text{chainsaw}, \text{O})] \text{ CAUSE } [\text{BECOME cut down}' (\text{tree})]]$$

In the LS in (4), the first, highest-ranking effector is *Alex* and is read as an agent. *Chainsaw* is a lower-ranking effector and is manipulated by the



In previous RRG-work, ISA was used as a diagnostic to distinguish instruments from implements and thus, to justify the different types of LS associated with them. In Van Hooste (2018), it is argued that this is inadequate for a simple reason: The acceptability of (1a-1b) versus the unacceptability of (1g-1h) would then suggest that there is causal embedding in English but not in Dutch. In other words, English would feature a different type of LS than Dutch in this case. As the logical structures are intended to be (quasi-)universal structures, this would be highly problematic. Rather than using ISA as a diagnostic, a paraphrase aimed at making the LSs explicit is proposed. Ergo, when the paraphrase positively identifies an instrument, it does not mean that ISA is possible. Rather, the causal embedding that the paraphrase tests for is a necessary precondition for the construction. However, the possibility to apply ISA further depends on a language-specific constraint, connected to the semantic properties of the referents relative to the LS's requirements. This will be explored in Section 3. Even though implement-LSs are equally interesting, they transcend the goals of the present paper. As such, only instances of instruments are explored in this paper.

### 3. Capturing the semantic properties of referents

In the previous section, I argued that argument slots in the logical structures are open to different kinds of fillers and that a language-specific constraint governs which arguments can fill the argument slots. Simply postulating a list of nouns that can and a list of nouns that cannot occur in these slots would be very ad hoc. Rather, it is clearly preferable to arrive at an approach where the referents of nouns are ranked along a principled system. To do this, I proposed (Van Hooste 2018) a scale consisting of two axes. The x-axis is a revised form of a more frequently used animacy hierarchy. I adopted the animacy hierarchy that was used in Van Valin & Wilkins (1996: 314-315) and elaborated on it by providing more detailed accounts of several of its echelons. It is given in (8).

(8) Non-entity < abstract entity < concrete entity (immobile) < concrete entity < animate entity < (pseudo-)sentient entity < anthropomorphic entity < 3<sup>rd</sup> person < 2<sup>nd</sup> person < 1<sup>st</sup> person

The second hierarchy is a scale that essentially captures how independently a given referent can perform an activity. Van Hooste (2018) introduced it as the autonomy hierarchy. It is given in (9).

- (9) Plain artifact < specifically tailored < Semi-autonomous < Autonomous (proper) < Group < Organization < Hive/individual human < Para-autonomous

To differentiate between different levels of AUTONOMY, Van Hooste (2018) suggested a classification of referents based on two features: [ $\pm$ independent] and [ $\pm$ controllable]. The former feature refers to whether the entity is capable of potential independent action. Decision-making abilities are typical of this class. For example, an antivirus program has to be initiated by a wielder, but carries out the activity independently. In some cases, the program must make independent decisions on how to carry out the assignment it was given by the wielder. *Controllability* refers to the feature of being controllable by a wielder. An antivirus program is controllable in that it has to be ultimately given commands by the wielder. Referents like these rank as *autonomous (proper)*. Semi-autonomous entities like chainsaws rank as [+controllable] because they have to be manipulated by a wielder as well. However, contrary to computer programs, they are also [-independent] because they cannot act – at any point during the event – without the direct intervention of a wielder. This is reflected in the fact that the control that is exerted over them is continuous throughout the event and they do not possess decision-making abilities. Storms or other meteorological phenomena are not controllable and can – due to their inherent kinetic potential – act independently. Note that this classification only classifies more complex inanimates and excludes animals and humans. Its primary purpose is to provide a more sharply delineated internal classification of machines and AI-driven referents. An overview is given in Table 1.

Table 1: Feature matrix for the principle levels of autonomy

	[+controllable]	[-controllable]
[+independent]	Autonomous	Para-autonomous
[-independent]	Semi-autonomous	<i>Irrelevant</i>

Consider by way of illustration of the larger actionality scale the difference between (10a) and (10b).

- (10) a. ???The angry mob carefully planned the assault.  
 b. The army brigade carefully planned the assault.

In (10), both *angry mob* and *army brigade* denote groups of humans, but the adverb *carefully* is much less compatible with the former, as an

*angry mob* by its very definition lacks internal cohesion and cannot – as a consequence – perform a planning act in a careful way. In Van Hooste (2018), both hierarchies were treated as axes, yielding a coordinate system. This allows for a characterization of semantic space in terms of regions rather than fixed points. Typical instruments occupy a mid-range position on the composite scale (the ACTIONALITY SCALE) and typical agents (i.e. humans) occupy higher positions. There is an interesting section in between these two clusters, which is occupied by non-human entities that can be conceptualized as instrument-effectors and as initial effectors in a causal chain. That is to say, they can be wielded by a manipulating entity and they can themselves manipulate an instrument. Van Hooste (2018) calls this class *pseudo-agents* and typical members of this class are complex mechanisms and AI-driven inanimates.

The language-specific constraint (the ACTIONALITY CONSTRAINT) governing ISA can be captured in terms of the composite scale: The instrument has to be within the correct portion of semantic space before ISA can take place. Take the examples in (1): ISA is possible with the instrument referent *chainsaw* but not with *axe*. The former occupies a higher position on the composite scale than the latter. Chainsaws have an engine, they have moving parts and when held in the correct position to a certain substance, the self-moving cutting surface will perform the cutting action. An axe, by contrast, does not have complex internal mechanics. In terms of the autonomy hierarchy, it ranks as *specifically tailored*, because it was designed to have a sharp cutting surface but it is the wielder whose force is directly translated to the tool. A *stick* would rank as a *plain artefact*. It can be used in an instrument-like manner, but it was not designed for a specific task. Dutch does not allow for ISA in (1), because the argument slots in its LS have tighter requirement than its English counterparts. The past sections provided a very brief overview of the tools proposed in Van Hooste (2018) and is an oversimplification in many respects. Rather than exploring the theoretical side of instruments and instrument-related alternations, the purpose of this paper is to ascertain the tightness of the ISA-requirement across languages: Which language has very loose requirements? Which ones simply do not allow for ISA? Where do languages put the cut-off point?

#### 4. ISA as a construction

In the literature, there have been several proposals to capture *how* ISA works. Interestingly, most of these approaches do not tackle the issue of the motivation behind ISA (Schlesinger 1989 is a notable exception). In



Van Hooste (2018), it was argued that ISA is a construction that is applied in order to make the instrument-effector the topic. In other words, its communicative goal is to obscure the instigating effector, on the account of it being irrelevant or even unknown. Apart from the language-specific constraint, ISA is governed by a further condition. For instance, to exclude implausible or non-sensical instigators, Van Hooste (2018) proposed the so-called NATURAL EVENT CONDITION (or: *NEC*), which captures the conceptual necessity for a direct causal link between an instigator and an instrument-effector. Without such a constraint, highly unlikely proposed instigators such as the one in (11b) would be possible.

- (11) a. The grenade destroyed the shed in 2015.  
 b. \*The Austro-Hungarian general destroyed the shed in 2015 with a grenade.

The example in (11b) identifies *Austro-Hungarian general* as a very unlikely conceptual instigator for (11a). This is important, as this implies that *grenade* in (11a) is not an intermediate effector, but much rather the filler of the highest-ranking slot in the LS. The NEC explains why sentences like the one in (11b) and others like it are strange (or even unacceptable) and, at the same time, puts a limit on context. The NEC is a universal constraint, whereas the actionality constraint is language-specific.

In RRG, constructions are captured in terms of constructional schemas. The advantage of such schemas is that they refer to general principles of the theory while capturing a high degree of specificity. A constructional schema for the English-ISA construction is given in Table 2.

Table 2: Constructional schema for ISA (English)

Construction: English Instrument as Actor-construction	
Syntax:	template: default PSA: Standard rules for accusative systems
Morphology:	PSA: no explicit morphology Verb agreement: Default
Semantics:	(1) x-argument of initial <i>do'</i> is unspecified (2) Actor-macrorole is assigned to highest specified x-argument of <i>do'</i> (3) highest specified x-argument of <i>do'</i> must have minimum actional status as defined by the argument position
Pragmatics:	(1) Instrument-effector is default topic (2) Natural Event Condition must be met

## 5. A cross-linguistic sample & example contexts

For the purposes of this paper, I used questionnaires and conducted interviews with native speakers of several languages. As this paper is intended as a preliminary exploratory study only, a handful of very clear examples was chosen. ‘Very clear’ in this respect refers to examples with change of state verbs, such as *destroy*, *cut down* and the like. All verbs were tested with the aktionsart tests standardly used by RRG (Van Valin & LaPolla 1997: 91ff., Van Valin 2005: 34ff.). All examples meet the NEC.

In each of the examples, the instrument-referent is raised with respect to its animacy-autonomy ranking. The goal of this change is to identify how high a referent must be before ISA becomes possible. The basic LSS underlying the example sentences can be summarized schematically as in (12).

- (12) [**do'** (x, Ø)] CAUSE [[**do'** (y, Ø)] CAUSE [INGR/BECOME pred' (z)]]

I have selected a number of languages that are all spoken in Europe. Table 3 contains an overview of the language explored for this paper, together with their genetic affiliation.

Table 3: Overview of languages investigated

Language	Genetic affiliation
Afrikaans	West-Germanic
Basque	Isolate
Bulgarian	South-Slavic
Dutch	West-Germanic
English	West-Germanic
French	Gallo-Romance
German	West-Germanic
Icelandic	North-Germanic
Irish	Celtic
Italian	Italo-Romance
Norwegian	North-Germanic
Portuguese	Ibero-Romance
Romanian	Eastern Romance
Spanish	Ibero-Romance

In the next section, I will begin by looking into the languages belonging to the Germanic family. From there, I will explore the other languages included in Table 3. Even though this paper aims at a broader typological perspective, Germanic will take a central place.

## 6. Exploring the language material

Taking an instrument whose referent ranks as specifically tailored reveals that neither Dutch, nor German, nor English allows for ISA. That is to say, the referent is too low on the actionality scale.

- (13) a. John cut **down** the tree with the axe.  
 b. \*The axe cut **down** the tree.  
 c. *Jan vel-de de boom met de bijl.*  
 Jan cut **down**-PST.3SG DEF tree with DEF axe.  
 ‘Jan cut **down** the tree with the axe.’ (Dutch)  
 d. \**De bijl vel-de de boom.*  
 DEF axe cut **down**-PST.3SG DEF tree  
 ‘The axe cut **down** the tree.’  
 e. *Jan fällt-te den Baum mit der Axt.*  
 Jan cut **down**-PST.3SG DEF tree with DEF axe.  
 ‘Jan cut **down** the tree with the axe.’ (German)  
 f. \**Die Axt fällt-te den Baum.*  
 DEF axe cut **down**-PST.3SG DEF tree  
 ‘The axe cut **down** the tree.’

However, this does not mean that specifically tailored instruments can never undergo ISA. For example, *knife* can undergo ISA in the English sentence *John cut the bread with the knife* but not in its Dutch and German equivalents. This shows that (1) English is somewhat more permissive with respect to ISA and (2) ISA is a complex interplay of factors. If the instrument-referent is higher on the actionality scale, ISA becomes possible in some languages: The referent in (1a-1b) and (1g-1f) and in (14) ranks as semi-autonomous. Yet, despite this higher status, only English allows for ISA. By contrast, German and Dutch do not.

- (14) a. *Jan fällt-te den Baum mit der Kettensäge.*  
 Jan cut **down**-PST.3SG DEF tree with DEF chainsaw.  
 ‘Jan cut **down** the tree with the chainsaw.’ (German)  
 b. \**Die Kettensäge fällt-te den Baum.*  
 DEF chainsaw cut **down**-PST.3SG DEF tree  
 ‘The axe cut **down** the tree.’

Taking other predicates reveals that German is less permissive than Dutch (15a-15f & 15g-15l). English, again, allows for ISA.

- (15) a. Sara destroyed the barn with the cannon.  
 b. The cannon destroyed the barn.  
 c. *Sara vernietig-de de schuur met het kanon.*  
 Sara destroy-PST.3SG DEF barn with DEF cannon  
 ‘Sara destroyed the barn with the canon.’ (Dutch)  
 d. *Het kanon vernietig-de de schuur.*  
 DEF cannon destroy-PST.3SG DEF barn  
 ‘The cannon destroyed the barn.’  
 e. *Sara vernichte-te die Scheune mit der Kanone.*  
 Sara destroy-PST.3SG DEF barn with DEF cannon  
 ‘Sara destroyed the barn with the canon.’ (German)  
 f. *\*Die Kanone vernichte-te die Scheune.*  
 DEF cannon destroy-PST.3SG DEF barn  
 ‘The cannon destroyed the barn.’  
 g. John killed Bill with the arrow.  
 h. The arrow killed Bill.  
 i. *Jan dood-de Bill met de pijl.* (Dutch)  
 Jan kill-PST.3SG Bill with DEF arrow  
 ‘Jan killed Bill with the arrow.’  
 j. *De pijl dood-de Bill.*  
 DEF arrow kill-PST.3SG Bill  
 ‘The arrow killed Bill.’  
 k. *Jan töte-te Bill mit dem Pfeil.* (German)  
 Jan kill-PST.3SG Bill with DEF arrow  
 ‘Jan killed Bill with the arrow.’  
 l. *\*/?Der Pfeil töte-te Bill.*  
 DEF arrow kill-PST.3SG Bill  
 ‘The arrow killed Bill.’

Judging by these (preliminary) examples, Dutch seems to be in between German and English as far as ISA’s productivity is concerned. This provides further evidence for the often-cited observation that Dutch is in between English and German and is, as it were, ‘in the middle’ (e.g. Weerman 2006). Extending the scope beyond West-Germanic, Norwegian allows ISA more or less to the same degree as Dutch. Icelandic, on the other hand, does not allow for the alternation any more than German (given in (16a-16b)). This is illustrated in (16c).

- (16) a. *Jan fel-te tre-et med en motorsag.*  
 Jan cut down- PST.3SG tree- DEF with DEF chainsaw  
 ‘Jan cut down the tree with the chainsaw.’ (Norwegian)
- b. *Motorsag-en fel-te tre-et.*  
 chainsaw- DEF cut down- PST.3SG tree- DEF  
 ‘The chainsaw cut down the tree.’
- c. *\*Keðjusög-in fell-d-i tré-ð.* (Icelandic)  
 chainsaw- DEF.NOM.SG fell- PST-3SG tree- DEF.ACC.SG  
 ‘The chainsaw cut down the tree.’

Thus far, an initial scale of acceptability for ISA can be drawn up. This is given in (17).

- (17) English > Dutch/Norwegian > Icelandic/German

As the hierarchy in (17) suggests, there does not seem to be a direct phylogenetic aspect to the variation. That is to say, the North- vs. West-Germanic divide does not seem to directly translate to differing acceptability ratings of ISA. Icelandic and German do, however, share a trait: They both have more extensive morphology than the other Germanic languages. Whether this is directly relevant for the acceptability of ISA is unclear at this point and requires a large-scale study that is beyond the scope of the present study. It is, however, intriguing that other languages with extensive morphology are not very productive (or not at all) with respect to ISA. Russian, Serbian and Lithuanian, for example, are all languages with extensive morphology and they all strongly disprefer ISA. Only referents that occupy a rather specific portion on the autonomy scale are allowed as actors. These are typically referents such as chemicals that, once administered or deployed, function entirely autonomously.

Bulgarian morphology is less extensive than that of the other Slavic languages, yet, ISA is equally impossible:

- (18) *\*Rezachka-ta otseche dŭrvo-to.* (Bulgarian)  
 Chainsaw-DEF cut down tree- DEF  
 ‘The chainsaw cut down the tree.’

The hierarchy in (17) can be updated by adding Slavic as a single entry, given in (19). Bear in mind that this hierarchy is very coarse. That is to say, it can (and must) be refined by looking at many more example contexts. It is quite likely that the Slavic languages show some differences between them. Yet, from the limited data set that is explored here, it is warranted to treat Slavic languages as a monolithic class.

## (19) English &gt; Dutch/Norwegian &gt; Icelandic/German/Slavic

Within the Romance family, French seems to be fairly restrictive concerning ISA, contrary to Portuguese. Compare (20a-b) with (20c-d).

- (20) a. *L'homme a abbatu l'arbre*  
 DEF =man AUX.3SG cut down.PTCP DEF =tree  
*avec la scie à chaîne.*  
 with DEF chainsaw  
 'The man cut down the tree with the chainsaw.' (French)
- b. \**La scie à chaîne a abattu l'arbre.*  
 DEF chainsaw AUX.3SG cut down.PTCP DEF =tree  
 'The chainsaw cut down the tree.'
- c. ● *homem abateu a árvore*  
 DEF man cut down.PST.3SG DEF tree  
*com uma motosserra.*  
 with INDEF chainsaw  
 'The man cut down the tree with the chainsaw.' (Portuguese)
- d. *A motosserra abateu a árvore.*  
 DEF chainsaw cut down.PST.3SG DEF tree  
 'The chainsaw cut down the tree.'

Even though the order of French and Portuguese relative to each other is clear, it is necessary to examine further material to determine their absolute positions on the larger hierarchy. Consider the examples in (21):

- (21) a. ● *John corta o pão com uma faca.*  
 DEF John cut.PST.3SG DEF bread with INDEF knife  
 'John cut the bread with a knife.' (Portuguese)
- b. *A faca cortou o pão.*  
 DEF knife cut.PST.3SG DEF bread  
 'The knife cut the bread.'
- c. *Jean a coupé le pain avec un couteau.*  
 Jean AUX.3SG cut.PTCP DEF bread with INDEF knife  
 'Jean cut the bread with the knife.' (French)
- d. \**Le couteau a coupé le pain.*  
 DEF knife AUX.3SG cut.PTCP DEF bread  
 'The knife cut the bread.'

Portuguese exhibits an acceptability of ISA that is similar to that of English as evidenced by the differences in grammaticality between (21a-

21b) and (21c-21d). Based on the examples above and others like it, an updated version of the ISA-hierarchy is given in (22).

- (22) English/Portuguese>Dutch/Norwegian/French>  
Icelandic/German/Slavic

Establishing the positions of English and Portuguese as similar, does not justify the position of French in the middle section of the hierarchy. The examples in (20) and (21) only provide evidence for the position of Portuguese and for French as occupying a position lower than that. By comparing French to German and Dutch, it is revealed that French shows similar behavior to Dutch (see (15) to compare). It does not allow for (20b), but it does allow for (23).

- (23) a. *Jean a tué Bill avec une flèche.* (French)  
Jean AUX.3SG kill.PTCP Bill with INDEF arrow  
'Jean killed Bill with an arrow.'  
b. *La flèche a tué Bill.*  
DEF arrow AUX.3SG kill.PTCP Bill  
'The arrow killed Bill.'

From an actionality point of view, it seems counterintuitive that a more high-ranking referent such as *chainsaw* cannot undergo ISA but a lower-ranked one, such as *arrow* can. It has been frequently observed (e.g. DeLancey 1984, Cruse 1973) that under the appropriate context, referents can be interpreted as being imbued with some form of kinetic energy. Languages frequently take recourse to this type of 'eventive construal'. For example: (24a) is ungrammatical because the subject referent's actionality value is not high enough for the slot in the LS it attempts to fill.

- (24) a. \**Die Axt zerbrach die Scheibe.* (German)  
DEF axe broke DEF window pane  
'The axe broke the window.' (Alexiadou & Schäfer 2006: 44)  
b. *Die herunterfallende Axt zerbrach die Scheibe.*  
DEF falling-down axe broke DEF window pane  
'The falling axe broke the window pane.' (Alexiadou & Schäfer 2006: 44)

Under an eventive construal in (24b), the sentence becomes grammatical. In actionality terms, the referent is raised on the scale due to the introduction of the adjective *herunterfallend* (*falling down*). In Van Hooste (2018), I proposed to treat eventive construal as a type of *feature induc-*

tion. Feature induction is also at play in (23): An arrow ranks as *specifically tailored*. It is crafted to fulfil a certain purpose but it does not carry out any instructions. Rather, it is directly manipulated. Semi-autonomous entities need to be controlled as well, but they typically carry out some function controlled by a wielder. For example, a *chainsaw* carries out part of the action (the cutting motion) itself but it does so due to instructions coming from a wielder ([+controllable] in Van Hooste (2018)). Autonomous entities, by contrast, rank as [+controllable] and [+independent]. That is, they are governed by a wielder but they can or will act ungoverned at certain times. *Arrow* on the other hand, belongs to neither class of referent. In a sentence like *Sara cut the cheese with the arrow*, the referent is non-autonomous: It must be controlled completely and continuously throughout the cutting event. No part of the arrow carries out instructions and it certainly does not make independent decisions or the like. However, in the context of being in flight due to having been fired from a bow, the referent is raised in terms of actionality. It is raised to a similar level to that of *chainsaw*. The arrow moves and kills a target as a result of an initial manipulation. The arrow is not [+independent], however, as it is not capable of independent action. It is only because of this feature induction that (23) is acceptable.

Returning to Romance, how do other languages rank on the hierarchy? Romanian, like French, does not allow (25a-25b) or (25c-25d).

- (25) a. *Bărbat-ul a tăiat copac-ul cu o drijbă.*  
 man-DEF AUX.3SG cut down.PTCP tree-DEF with INDEF chainsaw  
 ‘The man cut down the tree with the chainsaw.’ (Romanian)
- b. \**Drijba a tăiat copac-ul.*  
 chainsaw.DEF AUX.3SG cut down.PTCP tree-DEF  
 ‘The chainsaw cut down the tree.’
- c. *Ion a tăiat pâinea cu un cuțit.*  
 Ion AUX.3SG cut.PTCP bread.DEF with INDEF knife  
 ‘Ion cut the bread with a knife.’
- d. \**Cuțit-ul a tăiat pâinea.*  
 knife-DEF AUX.3SG cut down.PTCP bread.DEF  
 ‘The knife cut the bread.’

Spanish does not allow for (25c-25d) either (given in (26a-26b)). Interestingly, Spanish does not entirely reject the *chainsaw*-example, contrary to Romanian and French. However, the informant stated that as soon as a wielder is peripherally present, it becomes acceptable. This is given in (26c-26d). Italian, on the other hand, patterns like French ((26e-26h)).



- (26) a. *John cortó el pan con un cuchillo.*  
 John cut.PST.3SG DEF bread with INDEF knife  
 ‘John cut the bread with a knife.’ (Spanish)
- b. \**El cuchillo cortó el pan.*  
 DEF knife cut.PST.3SG DEF bread.  
 ‘The knife cut the bread.’
- c. *El hombre cortó el árbol con una motosierra.*  
 DEF man cut.PST.3SG DEF tree with INDEF chainsaw  
 ‘The man cut down the tree with a chainsaw.’
- d. #*La motosierra cortó el árbol.*  
 DEF chainsaw cut.PST.3SG DEF tree  
 ‘The chainsaw cut down the tree.’
- e. \**Il coltello ha tagliato il pane.* (Italian)  
 DEF knife AUX.3SG cut.PTCP DEF bread.  
 ‘The knife cut the bread.’
- f. \**La motosega ha abbattuto l’albero.*  
 DEF chainsaw AUX.3SG cut down.PTCP DEF =tree  
 ‘The chainsaw cut down the tree.’
- g. *L’arciere uccise il soldato con una freccia.*  
 DEF=archer kill.PST.3SG DEF soldier with INDEF arrow  
 ‘The archer killed the soldier with an arrow.’
- h. *La freccia uccise il soldato.*  
 DEF arrow kill.PST.3SG DEF soldier  
 ‘The arrow killed the arrow.’

Judging from the examples in (26), the ISA-hierarchy can be updated once again. To reflect the behavior from Spanish, I have inserted it between Portuguese and French:

- (27) English Portuguese > Spanish > Dutch/Norwegian French/Italian > Icelandic/German/Slavic/Romanian

Looking outside of Europe, it is most intriguing that Afrikaans patterns differently from Dutch. Despite being much more closely related to Dutch than to English (from a phylogenetic point of view), it behaves in much the same way as English with respect to ISA. This is not entirely unexpected, however, bearing in mind the extensive contact situation with English that Afrikaans is in. The Afrikaans data is given in (28).

- (28) a. *Johan het die brood met ’n mes gesny.*  
 John AUX.PST DEF bread with INDEF knife PTCP -cut  
 ‘John cut the bread with a knife.’ (Afrikaans)

- b. *Die mes het die brood ge-sny.*  
 DEF knife AUX.PST DEF bread PTCP -cut  
 ‘The knife cut the bread.’
- c. *Die man het die boom met 'n kettingsaag*  
 DEF man AUX.PST DEF tree with INDEF chainsaw  
*af<ge>saag.*  
 cut down<PTCP >  
 ‘The man cut down the tree with the chainsaw.’
- d. *Die kettingsaag het die boom af<ge>saag.*  
 DEF chainsaw AUX.PST DEF tree cut down<PTCP >  
 ‘The chainsaw cut down the tree.’

Irish is a language that has been, like Afrikaans, in close contact with English for an extended period of time. Looking at Irish provides further support ISA as an areal phenomenon, centered on the Atlantic Ocean. Just like Afrikaans, English and Portuguese, Irish and Basque allow for both examples that make up the core of this exploratory study:

- (29) a. *Ghearr John an tarán le scian.* (Irish)  
 Cut.PST Sean DET bread with knife  
 ‘John cut the bread with the knife.’
- b. *Ghearr an scian an tarán.*  
 Cut.PST DET knife DET bread  
 ‘The knife cut the bread.’
- c. *Leag an fear an crann le sábh slabhrach.*  
 Fall.PST DET man DET tree with chainsaw  
 ‘The man cut down the tree with the chainsaw.’
- d. *Leag an sábh slabhrach an crann.*  
 Fall.PST DET chainsaw DET tree  
 ‘The chainsaw cut down the tree.’
- e. *Jon-ek ogi-a labana-z ebaki zuen.*  
 John-ERG bread-DET(ABS.SG) knife(INDEF)-INS cut.PTCP AUX  
 ‘John cut the bread with the knife.’ (Basque)
- f. *Labana-k ogi-a ebaki zuen.*  
 knife(DET.SG)-ERG bread-DET(ABS.SG) cut.PTCP AUX  
 ‘The knife cut the bread.’
- g. *Gizon-a-k motozerra-z zuhaitz-a.*  
 man- DET(SG)-ERG chainsaw(INDEF)-INS tree-DET(ABS.SG)  
*moz-tu du.*  
 cut-PTCP AUX  
 ‘The man cut down the tree with a chainsaw.’

h. <i>Motozerra-k</i>	<i>zuhaitz-a</i>	<i>moz-tu.</i>
Chainsaw.DET(SG)-ERG	tree- DET(ABS.SG)	cut-PTCP
'The chainsaw cut down the tree.' <sup>1</sup>		

A final version of ISA-acceptability hierarchy is given in (30).

- (30) English Portuguese/Irish/Afrikaans/Basque>Spanish>  
Dutch Norwegian French/Italian>Icelandic/German/Slavic/Romanian

In this paper, I have used two sets of example sentences. There are essentially three possible acceptability patterns: (1) both are accepted by a given language, (2) none are accepted or (3) only one is accepted. It was shown that if only one example was judged grammatical, it is always the one containing the instrument-referent that ranks highest on the actionality scale. The hierarchy in (30) contains four echelons, rather than three, however. This is the case because Spanish seems to be more sensitive to contextual factors. One of the examples is perfectly acceptable, but the wielder must be peripherally present. In the ISA-construction, the wielder is, by the very nature of the construction, always implied. However, this implied manipulating entity can be implied to different degrees, with a previous textual mention as a stronger implication. I have summarized the results in Table 4 and I have indicated the contextual requirements of Spanish with an asterisk. The two primary example sets featured a knife and a chainsaw. I have named these sets accordingly in Table 4.

---

<sup>1</sup> If the participle *bota* (pulled down) is used, ISA becomes far less acceptable. This is essentially a manifestation of the issue of *predicate variation*.

Table 4: Cross-linguistic acceptability of ISA

Language	<i>Knife-set</i>	<i>Chainsaw-set</i>	<i>Other</i>
Afrikaans	Yes	Yes	/
Basque	Yes	Yes	/
Dutch	No	No	Yes
English	Yes	Yes	Yes
French	No	No	Yes
German	No	No	No
Icelandic	No	No	/
Irish	Yes	Yes	/
Italian	No	No	Yes
Norwegian	Yes	Yes	/
Portuguese	Yes	Yes	/
Romanian	No	No	/
Slavic	No	No	/
Spanish	No	Yes*	/

## 7. Conclusion: The typological prevalence of ISA

In this paper I have explored a limited amount of example sentences featuring ISA in a number of languages. Two main findings emerge: First, more complex mechanisms as instrument-referents are better candidates than less complex ones, which confirms Schlesinger's (1989) claim. Second, the acceptability of ISA increases westward, with the most liberal languages being situated around the Atlantic Ocean. There thus seems to be an areal dimension to ISA with English as a catalyst. For instance, whereas Dutch is more restricted, Afrikaans patterns almost identically to English. Slavic languages, by contrast, all strongly disprefer ISA. Languages like French and Dutch seem to occupy a middle position as far as ISA is concerned.

Despite offering interesting cross-linguistic insights, this study is only very preliminary. Only a handful of example contexts were investigated and only for a handful of languages. Therefore, the findings in this study have to be corroborated and refined by (1) looking at many more languages and (2) investigating many more contexts.

## Acknowledgements

This paper was made possible by the support of the CRC 991 'The Structure of Representations in Language, Cognition and Science', funded by

the Deutsche Forschungsgemeinschaft, hosted at the Heinrich Heine University Düsseldorf. I would also like to express my gratitude to Lea Kawalletz for her insightful comments on an earlier draft of this paper.

## Abbreviations

ISA	Instrument-Subject Alternation
LS	Logical Structure(s)
NEC	Natural Event Condition
RP	Reference Phrase

## References

- Alexiadou, Artemis & Schäfer, Florian. 2006. Instrument subjects are agents or causers. *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 40–48. Somerville, MA: Cascadia Press.
- Grimm, Scott. 2005. *The lattice of case and agentivity*. PhD dissertation, Universiteit van Amsterdam.
- Grimm, Scott. 2013. The bounds of subjecthood: Evidence from instruments. *Proceedings of the 33rd Meeting of the Berkeley Linguistic Society*, 1–12.
- Schlesinger, Izchak. M. 1989. Instruments as agents: on the nature of semantic relations. *Journal of Linguistics* 25: 189–210.
- Van Hooste, Koen. 2018. *Instruments and related concepts at the syntax-semantics interface*. Düsseldorf: DUP.
- Van Valin, Jr., Robert. D. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Van Valin, Jr., Robert. D., & LaPolla, Randy. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
- Van Valin, Jr., Robert. D., & Wilkins, David. P. 1996. The case for ‘Effector’: Case roles, agents and agency revisited. In M. Shibatani & S. A. Thompson (eds.), *Grammatical constructions: Their form and meaning*, 289–322. Oxford: Oxford University Press.
- Webb, James. 2008. *Instruments in LFG’s argument-structure*. University of Oxford.
- Weerman, Fred. 2006. It’s the economy, stupid! Een vergelijkende blik op men en man. In M. Hüning, U. Vogl, T. van der Wouden, & A. Verhagen (eds.), *Nederlands tussen Duits en Engels*, 19–47. Leiden: SNL.

# TWO-THEME CONSTRUCTIONS AND PREPOSITION ASSIGNMENT IN SPANISH

SERGIO IBÁÑEZ CERDA

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Abstract

This paper analyzes Spanish three-place verbs with two themes in their Logical Structure (LS) within Role and Reference Grammar [RRG]. In contrast to what happens in English, the non-undergoer theme is not always marked by the preposition *con* ‘with’, but it can also be introduced by *por* ‘for’. We show that the assignment of these prepositions is semantically guided, as both of them appear in a specific LS context. We also show that the preposition *de* ‘of’ marks the non-default projections of non-undergoer themes and in this sense covers a part of the functional domain of the preposition *with* in English. New assignment rules are proposed for these prepositions.

## Keywords

Preposition assignment, two-theme constructions, three-place verbs, Spanish prepositions, syntax-semantics interface.

## 1. Introduction

Three-place predicates have long been discussed in RRG (Van Valin and LaPolla 1997; Van Valin 2007; Kailuweit 2008; Haspelmath 2008). However, the kind of phenomena which are always analyzed in the literature are relative to either ditransitive, transfer or locative verbs, such as the ones shown in (1a), (1c) and (1e), respectively:

- (1) a. Pat gave the book to Kim.  
b. Pat gave Kim the book.  
c. Maurice presented the book to Elenor.

- d. Maurice presented Elenor with the book.
- e. Henry loaded the hay on the truck.
- f. Henry loaded the truck with the hay.

Of particular interest regarding these predicates have been the alternative constructions shown in (1b), (1d) and (1f), which imply the non-default coding of the non-subject or non-PSA (Privileged Syntactic Argument) arguments. As the current RRG analysis (Van Valin, 2007) posits, the theme argument in (1a) *the book*, (1c) *the book* and (1e) *the hay*, is the undergoer, but in (1b) it is a non-macrorole direct core argument, while in (1d) and (1f) it is an oblique core argument introduced by the preposition *with*.

Both of these kinds of constructions, the canonical and the alternative ones, are adequately handled by the RRG linking system. On the one hand, starting from the logical structure (LS) in (2), this system specifies, in the default case, that the leftmost argument (*x*) is selected as the actor and as the PSA; the rightmost argument (*z*) is selected as the undergoer, and the (*y*) argument, in the case of ditransitive and transfer verbs, is assigned dative case, through the preposition *to*, and in the case of locative verbs, such as *load* in (1e), is assigned a goal preposition, as *on* or *in*.

(2) [ $\text{do}'(x, \emptyset)$ ] CAUSE [INGR predicate' (*y, z*)]

On the other hand, in the non-default cases, the *x* argument still is the PSA, but now the *y* argument is selected as the undergoer and the *z* argument is assigned the preposition *with*, which as proposed by Van Valin and LaPolla (1997), marks all the arguments competing for a macrorole status but not chosen as one, except in the case of ditransitive verbs, whose theme is left unmarked, as in (1b).

So, these non-default constructions imply an environment where two distinct arguments are 'competing' for undergoer status. In these cases, which are the ones always discussed in the literature, the competition takes place between a theme and a goal/recipient argument; that is, the competition is between two asymmetrical participants, as in cognitive terms the theme tends to function as a figure with respect to a locative ground, which usually is instantiated by the goal or the recipient arguments. That is why the selection of the theme as undergoer is viewed as the default choice and its coding as non-macrorole argument is viewed as marked.

There are, nevertheless, other cases of three-argument, or multiple-argument, constructions where the competition for undergoer status takes

place between symmetrical participants, as is the case of the two-theme constructions exemplified in (3) for Spanish:

- (3) a. *El presidente reunió a los empresarios con los trabajadores.*  
 ‘The president gathered the businessman with the workers.’  
 a’. *El presidente reunió a los empresarios y a los trabajadores.*  
 ‘The president gathered the businessman and the workers.’  
 b. *Lucrecia comparó a Leonardo con Miguel Ángel.*  
 ‘Lucrecia compared Leonardo with Miguel Ángel.’  
 b’. *Lucrecia comparó a Leonardo y a Miguel Ángel.*  
 ‘Lucrecia compared Leonardo and Miguel Ángel.’  
 c. *John intercambió los lentes por el sombrero con George.*  
 ‘John exchanged the sunglasses for the hat with George.’  
 c’. *John y George intercambiaron los lentes y el sombrero.*  
 ‘John and George exchanged the sunglasses and the hat.’

The non-prime examples show the three-argument projection of this type of verbs. In these cases, the effector is the actor and PSA; one of the themes is the undergoer, while the third argument, another theme, is coded as an oblique complement; the non-undergoer themes in (3a) and (3b) are introduced by the preposition *con* ‘with’, and the one in (3c) by the preposition *por* ‘for’. The prime examples show that the non-PSA arguments can be coded as a single complement, through a complex reference phrase [RP] with conjoined head nouns. This proves that both these arguments have the same semantic status, i.e., they both are themes. So, the three-argument constructions projected by these verbs obligatorily imply the selection of one of the themes as the undergoer and the coding of the other one as an oblique argument, but both of the themes can equally be selected as the undergoer.

There are two very interesting things about this kind of construction in Spanish: first, they exemplify cases of three argument predicates where the non-default preposition assignment rule for the non-PSA arguments seems to be obligatory, both in semantic and syntactic terms; that is, these constructions inherently imply an environment of competition between two theme arguments, so the competing rule, or instrumental case assignment rule, in RRG terms, is the default rule. No dative or locative prepositions are assigned in any case.

And second, the non-undergoer theme can be introduced by the preposition *con*, which in this respect seems to cover similar ground to that covered by *with* in English, but at the same time can be introduced by other prepositions, such as *por* ‘for’ in (3c), showing that the application



of this assignment rule is a little more complicated, and suggesting that the rule is not exclusively motivated by syntax. Van Valin (2007) has already pointed out that with removal and dispossession verbs the non-default coding of the theme argument results in an oblique argument introduced by the preposition *of*. The same holds true for similar verbs in Spanish, when the alternation is possible, as it is the preposition *de* ‘of’ and not *con* the one that appears in those contexts. But the more intriguing thing is that the three prepositions *con*, *de* and *por* can alternate in contexts where the competition rule applies, both in the case where the competition is between two themes (4a) and (4c), and between a theme and a locative argument (4b), but at the same time this alternation is restricted, as shown in the examples of (4):

- (4) a. *Clemente substituyó el vino por/con/\*de la cerveza.*  
 ‘Clemente substituted the wine for the beer.’  
 b. *María cargo el camion de/con/\*por manzanas.*  
 ‘María loaded the truck with apples.’  
 c. *Evaristo intercambió sus lentes por/\*con/\*de un libro con Amelia.*  
 ‘Evaristo exchanged his sunglasses for a book with Amelia.’

This suggests that the application of the non-default preposition assignment rule has some semantic basis, in the sense that it interacts with the semantic type of the predicates, allowing the appearance of one or two of these prepositions with certain predicates and one or two with others, in different combinations.

In summary, this paper has the goal of exploring further the two points sketched above with data from Spanish, that is: 1) to determine what preposition is assigned in the case of the several types of three-place predicates; and 2) to try to explain the semantic or syntactic motivation for the appearance of these prepositions.

In order to achieve this, in section 2 we introduce the current RRG approach to the preposition assignment of oblique core arguments; then, in section 3 we proceed to identify and analyze what the types of predicates are that have two themes in their LS. In section 3.1 we deal with the case of verbs of ‘putting together’; in 3.2. we analyze the verbs of ‘putting into a relation’; in 3.3 we look at the verbs of ‘exchange one entity for another’; verbs of ‘substitution’ are dealt with in 3.4; and in 3.5 we bring the verbs of ‘commercial event’ into the discussion. Finally, in section 4 we offer some conclusions.

## 2. The current RRG approach to preposition assignment in three-place constructions

In RRG traditional grammatical relations like subject and direct object have no theoretical status; rather, RRG posits a single, construction-specific, grammatical relation: the privileged syntactic argument (PSA), which in most English and Spanish constructions, but not all, is equivalent to the traditional subject. Non-PSA arguments are referred to, in terms of their morpho-syntactic properties, as direct or oblique arguments. The selection hierarchy and the principles for determine the PSA selection are given in (5) and (6), respectively:

(5) Privileged syntactic argument hierarchy

Arg of D > 1st arg of do' > 1st arg of pred' (x, y) >  
2nd arg of pred' (x, y) > arg of pred' (x)

(6) Accessibility to privileged syntactic argument principles:

- a. Accusative constructions: Highest ranking direct core argument in terms of (5) [default].
- b. Ergative constructions: Lowest ranking direct core argument in terms of (5) [default].

The hierarchy in (5) and the principles in (6) capture the fact that in languages like English and Spanish in an active voice clause with a transitive verb the actor is the PSA (6a), whereas in ergative languages like Dyirbal, in the same type of clauses, the undergoer is the PSA (6b).

Another important part of the linking system involves case and adposition assignment rules, which govern the morphosyntactic realization of arguments. The basic rules for direct core arguments in accusative languages, such as the ones discussed here, are given in (7):

(7) Case marking rules for accusative languages:

- a. Highest ranking core macrorole takes nominative case.
- b. Other core macrorole takes accusative case.

Non-macrorole arguments are normally oblique, and in languages that lack morphological case they are marked by adpositions. These adpositions are not idiosyncratically listed in the lexical entries of verbs, but rather, they are assigned by systematic rules; some of them have a semantic basis, i.e., they operate in specific verbal LS contexts. This approach to

adposition assignment was first presented in Foley and Van Valin (1984) and further developed in Jolly (1993) and Van Valin and LaPolla (1997).

Examples of adposition assignment rules that operate on a semantic basis are those of the prepositions *to* and *from* in English:

- (8) a. Assign *to* to non-MR *y* argument in logical structure segment: ...  
       BECOME/INGR **pred'** (*y*, *z*)  
 b. Assign *from* to non-MR *y* argument in logical structure segment: ...  
       BECOME/INGR **NOT pred'** (*y*, *z*)

The preposition *to* appears in contexts like the following:

- (9) a. Sally gave/sent/handed the box to Pat.  
 b. Sally showed the box to Pat.  
 c. Sally taught basketweaving to Pat.

All these examples have in common the general LS in (2), presented here as (10):

- (10) [**do'** (*x*,  $\emptyset$ )] CAUSE [INGR **predicate'** (*y*, *z*)]

What is different in their LS, is the predicate embedded under the INGRESSIVE operator:

- (11) a. ... INGRESSIVE **have'** (Pat, box) = *give, hand, send*  
 b. ... INGRESSIVE **see'** (Pat, box) = *show*  
 c. ... INGRESSIVE **know'** (Pat, basketweaving) = *teach*

The linking system specifies, in the default case, that the leftmost argument *x* is selected as the actor and the PSA; the rightmost argument *z* is selected as the undergoer, and the *y* argument, a non-macrorole argument, is assigned the preposition *to*, following rule (8a). Note that the state predicate embedded under a BECOME operator can be a possession, a perception or a cognition predicate, as well as a locative predicate, so the argument marked by *to* can be a possessor, a perceiver, a cognizer or a location; in this sense, *to* marks different types of semantic arguments, but it always appears in the same kind of LS environment, precisely the one foreseen in the rule (8a): BECOME/INGR **pred'** (*y*, *z*).

A similar analysis can be posited for *from*, which appears in examples like the following:

- (12) a. Sandy took/stole/bought the keys from Kim.  
 b. Pat drained the water from the pool.  
 c. Kim escaped from the burning house.

These examples have these content segments in their LS:

- (13) a. .... INGRESSIVE NOT **have'** (Kim, keys) = *take, steal, buy*  
 b. .... INGRESSIVE NOT **be-in'** (pool, water) = *drain*  
 c. .... INGRESSIVE NOT **be-in'** (burning house, Kim) = *escape*

In each of these LSs, *from* marks the first argument of the two-place state predicate, which is a non-macrorole core argument. The difference between these LSs and those in (11) is the presence of NOT. This difference in content between *to* and *from* was first proposed by Gruber (1965). Again, like *to*, *from* does not mark a single thematic relation, but rather it is assigned in the particular LS context BECOME/INGR NOT **pred'** (y, z), so the rule (8b) rightly predicts its appearance.

The best example of a preposition assignment rule that does not have a semantic basis is that of the preposition *with* in English, which is given in (14):

- (14) Assign *with* to non-MR *b* argument if, given two arguments, *a* and *b*, in a logical structure, with (1) both as possible candidates for a particular macrorole and (2) *a* is equal or higher (to the left of *b*) on the AUH, *b* is not selected as that macrorole.

The rule in (14) applies in the following single prime examples:

- (15) a. Sally presented the flowers [z] to Kim [y].  
 a'. Sally presented Kim [y] with the flowers [z].  
 a''. [**do'** (Sally, Ø)] CAUSE [INGR **have'** (Kim, flowers)]  
 b. Max loaded the olives [z] into his minivan [y].  
 b'. Max loaded his minivan [y] with the olives [z].  
 b''. [**do'** (Max, Ø)] CAUSE [INGR **be-in'** (minivan, olives)]

These examples, (15a') and (15b'), imply a marked assignment, given that the theme argument *z* has not been selected as undergoer; instead, it is marked by *with*. On the other hand, the non-prime examples obey rule (8a), which predicts the appearance of *to*. So, *with* is assigned just in contexts where two arguments can be selected for undergoer status, but where the marked option in terms of (8a) is the one that prevails.

There are two other main contexts where *with* is assigned: for marking an instrument (16a) and for introducing a comitative (16b). (See also Nakamura, this volume.)

- (16) a. Tom cut the bread with the knife.  
 a'. The knife cut the bread.  
 a''. [ $\text{do}'$  (Tom, [ $\text{use}'$  (Tom, knife)])] CAUSE [[ $\text{do}'$  (knife, bread)] CAUSE [BECOME cut' (bread)]]  
 b. Sandy went to the store with Kim.  
 b'. Sandy and Kim went to the store.  
 b''. [ $\text{do}'$  (Sandy, [ $\text{go}'$  (Sandy/Kim)])] & [INGR be-at' (store, Sandy/Kim)]

Both these arguments can also have the actor macrorole. In (16a), following the AUH, it is the leftmost argument, the effector, *Tom*, the one which is selected as actor, and as such it is projected as subject; then, the other argument, the instrument, *the knife*, is marked by the preposition *with*; but it is also the case that when the effector is not projected, the instrument can be selected as actor and then projected in the subject function, as in (16a'). (See Van Hooste, this volume.) A similar situation happens with the comitative, or co-effector argument. The LS in (16b'') has two co-effector referents, *Sandy and Kim*, and each one can be selected as actor and as subject; in (16b) *Sandy* is projected as subject and, applying the rule (14), *Kim* is marked by the preposition *with*; but in (16b') both referents are projected as subject, through a single complex RP. Hence, English *with* effectively appears in contexts where two arguments compete for macrorole status, be it the actor or the undergoer, and one is not selected, precisely the context predicted by the rule (14).

### 3. Two-theme verbs in Spanish

As mentioned before, besides the three-place constructions exemplified in (1) and in (15), in Spanish, there are several other types of three-argument predicates which do not follow the preposition-assignment rules provided in (8) and (14). These predicates, besides an effector-cause argument, have in their LS two additional arguments that are not in an asymmetrical relation, like themes are in relation to locatives and recipients, because those two arguments have the same semantic identity, i.e., they both are themes. In what follows, we identify and analyze these types of predicates.

### 3.1. Verbs of ‘putting together’

First, there are those causative predicates of change of location which inherently imply ‘putting two entities together in the same locative space’, as *juntar* ‘join’ or ‘put together’, *unir* ‘unite’ or ‘join together’, *reunir* ‘gather together’, *acoplar* ‘fit together’ or ‘couple’, *mezclar* ‘mix (up)’, *combinar* ‘combine’ or ‘match’. Some examples are provided in (17):

- (17) a. *Ramiro reunió a Simón con Elenor en el salón principal.*  
 ‘Ramiro got Simon together with Elenor in the main hall.’  
 b. *El gobierno unió a los trabajadores con los patrones en la misma agrupación.*  
 ‘The government united the workers with the bosses in the same association.’  
 c. *Roberta mezcló los pepinos con las zanahorias en la ensalada.*  
 ‘Roberta mixed up the cucumbers with the carrots in the salad.’

As can be seen, this kind of verbs can be constructed with four arguments: a cause-effector, two themes and a goal locative. Its semantics is inherently locative. It implies that the two themes end up sharing the same space, or at least, that they end up being together. Tentative LSs for these predicates are shown in (18):

- (18) a. [ $\text{do}'(x, \emptyset)$ ] CAUSE [[ $\text{do}'(y \wedge z, \emptyset)$ ] & [INGR **be-with'** (y, z)  $\wedge$  **be-in'** (w, (y  $\wedge$  z))]]  
 b. [ $\text{do}'(x, \emptyset)$ ] CAUSE [[ $\text{do}'(y \wedge z, \emptyset)$ ] & [INGR **be-with'** (z, y)]]  
 c. [ $\text{do}'(x, \emptyset)$ ] CAUSE [INGR **be-with'** (y, z)  $\wedge$  **be-in'** (w, (y  $\wedge$  z))]

(18a) is the LS for cases where the goal is coded or semantically implied. In a similar fashion to what Farrell (2009) proposes for the English *with*, we here also propose that the Spanish preposition *con* have a basic predicative meaning, comitative, that of ‘two entities being together’, as in *Juan está con María* ‘John is with Mary’, which can be formalized as a **be-with'** (z, y) predicate.<sup>1</sup> When this predicate is linked to a locative **be-in'** by the presence of a conjunction  $\wedge$  (‘and simultaneously’), like in the LSs above, the meaning of ‘being together in the same place’ is captured. (18b) accounts for the uses where no goal is implied. The second [ $\text{do}'(y \wedge$

<sup>1</sup> Here we follow Van Valin (personal communication), who in turn follows John Payne (personal communication 2004), in stating that the basic predicative meaning of *with* is comitative; so, we differ from Farrell (2009), who proposes that *with* is a kind of locative predicate. See also Nakamura (this volume).

$z, O$ ] predicate in (18a) and (18b) is postulated inasmuch as these verbs tend to appear, in actual use, in anticausative constructions, where both themes, when animate, have a self-mover interpretation as in *María se reunió con Sofía* ‘Mary got together with Sofía’. When both themes are inanimate, the resultant construction is a stative one, where only the [be-with’ ( $z, y$ )] predicate is projected, as in *El color negro se mezcla bien con el color blanco* ‘Black color mixes up well with the white color’. (18c) accounts for causative (or anticausative) cases with two inanimate themes and a goal.

In the examples in (17), the effector-cause argument is the actor and PSA; the goal is coded through the preposition *en*, as are all goals in Spanish in the context of an INGRESSIVE operator. As for the themes, one is selected as the undergoer and the other one appears coded as an oblique argument introduced by the preposition *con* ‘with’, which is the expected one in terms of the RRG linking algorithm, which specifies that a potential undergoer not selected as such has to be marked by that preposition. Nevertheless, as mentioned before, in Spanish, not all the ‘competing’ arguments for the undergoer role not selected as such are marked by *con*. They also can be marked by *por* and *de*. So, I rather propose that the appearance of *con* in the projection of these verbs has to do with the presence of the predicate be-with’ ( $z, y$ ) in their LS; with this semantic segment as part of the LSS, *con* is assigned automatically.

### 3.2. Verbs of ‘putting into a relation’

Second, there are predicates that are very closely related to the ones in the previous section. These are verbs like *relacionar* ‘relate’ or ‘put into a relation’, *conectar* ‘connect’, *contactar* ‘put in contact’, *asociar* ‘associate’ or ‘link’, *vincular* ‘bind together’ or ‘link’ and *comparar* ‘compare’, among others. Some examples are shown in (19):

- (19) a. *Cipriano conectó a Florentino con Fabián (en el proyecto).*  
 ‘Cipriano connected Florentino with Fabián (in the Project).’  
 b. *Rodolfo contactó a Bernardo con Laura (para un negocio).*  
 ‘Rodolfo contacted Bernardo with Laura (in order to do business [with her]).’  
 c. *Betty comparó la película con el libro.*  
 ‘Betty compared the film to the book.’

These verbs denote a similar process of putting together two entities, but this time, not into a physical space, but into an abstract space, that is,

the ‘generic space of a relation’. This signifies, as can be seen from the examples in (19), that the locative argument is less implied and can be completely excluded, both in semantic and syntactic terms. In (20), we posit similar LSs as the ones postulated above:

- (20) a. [ $\text{do}'(x, O)$ ] CAUSE [[ $\text{do}'(y \wedge z, O)$ ] & [INGR **be-with'** (y, z)  $\wedge$  **be-in'** (w, (y  $\wedge$  z))]]  
 b. [ $\text{do}'(x, O)$ ] CAUSE [[ $\text{do}'(y \wedge z, O)$ ] & [INGR **be-with'** (z, y)]]  
 c. [ $\text{do}'(x, O)$ ] CAUSE [INGR **be-with'** (y, z)]

Just the same, in the syntactic projection of these verbs, one of the themes is selected as the undergoer and the other one is projected as an oblique argument introduced by the preposition *con* ‘with’. What these verbs have in common with the verbs of ‘putting together’, besides the locative value, is a LS with an embedded **be-with'** (y, z) predicate, where the two arguments are themes. This is what, in structural terms, is behind the appearance of *con* ‘with’ as the preposition that introduces the theme coded as an oblique argument.

### 3.3. Verbs of ‘exchange one entity for another’

Then we have verbs like *intercambiar* ‘exchange’, *permutar* ‘permute’ or ‘swap’, *canjear* ‘exchange’ or ‘trade’ and *trocar* ‘exchange’ or ‘trade’, which are inherently reciprocal. They imply a complex structure where two acts of giving take place at the same time. In this sense, they have four arguments: two givers, which simultaneously are recipients, and two themes, each of them ending in the possession of each one of the recipients. Examples are shown in (21):

- (21) a. *Evaristo intercambió sus lentes por/\*con un libro con Amelia.*  
 ‘Evaristo exchanged his sunglasses for a book with Amelia.’  
 b. *Dolores canjeó dólares por/\*con euros con Anastasia.*  
 ‘Dolores traded dollars for euros with Anastasia.’

Tentative LSs for these verbs are shown in (22):

- (22) a. [[ $\text{do}'(x, O)$ ] CAUSE [[INGR **have'** (z, y)] & [INGR N●T **have'** (x, y)]]]  $\wedge$  [[ $\text{do}'(z, O)$ ] CAUSE [[INGR **have'** (x, w)] & [INGR N●T **have'** (z, w)]]]



- b. [[do' (x, O)] CAUSE [[INGR have' (z, y)] & [INGRN●T have' (x, y)]] IN-EXCHANGE-F●R [[do' (z, O)] CAUSE [[INGR have' (x, w)] & [INGRN●T have' (z, w)]]]

(22a) joins together two standard LSs for the verb *give*. Another proposal is the LS in (22b), which additionally has the operator IN-EXCHANGE-F●R that makes explicit that one sub-event of giving takes place, because the other one is also taking place.

As can be seen from the examples in (21), in the projection of these verbs one of the givers is selected as actor and as PSA; the other one is coded as an oblique argument introduced by the preposition *con* 'with', as it is also expected for co-effectors (Jolly 1993), following the rule in (14); one of the themes shows up as the undergoer and the other one appears as an oblique argument introduced, interestingly, not by *con*, but by the preposition *por* 'for', showing that the linking does not take place as expected when two arguments are also competing for the undergoer function. Several factors can be taken into account to explain this fact: first and most important, the two themes are not arguments of the same stative predicate, as the LS is composed of two equal structures of giving. Second, the themes don't end up located in the same locative space, but in the hands of different recipients, so the meaning of 'togetherness' is absent. Third, the meaning that is the structural base of these verbs is justly that of 'one for one' or that of 'in exchange for/of' ('one event for the other event'), which is one of the basic predicative meanings of the preposition *por*. This preposition is the one that has to be assigned, because it matches the structural meaning of these predicates.

### 3.4. Verbs of 'substitution'

Next, we have what we are calling the verbs of 'substitution'. These are verbs such as *sustituir* 'substitute', *suplantar* 'supplant', *reemplazar* 'replace', *suplir* 'replace', *relevar* 'substitute' and *alternar* 'alternate', among others. In (23) we have examples of clauses with these predicates:

- (23) a. *El ingeniero sustituyó la alberca por/con una jardinera (en el proyecto).*  
 'The Engineer substituted the swimming pool for a plant stand (in the design).'
- b. *El entrenador reemplazó a Maradona por/con Messi (en el campo de juego).*  
 'The coach replaced Maradona with Messi (on the football field).'

These verbs are similar to those of ‘interchange’ in the sense that they also have a complex LS, but in this case, the LS is composed of two different sub-events, one of ‘removing’ or of ‘taking something from’ and one of ‘putting’. See their tentative LS in (24):

- (24) a. [**do**’ (x, O)] CAUSE [[INGR N**OT** **be-in**’ (w, y)] &  
 [INGR **be-in**’ (w, z)]  
 b. [[**do**’ (x, O)] CAUSE [[INGR N**OT** **be-in**’ (w, y)]  
 IN-EX.CHANGE-F**OR** [INGR **be-in**’ (w, z)]]

As can be seen in the examples in (23), four arguments can be coded in the projection of this kind of verb. In a similar fashion to what happens with the other predicates under study, the first argument of **do**’ is the actor and PSA, one theme is selected as the undergoer and the other one is coded as an oblique complement introduced, in the unmarked case, by the preposition *por* ‘for’. Finally, the goal locative also appears as a prepositional complement (introduced by *en*). Again, we have a structural environment where two themes are competing for the undergoer role and, again, another preposition besides *con* is assigned as the unmarked choice. What are the reasons for this behaviour? Several factors are called into play: first, we have a complex structure where each of the themes belongs to a different event structure: one of them is the argument of the ‘removal’ event, and the other one is the argument of the ‘putting’ event. In fact, these two themes are not reciprocal, as are the ones of the ‘exchange’ verbs, their referents cannot be conjoined in a complex RP: *El entrenador sustituyó a Maradona y a Messi* ‘The coach substituted Maradona and Mesi’ is not a proposition equivalent to the one in the example (23b). More important, the two themes never share the same locative space, so, again, the meaning of ‘togetherness’ is not present. Additionally, the LS implies that the two sub-events depend reciprocally on each other (‘one for the other’), so that is why we posit the IN-EXCHANGE-F**OR** operator as part of it, which matches the semantic meaning of the preposition *por*.

Nevertheless, as can also be seen in the examples of (23), the preposition *con* ‘with’ can also appear marking the theme not selected as undergoer, but although this is a completely grammatical option, it is much less frequent in usage data. So, it is really the marked case. Now, there are two possible explanations for this appearance of *con*: a) It can be the case that this preposition has been (diachronically) extending its uses and is now competing with *por*, which could be possible because in these contexts both mark a theme not selected as undergoer; or b) it could be the case that the presence of *con* is an instance of another predicative use, an instrumen-

tal one, related to a meaning that could be something as ‘using something for’; if this would be the case, an example like (23b) would be paraphrased as *El entrenador utilizó a Messi para sustituir a Maradona* ‘The coach used Messi to substitute Maradona’. As this is out of the scope of this paper, we are leaving these two hypothesis open for further investigation.

### 3.5. Verbs of ‘commercial event’

The last class of predicates with multiple arguments we are addressing in this work is that of ‘commercial event’ verbs. Among these are *comprar* ‘buy’, *vender* ‘sell’, *pagar* ‘pay’ and *cobrar* ‘charge’. These verbs also have a complex LS where, at least, two events of giving take place reciprocally and simultaneously: one where the buyer gives money to the seller in exchange of some goods; and one in which the seller gives the goods to the buyer in exchange of some money. So, they have four arguments and in structural terms both the goods and the money are themes. Some examples are shown in (25):

- (25) a. *Juan compró un libro por 30 pesos.*  
 ‘John bought a book for 30 pesos.’  
 b. *Juan compró 30 pesos de libros.*  
 ‘John bought a book for 30 pesos.’  
 c. *Juan vendió un libro por 30 pesos.*  
 ‘John sold a book for 30 pesos.’  
 d. *Juan vendió 30 pesos de libros.*  
 ‘John bought a book for 30 pesos.’  
 e. *Juan pagó 30 pesos por los libros.*  
 ‘John paid 30 pesos for the books.’  
 f. *Juan pagó un libro en 30 pesos.*  
 ‘John paid a book in 30 pesos.’  
 g. *Juan cobró 30 pesos por un libro.*  
 ‘John charged 30 pesos for a book.’  
 h. *Juan cobró un libro en 30 pesos.*  
 ‘John charged 30 pesos for a book.’

As posited in Ibáñez (2018), these verbs have a shared complex LS that accounts for the fact that all the four arguments can be, in different constructions, selected as actor or as undergoer, or coded as an oblique argument. A portion of this shared LS that is equivalent to the meaning of *pagar* ‘to pay’ is shown in (26):

- (26) [**do'** (z, Ø)] CAUSE [[INGR **have'** (x, [**money'** (y))]]  $\wedge$   
 [[INGR **have'** (z, [**goods'** (w))]]  $\wedge$  [INGR **NOT have'** (z, [**money'**  
 (y))]]  $\wedge$  [INGR **NOT have'** (x, [**goods'**(w))]]

As can be seen from the examples in (25), all these verbs have, besides other interesting constructions, two kinds of basic projections: one in which the goods are selected as undergoer, as in (25a) for *comprar* 'to buy', and one in which the money is selected as such, as in (25b) for the same verb. In the first of these projections, which is the unmarked one, the non-undergoer theme is coded as an oblique argument introduced by the preposition *por*, as in examples (25a) (25c) (25e) and (25g). In the second type of construction, the non-undergoer theme is introduced by other prepositions, *de* 'of' in the case of *comprar* 'buy' and *vender* 'sell', and *en* 'in/on/at', in the case of *pagar* 'pay' and *cobrar* 'charge'. This behaviour is explained by the fact that although both the goods and the money are themes, they are not reciprocal, they cannot be conjoined in single RP, and they are not completely equivalent inasmuch as *vender* and *comprar* have an inherent prominence on the goods argument, while *pagar* and *cobrar* focus on the money argument; so, in structural terms each pair of verbs has a preference for the syntactic projection of one of these themes as the undergoer. The result is that effectively the relation of these two types of constructions is similar to that of the transfer and change-of-location verbs, which have one unmarked projection with the theme as undergoer and the goal/recipient as an oblique argument, and one marked construction with the locative as undergoer and the theme as an oblique argument, but in the case of the commercial event verbs these two projections are characterized by the presence of two themes: as said above, *vender* and *comprar* have an inherent semantic and syntactic prominence on the goods argument, while *pagar* and *cobrar* focus on the money argument, and this produces an asymmetry between the two theme arguments.

Again, for the same reason as with the 'exchange' and 'substitute' verbs, the preposition assigned in the unmarked construction is *por* 'for' and not *con* 'with': the predicates have a complex LS, the themes are not reciprocal and they never share the same space, as they end up in different hands, so, again, the meaning of 'togetherness' is not present. And all the same, the LS implies that the two sub-events depend reciprocally on each other, in the sense that one takes place 'in exchange of' the other, value or feature which matches the semantic meaning of the preposition *por*.

So, we now have seen that there are three classes of predicates, those we identify as 'exchange one entity for another', 'substitution' and 'commercial event' verbs, which in their syntactic projection have one theme

argument marked by *por* ‘for’, instead of *con* ‘with’. In all these cases *por* behaves like a non-predicative preposition: it is automatically assigned in a specific structural context.

#### 4. The competition between *de* and *con* in the non-default coding of oblique core argument

In terms of what we have presented so far, we can say that the assignment of prepositions with two-theme predicates, be it *con* or *por*, is systematic and structural, and that is ruled by structural and semantic factors, namely, the occurrence of LS segments that match the meaning content of those prepositions. The assignment rules can be formulated as in (27):

- (27) a. Assign *con* to non-MR *y* argument in logical structure segment: ...  
 BECOME/INGR **be-with**’ (z, y)  
 b. Assign *por* to non-MR *y* argument in logical structure segment: ...  
 BECOME/INGR **NOT pred**’ (z, y), part of a complex LS which includes the IN-EXCHANGE-FOR operator.

These are rules that apply in the default linking. As for the prepositions which mark the oblique argument in the marked constructions of the ‘commercial event’ verbs, such as those in (25b), (25d), (25f) and (25h) above, in which the less prominent theme, in lexical terms, is selected as undergoer, it can be *de* ‘of’ or *en* ‘at/in/on’<sup>2</sup>. As proposed in Ibáñez (2019), *de* is certainly expected, inasmuch as it is, and not *con* ‘with’, the preposition that is more commonly found marking the non-default linking with predicates which does not have two themes and that have an asymmetrical relation between their arguments, as the case of communication, removal, locative and spray verbs, as shown in the examples of (28):

- (28) a. *La oficina de prensa informó la noticia al público.*  
 ‘The press office informed the notice to the public.’  
 a’. *La oficina informó al público ~~de~~?con/\*por la noticia.*  
 ‘The press office informed the public the notice.’  
 b. *Elia le perdonó su mala educación a Marcelo.*  
 ‘Elia forgave his bad manners to Marcelo.’

<sup>2</sup> For reasons of space, in this work we will not be dealing with the occurrence of preposition *en* in this kind of context.

- b'. *Elia perdonó a Marcelo de/por/\*con su mala educación.*  
 'Elia forgave Marcelo his bad manners.'
- c. *Ramón denunció el robo de Martha.*  
 'Ramón denounced Martha's robbery.'
- c'. *Ramón denunció a Martha de/por/\*con robo.*  
 'Ramón denounced Martha for the robbery.'
- d. *Ludmila vació todo el líquido de la botella.*  
 'Ludmila emptied all the liquid from the bottle.'
- d'. *Ludmila vació la botella de/\*con/\*por todo su líquido.*  
 'Ludmila emptied the bottle of all its liquid.'
- e. *Leonardo cargó las manzanas en el camion.*  
 'Leonardo loaded the apples into the truck.'
- e'. *Leonardo cargó el camion de/con/\*por manzanas.*  
 'Leonardo loaded the truck with the apples.'
- f. *Fernando roció cerveza en todo el cuarto.*  
 'Fernando sprayed beer all over the room.'
- f'. *Fernando roció el cuarto de/con/\*por cerveza.*  
 'Fernando sprayed the room with beer.'

As can be seen from the prime examples in (28), *de* 'of' is the preposition that, in effect, can appear in all contexts of non-default projection of a theme, when the competition for the undergoer function is between asymmetrical arguments. *Con* 'with' and *por* 'for' have a more reduced area of operation<sup>3</sup>. Besides, *de* is also the preposition that appears marking the non-macrorole argument of the majority of two-place M-intransitive verbs in Spanish, those as *carecer* 'lack', *disfrutar* 'enjoy', *padecer* 'suffer', *cuidar* 'look after' or 'take care of', *necesitar* 'need', *sospechar* 'suspect', *acordarse* 'remember' or 'recall', *constar* 'consist of', *disponer* 'have' or 'make use of', etc., that in accordance to Ibáñez (2019) have a lexical specification for selecting their first argument as undergoer, and as PSA, which is a marked choice in terms of the linking system, as they are stative predicates.

One way of approaching these facts is to consider *de* as the basic preposition for marking non-prototypical argument projections. Following Ibáñez (2019), we can posit a rule like (29) for its assignment to verbal arguments:

<sup>3</sup> The instances of *por* in the examples (28b') and (28c') are probably predicative uses related to the cause meaning of 'because of', but establishing this properly is beyond the scope of this work.

(29) Assign *de* to non-macrorole arguments in non-default linkings

Still, the preposition *con* appears in contexts where *de* cannot: a) marking the co-effector not chosen as actor in comitative constructions, as in *Lola fue al cine con/\*de Domingo* ‘Lola went to the movies with Domingo’; b) marking the effector not chosen as actor, i.e., the instrument, of ‘cut’ and ‘brake’ verbs, as in *Juan cortó el pan con/\*de el cuchillo* ‘John cut the bread with the knife’. That is, *con* has a particular niche of projection in contexts that suppose a competition of two arguments for the actor macrorole. For these cases, a similar rule to that of the English *with* rule still can be proposed:

- (30) Assign *con* to non-MR *b* argument if, given two arguments, *a* and *b*, in a logical structure, with (1) both as possible candidates for the actor macrorole and (2) *a* is equal or higher (to the left of *b*) on the AUH, *b* is not selected as that macrorole.

This rule would be complementary to a slightly modified version of (29):

- (31) Assign *de* to non-macrorole arguments in non-default **undergoer** selections.

So, there seems to be, at least, two different *con* prepositions in Spanish, in a similar fashion to what is proposed by Nakamura (this volume) for the English *with*: a) one comitative, which have a basic predicative meaning (‘be-with’) and which can also be assigned in structural contexts that have a **be-with** segment as part of it; and b) an instrumental one, which is assigned in the context of LSs where two arguments are in competition for the actor macrorole.

The appearance of *con* in the same contexts as *de*, as in the examples of (28e) and (28f) –locative and spray verbs, respectively– can be explained in two ways: a) both prepositions, *con* and *de*, are non-predicative prepositions equally apt for marking the non-default coding of the theme of some three-place predicates, and in this sense they are in a synchronic and diachronic competition for gaining more structural scope; or b) *de* is the basic and unmarked choice for non-default undergoer selections and the presence of *con* supposes an additional predicative meaning, probably related to the instrumental one, as we suggested in section 3.4., and as has also been proposed by Farrell (2009). We leave this issue for further investigation.

## 5. Conclusions

Besides the typical transfer and locative verbs, there are additional classes of three-place predicates. In particular, in this paper we have been focusing on the analysis of those Spanish verbs that have two themes in their LS, as the verbs of ‘putting together’, the verbs of ‘putting into a relation’, the ‘exchange’ and ‘substitute’ verbs, and the ones that denote a commercial event. All these predicates involve an inherent context of competition between the two themes for the undergoer role. Contrary to what could be expected from the current RRG algorithm in languages that have an inventory of prepositions similar to that of English, the non-undergoer theme is not always coded introduced by the instrumental preposition *con* ‘with’, but can also be introduced by the preposition *por* ‘for’. As shown throughout this paper, the appearance of each one of these two prepositions is guided by the semantic content in the LS of the different verbs. In this sense, the preposition *con* in Spanish seems not to completely lack semantic content, as it has been proposed in RRG for the English *with* (Van Valin and LaPolla 1997). A basic sense of ‘togetherness’ can be posited for it and can be captured by means of a predicate **be-with** (x, y). This sense effectively matches a part of the LS of the verbs of ‘putting together’ and those of ‘putting into a relation’, which is what is rightly behind the use of *con* to mark the oblique argument of those verbs. As for the preposition *por* ‘for’, it has, among some other predicative meanings, one that can be paraphrased as IN-EXCHANGE-FOR, which can also be posited as a kind of a semantic operator, part of the structural meaning of the ‘exchange’, ‘substitute’ and ‘commercial event’ verbs. All of these predicates have a complex LS composed of two transfer events, in the case of the ‘exchange’ and ‘commercial event verbs’, and two change of location events (one of putting and one of removal), in the case of ‘substitution’ verbs; and they have two themes that are not reciprocal and that never share the same ‘space’, as they end up in different hands or locations, so they do not imply the meaning of ‘togetherness’ which is at the core of the verbs of ‘putting together’ and ‘putting into a relation’ meaning.

Besides *con* ‘with’ and *por* ‘for’, in Spanish, the preposition *de* ‘of’ can also appear marking an oblique argument in three-place constructions where there is a ‘competition’ between two asymmetrical arguments (theme/location; theme/recipient and others) for the undergoer role and a non-default linking takes place. In fact, *de* appears in more structural contexts to mark non-default coding of non-undergoer themes than the expected instrumental-comitative case preposition *con*. In this context, we have proposed two complementary rules for covering the same functional



domain as the rule for the assignment of the preposition *with* English: one rule for the assignment of preposition *de* for marking themes not chosen as undergoers, and one rule for marking co-effectors not chosen as actors.

## 6. References

- Farrell, Patrick. 2009. The preposition “with” in Role and Reference Grammar. In Guerrero, Lilian, Ibáñez-Cerda, Sergio & Belloro, Valeria (eds.), *Studies in Role and Reference Grammar*, 179-202. México: UNAM.
- Foley, William A. & Van Valin, Robert D. Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Gruber, Jeffrey S. 1965. *Studies in lexical relations*. Cambridge: MIT (Doctoral dissertation). Published with revisions as *Lexical structures in syntax and semantics*. Amsterdam: North-Holland. 1976.
- Haspelmath, Martin. 2008. Ditransitive constructions: Towards a new Role and Reference Grammar account? In *Investigations of the syntax-semantics-pragmatics interface*, Van Valin, Robert D. Jr. (ed.), 75-100. Amsterdam: John Benjamins.
- Ibáñez Cerda, Sergio. 2018. Commercial event verbs in Spanish. A corpus-based study. In Kailuweit, Rolf, Künkel, Lisann & Staudinger, Eva (eds.), *Applying and expanding Role and Reference Grammar*, 213-235. Freiburg: FRIAS.
- Ibáñez Cerda, Sergio. 2019. Marcación preposicional de complementos oblicuos con la preposición *de* en español. In Guerrero, Lilian (ed.), *Adposiciones y elementos de su tipo en lenguas de América*, 589-633. México: IIFL-UNAM.
- Jolly, Julia. 1993. Preposition assignment in English. In Van Valin, Robert D. Jr. (ed.), *Advances in Role and Reference Grammar*, 275-310. Amsterdam and Philadelphia: John Benjamins.
- Kailuweit, Rolf. 2008. A RRG description of locative alternation verbs in English, French, German and Italian. In Kailuweit, Rolf, Wiemer, Björn, Staudinger, Eva & Matasović, Ranko (eds.) *New applications of Role and Reference Grammar: Diachrony, grammaticalization, Romance languages*, 328-355. Cambridge: Cambridge Scholars Publishing.
- Van Valin, Robert D. Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D. Jr. 2007. The Role and Reference Grammar analysis of three-place predicates. *Suvremena Lingvistika* 33.1(63): 31-64.

Van Valin, Robert D. Jr. & LaPolla, Randy. 1997. *Syntax, structure, meaning and function*. Cambridge: Cambridge University Press.



**PART 3.**  
**SYNTACTIC STRUCTURE**



# AN RRG ANALYSIS OF NON-ICONIC WORD ORDER IN THE BAMUNKA REFERENCE PHRASE

CIARA ANDERSON

TRINITY COLLEGE DUBLIN, THE UNIVERSITY OF DUBLIN

## Abstract

The notions of 'adjective' and parts-of-speech more generally, have been a matter of some debate in the field of typology. Authors such as Dryer (1992) have used heavily semantic-first definitions when it comes to cross-linguistic comparisons of categories such as adjective. Rijkhoff (2002; 2004) has challenged this, however, in his theory that internal NP ordering patterns iconically reflect the underlying semantic structure of the NP. In such cases, loosely defined categories of terms such as adjective and numeral render this theory unworkable. An analysis of the RP in Bamunka, a Ring language of Grassfields Bantu, through the lens of Role and Reference Grammar (RRG), lends support to Rijkhoff's claim that purely semantic definitions for such categories may not be helpful in typological analyses.

## Keywords

Bamunka, Grassfields Bantu, iconicity, parts-of-speech, adjective.

## 1. Rijkhoff's iconicity principle

Rijkhoff (2004) has proposed that Noun Phrase (NP) internal ordering patterns iconically reflect the underlying (semantic) structure of the NP. He has hypothesized that the few non-iconic basic NP-internal ordering patterns that have been attested do not involve simple, whole (integral) NPs. This deals with descriptive rather than referential modifiers. His breakdown of the layered structure of the noun phrase consists of a model of the NP descriptive modifiers can be distributed over three nested layers.

[T]he quality layer contains the head noun and accommodates modifier categories that only relate to the property that is designated by the noun (qualifying modifiers): nominal aspect markers and (typically) adjectives; the quantity layer contains the quality layer and accommodates modifier categories (quantifying modifiers) having to do with number distinctions (singular, plural) and cardinality (one, two, etc.); the location layer contains both the quality and the quantity layer and accommodates modifier categories specifying properties concerning the location of the referent (localizing modifiers), such as demonstratives and relative clauses (Rijkhoff 2004: 170).

Rijkhoff (2008: 85–87) provides an updated five-layered model of the NP which takes into account a 0-layer under the category of *Kind* accounting for classifying operators and satellites ( $\omega_0$  and  $\tau_0$ ) wherein nominal aspect is accounted for under the  $\omega_0$  categories and qualifying operators have thus been omitted. A Discourse-Referential layer is also present. Figure 1 outlines the descriptive modifier layers of this updated model.

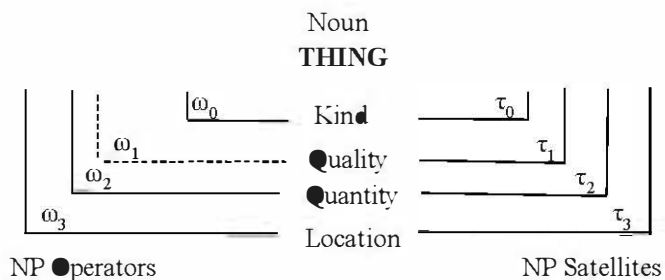


Figure 1: Descriptive modifiers in the layered structure of the NP  
(Adapted from Rijkhoff 2008: 65, 81)

Rijkhoff's (2002; 2004) hypothesis claims that demonstrative (dem), numeral (num) and adjective (A) are ordered according to the scope of semantic relations as seen in the layered organization of the underlying structure seen above, only eight of the 24 logically possible are predicted to occur.

- (1) dem num A N    dem A N num    num A N dem    A N num dem  
 dem num N A    dem N A num    num N A dem    N A num dem  
 (Rijkhoff 2004: 175).

The remaining combinations are viewed as non-iconic. Rijkhoff (2002: 273) points out that Hawkins (1983: 119–120) put forth a modified version

of Greenberg's (1966: 87) Universal 20 when he came across two Niger-Congo languages with a non-iconic basic pattern in the NP: Aghem [N A dem num] and Noni (which has [N dem num A] as well as [N dem A num]).

**Universal 20'.** When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they (i.e., those that do precede) are always found in that order. For those that follow, no predictions are made, though the most frequent order is the mirror-image of the order for preceding modifiers. In no case does the adjective precede the head when the demonstrative or numeral follow (Hawkins 1983: 119–120).

Explanations for such discrepancies in these and other languages include the miscategorisation of parts of speech, for instance, verbal and nominal elements being classified as adjectives based on semantic-first criteria (Rijkhoff, 2004: 178–179). Thus, such cases actually involve constituents such as phrasal constructions or appositions, resulting in a complex NP (Rijkhoff 2002: 327). While Rijkhoff's (2002) approach is largely theory independent, he suggests some common ground between his approach of layered representations and that of Van Valin and La Polla's (1997) Role and Reference Grammar (RRG) framework (Rijkhoff 2002: 4). This paper will address Bamunka's place in the above iconicity theory through the lens of RRG. The RRG framework is a particularly useful approach here in that it does not assume an a priori set of universal lexical categories, such as Noun, Verb, Adjective and Pre/Postposition. It further acknowledges that, in many languages, adjectives are in fact subclasses of verbs or nouns (Van Valin 2008: 161–165). The ability to represent the contents of a referring expression within the context of the notion of a category neutral nucleus deemed Reference Phrase (RP) rather than with the constraint of assigning only a nominal nucleus as the head of an NP is also beneficial (ibid: 167–168).

## 2. The Bamunka language

The Bamunka language is one of 279 indigenous languages spoken in Cameroon. The Ethnologue classifies Bamunka as follows: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Wide Grassfields, Narrow Grassfields, Ring, South (Ethnologue 2018) The word



order of Bamunka is typically SV and, relevant to the present study, it displays a clearly defined noun class system.<sup>1</sup>

## 2.1 Bamunka noun classes

A shared characteristic of languages belonging to the GB family is the classification of nouns according to a complex noun class system. There is evidence of semantically motivated membership of noun classes in Bamunka, though this is not clear-cut (Ingle 2013: 22–25). Membership of a particular noun class is usually denoted by a corresponding affix and the presence of a concord consonant and tone. The noun may occur in an A-form or B-form depending on its focus. A-forms will take the corresponding class marker as a suffix denoting that it is in-focus while B-forms will take the class marker as a prefix indicating that it is out of focus. Bamunka has a total of eight noun classes. Classes 7, 9 and 19 are singular while 2, 6a, 8, 10 and 13 are plural classes. (Ingle 2013: 14–19). Table 1 provides a brief sample of the noun class system with the relevant affixes.<sup>2</sup>

---

<sup>1</sup> Abbreviations: A ‘adjective’, AM ‘associative marker/morpheme’, ANAPH ‘anaphoric demonstrative’, ANM ‘animate’, ATTR ‘attribute’, AUX ‘auxiliary’, CX ‘class where X is relevant class number’, CONJ ‘conjunction’, DEIC ‘deictic’, DEM ‘demonstrative’, GB ‘Grassfields Bantu’, INANM ‘inanimate’, INDEF ‘indefinite’, M ‘modifier’, MP ‘modifier phrase’, N ‘noun’, N1 ‘1st noun in associative noun phrase’, N2 ‘2nd noun in associative noun phrase’, NASP ‘nominal aspect’, NUC ‘nucleus’, NUM ‘numeral’, NUM ‘number’, P ‘preposition’, PP ‘prepositional phrase’, PRED ‘predicate’, QUANT ‘quantifier’, R ‘reference’, REAL ‘realis’, RP ‘reference phrase’, TNS ‘tense’, V ‘verb’.

<sup>2</sup> Note that surface pitch marking in the Bamunka examples reflect Ingle’s (2013) work and have since undergone revision based on personal correspondence with Jane Ingle in April, 2021.

Table 1: Noun class system in Bamunka  
(adapted from Ingle 2013: 19, 82).

Class Number	Affix	Example	Gloss
2	<i>bà-</i>	<i>bà-kaá</i>	Monkeys
6a	<i>-má</i>	<i>nuò-<sup>+</sup>má</i>	Birds
7	<i>-ká</i>	<i>tyú-ká</i>	Tree
8	<i>-bá</i>	<i>tyú-bá</i>	Trees
9	-	<i>ndúú</i>	Cup
10	<i>-há</i>	<i>ndúú-<sup>+</sup>há</i>	Cups
13	<i>-tá</i>	<i>vvi-tá</i>	Mosquitos
19	<i>-há</i>	<i>nuò-<sup>+</sup>há</i>	Bird

## 2.2 Gender

Both double and single class gender nouns are present in Bamunka. Gender pairings refer to the pairing of a singular noun class with a plural noun class. In other words, nouns from a given singular class form a pair with nouns from a specific plural class(es) resulting in a gender (Ingle 2013: 21).

## 3. Modifiers in Bamunka

Ingle (2013: 25–28) has identified two systems of agreement in Bamunka. The first relates to noun class agreement with modifiers. The second refers to a reduced level of agreement with animacy and number in the case of emphatic forms, modifiers acting as head nouns and quantifiers. The basic order of modifiers in Bamunka is as follows in Figure 2:

DEF DET - NOUN - ATTR -  $\left\{ \begin{array}{l} \text{POSS DEM} \\ \text{INDEFDET} \end{array} \right\}$  - QUANT - REL. CLAUSE

Figure 2: Basic order of modifiers in Bamunka (Adapted from Ingle 2013: 30)

It's important to note that Ingle (2013: 30) has found that some attributes come before the noun, and in such cases are viewed as the head noun of an associative construction and that the indefinite determiner does not occur together with the possessive or demonstrative in a noun phrase.

This ordering of constituents does not fall in line with the predictions of Rijkhoff's (2002; 2004) hypothesis which suggests that Noun Phrase (NP) internal ordering patterns iconically reflect the underlying (semantic) structure of the NP. We do not see evidence of any of the 8 predicted iconic patterns in this regard, but similar to Hawkin's (1983) finding in Aghem, also a Grassfields Bantu language, we see an [N A dem num] order of modifiers. See the following illustration in Bamunka:

(2) Bamunka

*mú<sup>+</sup>t̃ f̃f̃ẽ h-uá h-ɔ̀ɔ̀y ibuí*  
 farm[C10] new C10-their C10-these two  
 'these their two new farms' (Ingle 2013: 30)

This begs the question as to whether Rijkhoff's (2002) cautioning against semantically based definitions of such elements, for instance, those offered by Dryer (1992), is warranted. Rijkhoff (2015: 648) notes that, because formal categories are largely too language specific, typologists have often followed Greenberg (1966), who applied semantic criteria to identify the varying types of constituents in his cross-linguistic study of ordering patterns. As a result, his semantic category Adjective largely includes all the forms and constructions that are understood as adjectives in English, regardless of whether the language has a distinctive adjectival class. Some, for instance, instead use abstract nouns expressing a concept like "big-ness<sub>N</sub>" as found in particular Chadic and Bantu languages, or stative verbs along the lines of "be.big<sub>v</sub>", as in a number of SouthEast Asian languages, to express adjectival notions (Rijkhoff 2015: 648). The result being that semantically based categories may paint too broad a brush as evidenced in deviations from predictable typological patterns. It may prove that what has been regarded as an adjective in Bamunka may be better defined in alternative terms such as those relating to verbs or associative noun phrases.

#### 4. The associative noun phrase in Bamunka

Before deciding on how best to represent adjectival-like modifiers in Bamunka, a brief overview of the associative NP construction and its semantics based on Ingle (2013) will be outlined. The associative NP in the Ring languages has a general pattern of a head noun (N1) and the dependent noun (N2) along with an associative marker (AM). In Bamunka, the class suffix of the N1 is deleted and the AM is expressed as an explicit morpheme unless the N1 comes from a class without suffixes. The AM generally takes the same form as the suffix deleted from the N1. The

marker commonly occurs between the two nouns in related languages, whereas in Bamunka it occurs after. Additionally, in Bamunka, if a given class has an A-form/B-form distinction (B-form being the out-of-focus form), the N2 will occur in B-form, that is, taking a prefix rather than a suffix (Ingle, 2013: 79). A similar basic pattern of N1, N2 and associative marking is seen in Grassfields languages such as Bafut and Obang (Asohisi 2015: 151–152; Tamanji 2009: 89–90). These associative constructions can denote meanings such as part-whole relationships, purpose or possession.

## (3) Bafut

*fí-kùù*      *f-í*      *ù-dânwí*  
 C19-bench C19-AM C9-church  
 ‘a church bench’ (Tamanji 2009: 93)

N1: *tóŋ-ká* box-C7    N2: *sè-bǎ* comb-C8

## (4) Bamunka

*tóŋ*      *bá-sè*      *kǎ*  
 box[C7] C8-comb C7.AM  
 ‘box of combs’ (Ingle 2013: 82)

N1: *tóŋ-ká* box-C7    N2: *yó-<sup>+</sup>há* snake-C10

## (5) Bamunka

*tóŋ*      *yó-<sup>+</sup>há*      *ká*  
 box[C7] C10-snake C7.AM  
 ‘box of snakes’ (Ingle 2013: 85)

(4) and (5) illustrate examples of associative NPs in which the N2 comes from a class with and without a B-form, respectively.

Pavey (2010: 182–183), too, points out that noun phrases may contain more than one noun and that such constructions may express association, relation, or a part-whole relationship. Rather than placing the second noun in the periphery of the nucleus as a modifier in an RRG-like representation, the dependent noun is treated as an argument of the head noun, just as predicates may take an argument in the clause structure. In line with Pavey’s definition of such constructions, both Hyman (1979: 35) and Schaub (1985: 298) point to NP constructions with an associative marker as expressing possession or other genitive functions in the Ring languages of Aghem and Babungo, respectively. Ingle (2013: 92) further elaborates on functions of the associative NP in Bamunka including “part-whole” (‘wall of house’), “product-material” (‘bamboo fence’), and “object-

purpose” (‘cooking pots’). Hyman (1979: 34) points out that certain adjectival notions may also be expressed by the associative phrase in Aghem. Many of what are deemed “adjectives”, in such cases, and in Bamunka in particular, take these associative structures and therefore may best be realised as such in the RRG representation while others may best be placed in the periphery of the nucleus.

## 5. Adjectives in Bamunka

In Bamunka, adjectives can have a nominal or verbal basis. They are either nouns describing other nouns in a type of associative noun phrase, or they are formed from reduplicated verb stems or verbs with a verbal extension (Ingle, 2013, 62–71). These measures may be diagnostic of the markedness of nouns and verbs in other functions. In fact, Rijkhoff (2002: 16, 122–139) adopts a modified form of Hengeveld’s (1992 a, b) approach to word classes in distinguishing the presence of a large closed adjectival class versus a nominal or verbal class with further measures taken on it to function in the role of modifier.

An *adjectival predicate* is a predicate which, without further measures being taken, can be used as a modifier of a nominal head (Hengeveld 1992b: 58).

Matasović (2001: 2), too, in an analysis of adjective phrases via the RRG framework acknowledges that in a number of languages “adjectives do not exist as a separate word-class; words with adjectival meaning are expressed either as verbs, or as nouns.” It may therefore be the case that modifiers of the reference phrase in Bamunka are marked noun and verb forms being used in an extended function rather than a distinct closed word class in and of themselves. For more detailed discussion on Rijkhoff’s modified approach see (Rijkhoff 2002; Browne 2020: 100–102).

### 5.1 Modifiers based on nouns

As outlined by Ingle (2013: 62–70), these include attributes which occur as other nouns which appear before or after the noun being characterised. Noun-like elements that follow the noun in modification function are found in three colours and in compound nouns (verb-noun combinations). In fact, these three colours in Bamunka come directly from class 7 and when they modify a head noun they occur in their B-form (prefixed or out

of focus form) demonstrating strong characteristics of functioning as the second noun in an associative construction rather than as an attributive adjective. Verb-noun compounds also behave as the second noun in an associative construction.

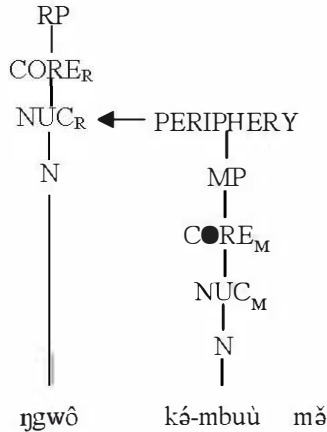


Figure 3: Noun functioning as modifier

(6) Bamunka

*mbuù-kǎ*

red-C7

‘red’ (Ingle 2013: 69)

(7) Bamunka

*ngwô ká-mbuù mǎ*

oil[C6a] C7-red C6a.AM

‘red oil’ (Ingle 2013: 69)

Attributive nouns that describe the colour of a preceding noun occur in their prefixed B-form. This is followed by an associative marker that agrees with the class of the head noun (Ingle 2013: 69). Thus, this construction appears to be in the form of an associative noun phrase showing the relevant agreement with a function of modification, rather than a distinct member of an adjectival word class acting as a modifier of the reference phrase. Drawing on the RRG framework, Pavey (2010: 182–183) points out that noun phrases may have more than one noun and that such constructions may express notions such as association and part-whole rela-

tionships between two nouns. Rather than putting the second noun in the periphery of the nucleus as a modifier we treat the dependent noun as an argument of the head noun, just as predicates may take an argument in the clause structure. Pavey notes that other devices such as case marking may also point to an associative or genitive construction. An English example would be through the use of an adpositional phrase as seen in Figure 4. The notion of RP rather than in NP is used in contrast to Pavey's account.

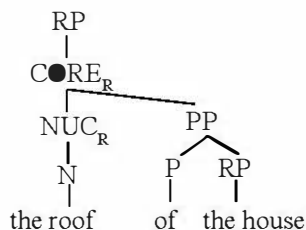


Figure 4: Associative NP in RRG framework adapted from Pavey (2010: 183)

Based on this, a modified version of Pavey's realisation of an associative NP in Bamunka will be used henceforth rather than the nuclear peripheral categorisation of the dependent noun.

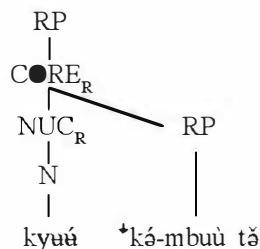


Figure 5: Class 7 Noun-like modifier in Bamunka

(8) Bamunka

*kyuú* \**ká-mbuù tǎ*  
 pot[C13] C7-red C13.AM  
 'red pots' (Ingle 2013: 70)

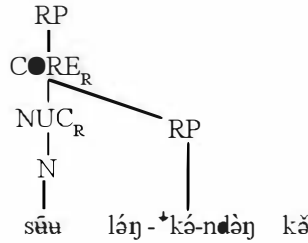


Figure 6: Verb-noun compound modifier in Bamunka

(9) Bamunka

*sũũ lɔŋ-ʰká-ndəŋ kǎ*  
 hoe[C7] work-C7-farm C7.AM  
 ‘hoe for farming’ (Ingle 2013: 70)

Difficulty comes in regarding noun-like modifiers that precede the noun. There are two instances of this, *vaã* ‘child’ with a diminutive function and *ŋkwẽ* ‘mother’ with an augmentative function (Ingle 2013: 71). Ingle suggests that while, semantically, they act as modifiers, grammatically, they are the head of an associative construction. The noun being described occurs as the N2, though because they are of classes 9 and 2 we don’t see agreement marking. This would pose a problem for the selection of the modifier as a peripheral element in the RRG analysis, as the N1 would take the position of nucleus and modifier simultaneously. However, an associative realisation accounts for both without imposing inaccurate semantics.

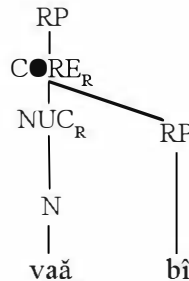


Figure 7: Noun-like modifier preceding noun in Bamunka



## (10) Bamunka

*vaã bi*  
 child[C9] goat[C9]  
 ‘small goat’ (Ingle 2013: 71)

Again, we do not see an explicit AM in this case as we are dealing with a class 9 noun which has  $\emptyset$  marking. There are two potential interpretations of such modifiers, as nouns in an associative NP or as modifiers undergoing semantic bleaching and potentially being grammaticalised into a classifier form; perhaps numeral classifier or noun class marker. For more detailed discussion on semantic bleaching in a case like this see Browne (2020: 167–168)

## 5.2 Modifiers based on verbs

A second set of modifiers in Bamunka are those derived from verbs. The data examples that follow are sourced from Ingle (2013: 63–65). The first type is those which are a reduplicated form of the original verb stem as follows.

Verb *chyii* ‘be sharp’      Adjective *chyii~chyii* ‘sharp’

## (11) Bamunka

*nyii*      <sup>+</sup>*chyii~chyii*  
 cutlass[C9] sharp~sharp  
 ‘sharp cutlass’ (Ingle 2013: 63)

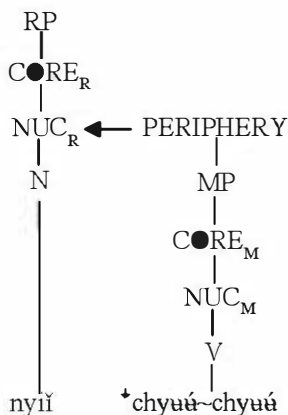


Figure 8: Reduplicated verb stem as modifier

Other adjectives also appear to be reduplicated forms, but their origin is uncertain. It is possible they may have been derived from verbs since lost from the language, but further research is needed (Ingle 2013: 63). While at first glance a peripheral allocation of the reduplicated verbal modifier seems appropriate, it must be noted that class 9 nouns do not take an AM. The distinctive patterning of such constructions from noun classes that take an explicit AM appears to be in line with that of the associative NP in which the suffix is deleted and follows the dependent element.

**Verb** *dà* ‘lengthen’      **Adjective** *dā~dā* ‘long’  
 (12) Bamunka  
*fiā*      <sup>+</sup>*dā~*<sup>+</sup>*dā* *tā*  
 road[C13] long~long C13.AM  
 ‘long roads’ (Ingle 2013: 63)

A further argument could be made for the nominalisation of the modifying element here as associative NPs tend to denote meanings roughly translated as ‘X of Y’, thus, a more accurate interpretation may be along the lines of ‘cutlass of sharpness’ or ‘road of length’ rather than a peripheral element carrying out the function of modification. This would yield a more uniform interpretation of an associative representation. Thus, we see evidence for further measures being taken on a verb root in order to operate in the function of modifier, rather than the existence of a distinctly specialised adjectival class.

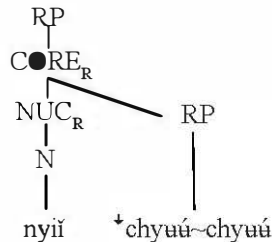


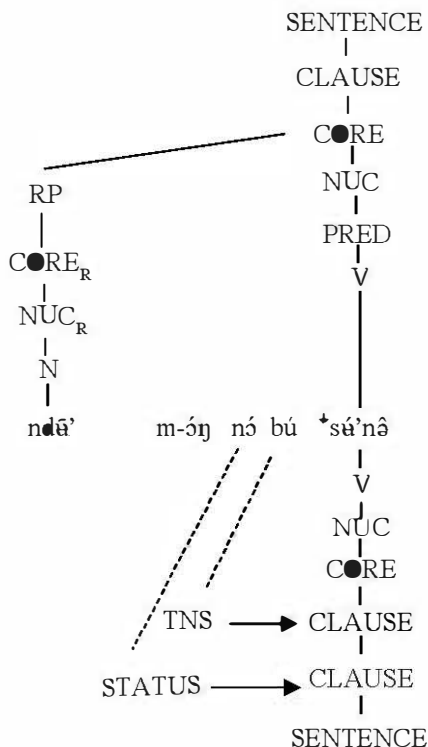
Figure 9: Reduplicated verb as N2 in associative construction

The second type of modifier in Bamunka is derived from verbs are stative verbs with the verbal extensions (*-hā*, *-nā*). These are used to describe the characteristics of the noun in three ways, in verb form (in the position of predicate), in attributive form (positioned as an attribute in a noun phrase and demonstrating agreement with the modified noun) and used with a copula verb as a head noun or verb form (Ingle 2013: 64–65). This

brings to mind Rijkhoff's critique of the use the term *Adjective* when referring to the use of abstract nouns such as "bigness<sub>N</sub>" or stative verbs like "be.big<sub>V</sub>" express adjectival notions (Rijkhoff 2015: 648). It is also reminiscent of Van Valin's (2008) assertion that certain adjectives may in fact be a subclass of verbs. Examples of the forms of *sù'ná* 'be sweet' follow (Ingle, 2013: 65).

## (13) Bamunka

*ndú'*      *m-óŋ*      *nó*      *bú*      *\*sù'ná*  
 wine[C6a] C6a-1SG.P●SS REAL FUT be.sweet  
 'my wine will be sweet.' (Ingle 2013: 65)<sup>3</sup>

Figure 10: Predicative use of *\*sù'ná*<sup>4</sup>

<sup>3</sup> Based on personal correspondence with Jane Ingle in Aug 2020, the verb phrase particle *nó* is glossed as realis in contrast to Ingle (2013).

(14) Bamunka  
*ó †nó kǎ ndú' sù'ná má tá mǎ*  
 3SG REAL give wine[C6a] be.sweet C6a.AM to 1SG  
 'He gave me sweet wine' (Ingle 2013: 65)

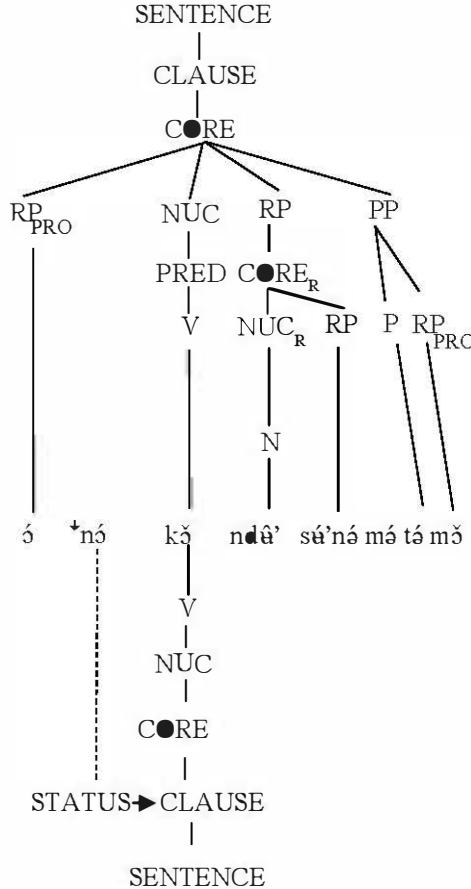


Figure 11: *sù'ná* as the N2 in an associative construction

In the above form the verbal modifier may perhaps be interpreted as being nominalised in that we again see an associative agreement marker occur-

<sup>4</sup> Note that the marking of illocutionary force in Bamunka requires further research on the layered structure of the clause.

ring after the element deemed attribute pointing to the pattern of an associative NP. For instance, Ingle (2013: 95) states that verbs or adjectives in such a construction may serve as “the associative noun”.

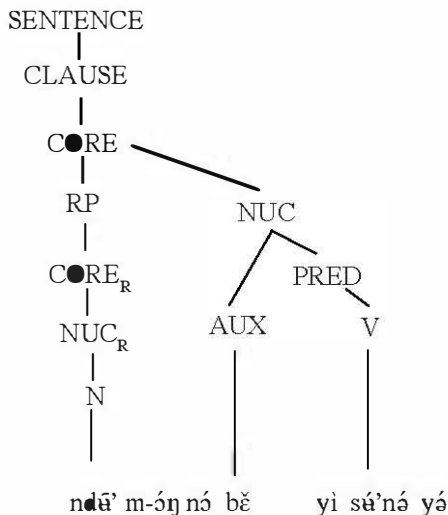


Figure 12: Copular use of *sù'ná*

(15) Bamunka

*ndú' m-óŋ nó bě yi sù'ná yá*  
 wine[C6a] C6a-1 SG.POSS REAL be SG sweet SG  
 'My wine is sweet' (Ingle 2013: 65)

While a number of languages use non-verbal predicates to express such notions, there is evidence of the use of verbal predicates for existential functions (Pavey 2010: 59). Again, we do see evidence of a verb root which is modified in the form of an associative NP construction to perform the function of attribute rather than the presence of a distinct, closed adjectival class.

### 5.3 Real adjectives in Bamunka

Ingle (2013: 67) points out that there is an attribute *kwétāŋ* 'big' which is not an instance of reduplication and does not have a known verb form. While it is may be a noun, its derivation is unclear, thus we may be dealing with a real adjective. It is also highlighted that other adjectives are evidently reduplicated forms, but their actual source is uncertain. They may have

been derived from verbs that have been lost from the language, or they may have been derived from other forms (Ingle 2013: 63). While not clearly members of a distinct adjectival class they at least raise questions as to their nature.

(15) Bamunka

*múts* *fě~fě*  
 farm[C9] new~new  
 ‘new farm’ (Ingle 2013: 64)

The analysis thus far suggests that when it comes to modifiers of the reference phrase in Bamunka, there is strong evidence for forms which are morpho-syntactically altered forms of nouns and verbs through processes such as reduplication and syntactic positioning. The existence of a distinct adjectival class as per Rijkhoff’s (2002) modified version of Hengeveld’s (1992 a, b) approach to parts of speech does not appear to be present. Evidence for a small, closed class of true adjectives is unclear, with the origin of a selection of derived forms unknown.

#### 5.4 The associative RP in Bamunka and Rijkhoff’s layered model

It has been noted that in an earlier version of Rijkhoff’s analysis, non- iconic ordering in the layered structure of the NP may be due to a misallocation of certain elements as simplex adjectives and numerals. Interestingly, the associative NP construction displayed a number of parallels with Rijkhoff’s (2008: 80) updated version of the layered structure of the NP which incorporates a  $\tau_0$  layer relating to “classifying satellites” which points to the kind or subclass of the entity denoted by the head noun. In English this would include the following the examples outlined in (16).

(16) English

- a. a **corporate** lawyer
- b. a house **of sin** (Rijkhoff 2008: 80)

Rijkhoff’s (2008: 83) further notes that the semantic range of such classifying satellites as noted by Halliday (2014: 377) include material, purpose and function. Thus, both syntactic and semantic similarities are demonstrated between these and the associative construction in Bamunka which include semantic functions such as “product-material” and “object-purpose” (Ingle, 2013: 92). Further to that, where Rijkhoff (2008: 83) points

to the difficulty in drawing a line between “a noun + classifying satellite combination and compounds or quasi-compounds”, Ingle (2013: 92) does the same in acknowledging that in Bamunka “more research may be required to determine whether a construction is a compound noun or a noun phrase/ associative noun phrase.” So, while the ordering of noun and “adjective” is iconic at the surface level, it appears that we are not dealing with a definitive member of a distinct adjectival class, but rather nouns and nominalized verbs functioning as classifying modifiers within the syntactic structure of an associative NP. Rijkhoff (2008: 83) notes that such classifying satellites can take various forms and constructions including adjectives, nouns and genitive NPs.

## 6. The demonstrative

In addition to an anaphoric demonstrative, the two main distinctions in Bamunka are proximal and distal demonstratives which both occur after the noun. With the exception of class 9 nouns which do not have an affix, the head noun suffix is deleted and a concord prefix is placed on the demonstrative (Ingle 2013: 31).

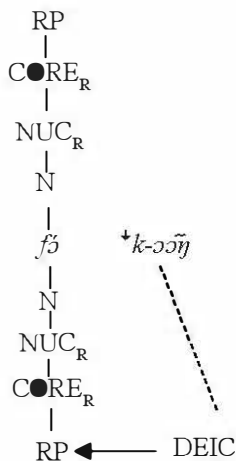


Figure 13: Proximal demonstrative in Bamunka

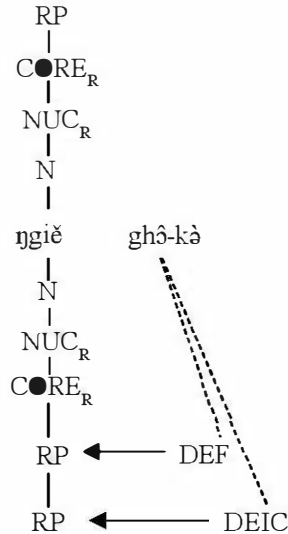


Figure 14: Anaphoric demonstrative in Bamunka

**Proximal**

(17) Bamunka

fɔ̃      \*k-ɔ̃ŋ  
 thing[C7] C7-PROX  
 ‘this thing’ (Ingle 2013: 31)

(18) Bamunka

ŋgiě      ghô-kà  
 language[C7] ANAPH-C7  
 ‘that language’ i.e. ‘those words (referred to already)’  
 (Ingle 2013: 32)

The anaphoric demonstrative identifies something that has already been mentioned and it takes a noun class concord suffix if one is present. It appears to be similar in function to the definite determiner *mbé* (Ingle 2013: 37–38, 44).

As regards the demonstrative, we do not observe evidence of the presence of a complex constituent or misnamed category, and so, this study will move on to an examination of the numeral in Bamunka.



## 7. The numeral in Bamunka

Numerals from one to ten in Bamunka follow the noun and show no agreement with the class of the head noun. Three, however (numbers ‘two’, ‘three’, and ‘five’), do take a prefix depending on the animacy/inanimacy of the head noun, the prefixes *bà-* or *i-*, respectively (Ingle 2013: 50–51).

(19) Bamunka

*njaá-<sup>+</sup>há i-buú*  
house-C10 INANM-two  
‘two houses’ (Ingle 2013: 50)

(20) Bamunka

*bá-<sup>+</sup>ké bə-buú*  
C2-female ANM-two  
‘two women’ (Ingle 2013: 51)

The question of whether the Bamunka numeral can be regarded as a simple constituent of the NP, or in Rijkhoff’s (2002: 171, 335) terms a complex or “embedded” constituent, arises in numerals above ten. The word *njuú* ‘digit’, a probable gender 9/10 noun, is used in the formation of numbers above ten. A number between one and nine follows *njuú* and agreement marking is absent. This is followed by the conjunction *nə* ‘and’ along with the word *wúŋ* ‘ten’ (Ingle 2013: 52). This points to the existence of a complex NP when such elements modify another noun.

(21) Bamunka

*njuú buú nə wúŋ*  
digit two and ten  
‘twelve’ (Ingle 2013: 52)

While Ingle points out that the elements *nə wúŋ* are not always used in actual speech as seen in (22), the structure itself appears complex in nature.

(22) Bamunka

*njuú buú*  
digit two  
‘twelve’ (Ingle 2013: 53)

- (23) Bamunka  
*wūy buĩ*  
 ten two  
 ‘twenty’ (Ingle 2013: 53)

As outlined by Ingle (2013:53–55), decades are formed using the word *wūy* ‘ten’ plus a numeral from one to nine as in (23). Numbers above each decade are formed in the same complex manner mentioned above with the conjunction *nà* ‘and’ along with the decade *wūy* + numeral. However, in these cases, the presence of *wūy* and the numeral is obligatory. Numbers beyond one hundred are formed with *ghyui* ‘one hundred’ or *ghyui-tá* + numeral for multiples of hundreds. These are followed by the conjunction *nà* ‘and’ and a numeral. A ‘thousand’ is expressed by *ηkóη* (plural *ηkóη-há*), a gender 9/10 noun, and numeral classifier *mbyuĩ* is required to denote ‘one thousand’. For further details on numbers above one thousand see Ingle (2013: 55–57).

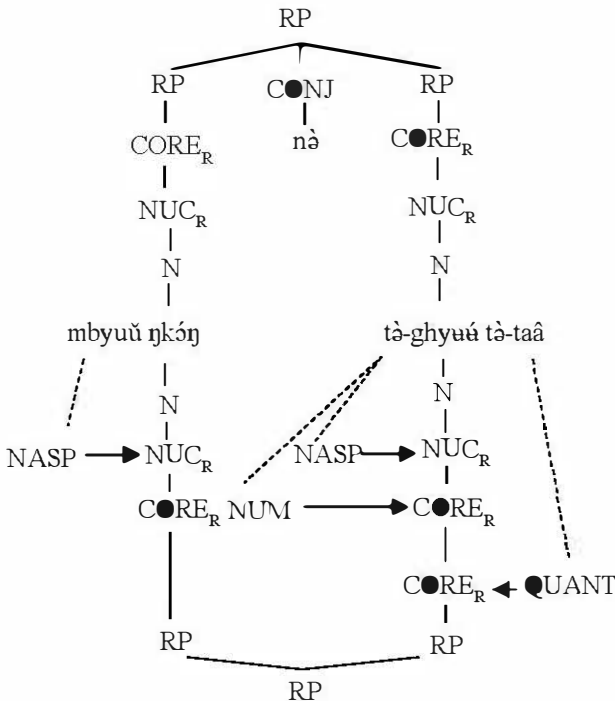


Figure 15: Complex numeral in Bamunka

Van Valin and La Polla (1997: 492) have pointed out that nexus-juncture relations found in the clause may also be identified in a similar fashion in the LSNP. They posit that the simplest example of NP-level linkage is found in conjoined NPs such as *the woman and the man*. Those sharing a determiner but independent of other operators are deemed NP cosubordination and those that take the full range of independent operators are deemed NP coordination. While a numeral may not take all NP operators, evidence from older speakers suggest that in the recent past complex numerals took independent agreement marking. In place of an inanimate agreement prefix in (25), the numeral in (24) has a prefix agreeing with the noun class of the preceding noun *ghyuu<sup>+</sup>tá* ‘hundreds’ in its B-form with numerals ‘one’, ‘two’, ‘three’ and ‘five’. With other numerals, all speakers use the A-form of *ghyuu<sup>+</sup>tá* ‘hundreds’, without agreement on the numeral (Ingle 2013: 56). Browne (2020: 217–220) has argued for the potential of noun class affixes to be categorised as nuclear aspect operators. Thus, we provisionally see complex numerals in Bamunka, at least in the recent historical sense among older speakers, as an example of a complex NP showing NP coordination with independent operators. The analysis in Figure 16 uses the updated terminology of RP.

Among younger speakers this has become simplified with the majority using the A-form of *ghyuu<sup>+</sup>tá* ‘hundred’ and an inanimate prefix on the numeral (Ingle 2013: 56).

(24) Bamunka

*mbyuú ykóŋ                      nà tà-ghyuu<sup>+</sup>      tà-taâ*  
unit[C9] thousand[C9] and C13-hundred C13-five  
‘one thousand five hundred’ (Ingle 2013: 56)

(25) Bamunka

*mbyuú ykóŋ                      nà ghyuu<sup>+</sup>tá      i-taâ*  
unit[C9] thousand[C9] and hundred-C13 INANM-five  
‘one thousand five hundred’ (Ingle 2013: 56)

## 8. Possible explanations for non-iconic ordering in the Bamunka RP

At first glance the languages above are a non-iconic anomaly as regards Rijkhoff’s (2002; 2004) predictions, similar to those identified by Hawkins (1983), leading him to modify Greenberg’s twentieth universal. It may be that the Bamunka language has raised some challenges to this theory of iconicity, but an analysis of the data suggests some other explanations.

## 8.1 The questionable status of the adjective

While position of “adjective” is iconic, based on our observations and Rijkhoff’s issues with the misattribution of certain elements such as appositional NPs as adjectives, it’s worth highlighting that this may arise as an issue in future typological research. Rijkhoff (2004: 178–179) noted the necessity of rightly categorising elements within the scope of the simplex NP and that apparently non-iconic predictions regarding the word order in the integral NP can be due to a misattribution of word classes. Modifiers in Bamunka are often noun or verb-like in nature with a number of “adjectives” taking the agreement marking of an associative NP. An RRG analysis of the RP in Bamunka revealed that, rather than treating such modifiers as peripheral modifying phrases, they are better captured as an argument within the RP in an adapted version of Pavey’s (2010) approach to associative NPs and are often best understood semantically in terms of “X of Y”. Additionally, a comparison of the syntactic structure and semantic commonalities such as “object-purpose” and “product-material” revealed that they may best be regarded as nouns or nominalised verbs belonging to the  $\tau_0$  layer of classifying satellite, rather than as a member of the  $\tau_1$  layer of qualifying satellite as per Rijkhoff (2008). Thus, in relation to Rijkhoff’s (2002) modified version of Hengeveld’s (1992a, b) approach to word classes, which takes both semantic and syntactic factors into account, there is further evidence for Bamunka as a language that is lacking a large set of distinctive adjectives.

## 8.2 The numeral as a complex constituent

A brief sketch of the realisation of NP modifiers in Bamunka through the lens of RRG in fact lends support to Rijkhoff’s (2002; 2004) theory of non-iconic anomalies. Initially, the presence of the demonstrative intervening between “adjective” and numeral appeared to contradict the scopal principle which is expected to be reflected iconically in the realization of the linguistic expression. Two ordering features are predicted by this *Principle of Scope*. Firstly, that constituents in the scope of a certain modifier (part of the same semantic layer) are expressed in a continuous sequence. And secondly, that operators and satellites occur immediately before or after the material they have in their scope (Rijkhoff 2002: 313).

While Rijkhoff’s (2002; 2004) model of the layered structure has been updated, we would still expect that the demonstrative ( $\omega_3$ ) would not intervene between the numeral ( $\omega_2$ ) and head noun. We have suggested that the

majority of elements deemed as adjectives in Bamunka may best be described as classifying satellites ( $\tau_0$ ), but the matter of the numeral can also be addressed. With reference to a similar situation in Babungo, Rijkhoff (2002: 324, 325) suggests that this is not a counter-example in that numerals are expressed in the form of an embedded modifier, that is a quantifying satellite ( $\tau_2$ ) rather than an operator ( $\omega_2$ ) and that their syntactic properties can be explained due to the fact that numerals are subject to conflicting ordering preferences. He highlights that, according to the *Principle of Scope* and the *Principle of Head Proximity*, the preferred pattern is [N num.<sub>2</sub> dem]. However, the *Principle of Domain Integrity*, a competing principle, states that “constituents of the matrix domain...prefer not to be interrupted by embedded domains, such as lexical expressions of cardinality” (Rijkhoff, 2002: 325). Thus, with regard to cases which explicitly use nouns in numeral constructions such as *njuɔ* ‘digit’ and *ghyuu-tá* ‘hundreds’, there is evidence of a complex construction. This explains why constituents of the [N dem] matrix domain in (26) are not interrupted by embedded domains, such as lexical expressions of cardinality. Thus, we see the competing interactions of three ordering principles at work. Future research could further examine the possibility of redundancy of use of nominal elements in numerals 1–9, for instance, as appears to be the case in Babungo (Rijkhoff 2002: 171).

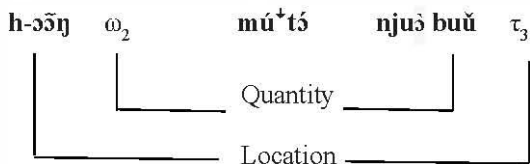


Figure 16: Quantity and location operators and satellites in Bamunka (Adapted from Rijkhoff, 2008: 65; 2002: 220–222)

#### (26) Bamunka

*mú'tá fɛfɛ h-ɔɔ njuɔ buɩ*  
 fam[C10]new C10-3PL.POSS C10-PROX digit two  
 ‘these their twelve new farms’ (Ingle 2013:53)

## 9. Summary

This analysis has raised a number of questions relating to Rijkhoff’s (2002; 2004) prediction that the ordering of modifying elements of the NP in a given language will iconically reflect the underlying layered semantic

structure. At first glance, it appears that Bamunka is a non-iconic anomaly here similar to those identified by Hawkins (1983), but an analysis of the data suggests alternative explanations.

As noted by Rijkhoff (2015), semantic first definitions of elements such as that of “adjective” can become too loosely defined. This was evident in an RRG breakdown of so-called adjectives in Bamunka in which we see verb and noun-like constituents functioning in the role of modification. The close alignment of such elements and the nouns they modify with the structure of the associative noun phrase in Bamunka further highlighted the problem of allocating such elements to a distinctive adjectival class. Additionally, Rijkhoff’s (2008) updated model of the layered structure of the NP indicates that we may in fact be dealing with classifying satellites in such cases.

An analysis of the numeral as a complex constituent in Bamunka and a quantity satellite rather than operator suggested that the occurrence of the demonstrative between numeral and noun could potentially be explained by means of competing principles. That is to say that in categorising the numeral as a quantifying satellite it is subject to the *Principle of Domain Integrity*, thus, explaining its apparently non-iconic positioning.

The above findings suggest that although Bamunka, at first, appears to be non-iconic with regard to Rijkhoff’s (2004) theory, the explanation may be found in the miscategorisation of constituents rather than his typological predictions. While authors such as Dryer (1992) have used heavily semantically based definitions for notions such as “adjective” for cross-linguistic purposes, the difficulties with utilising such a broad term have become apparent in the above analysis. Constituents which may be better labelled in verbal or noun-like terms at first suggest a non-iconic typological pattern, when in fact, it may be that they have been categorised in semantic terms outside of the scope of such typological predictions. As research into Bamunka is still at an early stage, further studies into adjectival-like elements and, in particular, those reduplicated forms whose origin is as yet unknown could strengthen the current findings. Additionally, an extended study of the construction of numerals via nominal elements and what range of operators they may take could shed further light on the RRG analysis made and the ordering principles at work within the Bamunka RP.

## Acknowledgements

Many thanks to Jane Ingle of SIL and to Robert Van Valin Jr. for their valuable feedback on this paper.

## References

- Asohisi, Melvice. 2015. *Structural and Typological Approaches to Obang Grammar*. Cologne: Rüdiger Köppe Verlag
- Browne, Ciara. 2020. *Accounting for the features of Shape and Transitivity in a Functional-Typological approach to the Ring Languages*. Dublin: Trinity College Dublin. (Doctoral thesis.)
- Dryer, Matthew, S., 1992. The Greenbergian word order correlations. *Language* 68 (1): 81–138. doi:10.1353/lan.1992.0028
- Ethnologue. 2018. *Bamunka*. (<https://www.ethnologue.com/language/bvm>) (Accessed 2018-12-22.)
- Halliday, M.A.K. 2014. *An Introduction to Functional Grammar* (revised by Matthiessen, Christian M. I. M.). Oxon: Routledge.
- Hawkins, John. A. 1983. *Word Order Universals*. New York: Academic Press.
- Hengeveld, Kees. 1992a. Parts of speech. In Fortescue, Michael, Harder, Peter & Kristoffersen, Lars (eds.), *Layered structure and reference in a functional perspective*, 29–55. Amsterdam: Benjamins.
- . 1992b. *Non-verbal predication: theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Hyman, Larry, M. 1979. Phonology and noun structure. In Hyman, Larry M. (ed.), *Aghem Grammatical Structure. Southern California occasional papers in linguistics 7*, 1–72. Los Angeles: Dept. of Linguistics. University of Southern California.
- Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, Joseph H. (ed.), *Universals of Language*, 73–113. Cambridge, MA: MIT Press.
- Ingle, Jane. 2013. *The Bamunka Noun Phrase*. Yaoundé Cameroon: SIL. (<https://www.sil.org/resources/archives/53981>) (Accessed 2020-08-15.)
- Matasović, Ranko. 2001. *Adjective Phrases*. (<http://www.acsu.buffalo.edu/~rrgpage/trg/Matasovic.pdf>) (Accessed 2020-08-21.)
- Pavey, Emma. 2010. *The Structure of Language. An Introduction to Grammatical Analysis*. Cambridge: Cambridge University Press.
- Rijkhoff, J. 2002. *The Noun Phrase*. Oxford: Oxford University Press.
- . 2004. Iconic and non-iconic word order patterns: on symmetry in the NP and counter examples to Universal 20. In Aertsen, Henk, Hannay, Mike & Lyall, Rod (eds.), *Words in their places: a Festschrift for J. Lachlan Mackenzie*, 169–180. Amsterdam: Free University, Faculty of Arts.

- . 2008. Layers, levels and contexts in Functional Discourse Grammar. In García Velasco, Daniel & Rijkhoff, Jan (eds.), *The Noun Phrase in Functional Discourse Grammar*, 63–116. Berlin: De Gruyter Mouton.
- . 2015. Word Order. In Wright, James D. (ed.), *International Encyclopedia of the Social & Behavioral Sciences Reference Work*, 644–656. Amsterdam: Elsevier.
- Shaub, Willi. 1985. *Babungo*. London: Croom Helm.
- Tamanji, Pius N. 2009. *A Descriptive Grammar of Bafut*. Cologne: Rüdiger Köppe Verlag.
- Van Valin, Robert D. & LaPolla, Randy J. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.
- Van Valin, Jr, Robert D., 2008. RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In Van Valin, Jr, Robert D. (ed.), *Investigations of the Syntax–Semantics–Pragmatics Interface*, 161–178. Amsterdam: John Benjamins.





# COSUBORDINATION

ROBERT D. VAN VALIN, JR.

UNIVERSITY AT BUFFALO, THE STATE UNIVERSITY  
OF NEW YORK

HEINRICH HEINE UNIVERSITY DÜSSELDORF

## Abstract

The theory of clause linkage in Role and Reference Grammar [RRG] is one of the most important and distinctive aspects of the theory. One of its significant features is the positing of a third syntactic linkage type, cosubordination, in addition to the two traditional linkage types, coordination and subordination, which has been widely adopted in the typological literature and used in many descriptive grammars. Nevertheless, its validity as a distinct linkage type has been questioned in Foley (2010) and Bickel (2010). The purpose of this paper is to evaluate their arguments and show that cosubordination is a valid concept, albeit more complex than originally supposed.

## Keywords

Clause linkage, nexus, embedding, juncture, information structure

## 1. Introduction

The theory of clause linkage in Role and Reference Grammar [RRG] is one of the most important and distinctive aspects of the theory. One of its most significant features is the positing of a third syntactic linkage type, cosubordination, in addition to the two traditional linkage types, coordination and subordination. While this notion has been widely adopted in the typological literature and used in many descriptive grammars, it has been criticized in Foley (2010) and Bickel (2010), who questioned its validity as a distinct linkage type. The purpose of this paper is to evaluate their arguments and argue that cosubordination is a valid concept, albeit more complex than originally supposed.

The discussion will proceed as follows. In section 2 there will be a brief review of the theory of clause linkage in RRG, followed in section 3 by a summary of Foley and Bickel's criticisms of the notion of cosubordination. In section 4 their arguments will be subjected to critical evaluation, and it will be argued that they are not a problem for the contemporary theory of clause linkage in RRG. Section 5 gives a summary and conclusions.

## 2. The RRG theory of clause linkage: a review<sup>1</sup>

The RRG theory of complex sentences has three main components: the layered structure of the clause [LSC], which supplies the units which are combined in complex sentences; the three syntactic linkage relations (coordination, subordination, cosubordination), which characterize the syntactic relationship between the units; and the interclausal semantic relations hierarchy, which deals with the semantic relationship between the units. Only the first two are relevant to this discussion. The units of clause structure (the nucleus, the core and the clause) define the levels of *juncture*: nuclear junctures involve the linking of nuclei, as in (1a,b), core junctures the linking of cores, as in (2a,b), and clausal junctures the linking of whole clauses, as in (3)<sup>2</sup>

- (1) a. *Tā* [<sub>N</sub>*qiāo*] [<sub>NP0</sub>] *le yī ge fānwǎn.* Mandarin Chinese  
 3sg hit break PRFV one CL ricebowl (Hansell 1993)  
 'He broke (by hitting) a ricebowl.'
- b. *Fu fase* [<sub>N</sub>*fɪ*] [<sub>N</sub>*isoe*]. Barai (Olson 1981)  
 3sg letter sit write [Papua-New Guinea]  
 'He sat writing a letter.'

<sup>1</sup> For detailed discussion of the RRG theory of juncture-nexus types, see Van Valin & LaPolla (1997), chapter 8, and Van Valin (2005), chapter 6.

<sup>2</sup> Abbreviations: AFD 'actual focus domain', ASP 'aspect', ASS 'assertion', C 'core', Cl 'clause', CL 'classifier', CLM 'clause-linkage marker', CMPL 'completive', CONT 'continuative', CUR.REL 'current relevance, all 'dual', DS 'different subject', DIR 'directional', DUR 'durative', EVID 'evidential', FUT 'future', IF 'illocutionary force', IND 'indicative', IPFV 'imperfective', LOC 'locative', LSC 'layered structure of the clause', NEG 'negation', NPST 'non-past', ns 'non-singular', N, NUC 'nucleus', POSS 'possessor', PRED 'predicate', PRES 'present', PRFV 'perfective', PRON 'pronoun', Q 'question', RP 'reference phrase', SEQ 'sequential', SIM 'simultaneous', SS 'same subject', STA 'status', TNS 'tense', TPAST 'today's past', TRANS 'transitive', YPAST 'yesterday's past'.

- (2) a. [<sub>C</sub> Max tried<sub>d</sub>] [<sub>C</sub> to fix his bicycle].  
 b. [<sub>C</sub> Max regrets [<sub>C</sub> asking Bill about it]].
- (3) [<sub>C1</sub> Mary bought fresh fish at the market] and [<sub>C1</sub> John will cook it].

The second relevant component is the syntactic relationship between the units, or *nexus* relation. The example in (3) exemplifies (clausal) coordination, and the one in (2b) illustrates a type of (core) subordination, in which one core (*asking Bill about it*) functions as a core argument of another core (*Max regrets*). Cosubordination, as the name implies, has features of both coordination and subordination. It is like coordination and unlike subordination, in that it is a flat structure (no embedding), but it is like subordination and unlike coordination in that the linked unit is dependent on the matrix (or ‘licensing’) unit in some way. In cosubordination the dependence concerns operators at the level of juncture. The contrast among the three nexus types at the clause level can be seen clearly in the following examples from Amele (Roberts 1988).

- (4) a. *Fred cum ho-i-an qa Bill uqadec h-ugi-an.*  
 yesterday come-3sg-YPAST but tomorrow come-3sg-FUT  
 ‘Fred came yesterday, but Bill will come tomorrow.’
- b. *Ija ja hud-ig-a eu nu, uqa sab mane-i-a.*  
 1sg fire open-1sg-TPAST that for 3sg food roast-3sg-TPAST  
 ‘Because I lit the fire, she cooked the food.’
- c. *Ho busale-ce-b dana age qo-ig-a.*  
 pig run.out-DS-3sg man 3pl hit-3pl-TPAST  
 ‘The pig ran out and the men killed it.’

In (4a) classic coordination at the clause level is exemplified: each clause is fully inflected and could stand on its own as an independent utterance. In (4b) each clause is fully inflected, but the first clause is marked by a subordinating conjunction, which makes it structurally dependent on the main clause; it cannot stand on its own as an independent utterance. This is a clear example of (adverbial) subordination. In (4c) the first clause lacks tense marking and therefore is dependent on the second clause for the expression of tense. Accordingly, the first clause cannot stand on its own as an independent utterance. Moreover, it is neither an adverbial modifier of the second clause, nor is it an argument (complement) of the verb in the second clause, which rules out an analysis of it as subordination. However, it is clearly different from the coordination example in (4a) as well, and so it does not fit into either of the traditional categories; it

is, then, an instance of cosubordination at the clause level. It is important to note that while the first clause in (4c) is dependent on the second for the expression of tense, it is not embedded in it, unlike the adverbial subordinate clause in (4b) or the gerund in (2b). Thus, a crucial idea underlying the RRG theory of clause linkage is that dependent does not entail embedded; there can be formal dependencies between units in a flat structure.

The three nexus types can occur in nuclear, core and clausal junctures, generating nine possible juncture-nexus relations. Cosubordination at the nuclear level is illustrated in (1a), in which two nuclei, *qiāo* 'hit' and *pò* 'break', form a single complex nucleus under the scope of the *le* perfective aspect operator, aspect being a nuclear operator. Cosubordination at the core level is exemplified in (2a), which can be seen clearly when a deontic modal operator, a core operator, is added, as in (5a).

- (5) a. Max must try to fix his bicycle.  
 b. Max must persuade Bill to fix his bicycle.

What Max is obliged to do in (5a) is not to try anything but rather to try to fix his bicycle, which means that the scope of *must* is over both cores. In contrast, in (5b) Max is obliged to persuade Bill of something, but Bill is not obliged to fix his bicycle, which means that *must* has scope over only the first core but not the second; hence (5b) is not an example of core cosubordination but rather of core coordination.<sup>3</sup> Thus the structures in (5a, b) do not involve embedding, hence they are not examples of subordination, contra the conventional wisdom regarding these constructions (see Van Valin 2005:189-90 for evidence against an embedding analysis).

### 3. Critiques of cosubordination

Foley (2010) and Bickel (2010) attempt to call into question the validity of the notion of cosubordination. They restrict their arguments to clausal cosubordination only, and Foley assumes the original version of the LSC presented in Foley & Van Valin (1984), which differs in certain crucial respects from the version developed in Van Valin (1993b) and subsequent work. The notion of the LSC at that time was rather different from the

---

<sup>3</sup> It's important to emphasize here that 'coordination' is an abstract linkage relation and not a grammatical construction; it should be distinguished from 'conjunction', which is a formal construction type. Coordination may be instantiated by conjunction, as in (3), but it is not restricted to cases of formal conjunction.

concept that is assumed today; it is presented in Figure 1.



Figure 1: The LSC in Foley & Van Valin (1984:224)

As can be clearly seen, there is no ‘clause’ or ‘sentence’ level, and the periphery surrounds the core and nucleus. The shift to the current conception of the LSC began with Van Valin (1987) and was first published in Van Valin (1990, 1993b). There was no formal representation of the LSC, and in particular the projection grammar formalism had not yet been developed; it was proposed in Johnson (1987) and ushered in the representation of constituents, operators and information structure in distinct projections, also first published in Van Valin (1993b). Furthermore, in the 1984 book there was no representation of complex sentences beyond labeled bracketings, which did not include any representation of operators.

Their criticisms center on two key issues: first, there seem to be cases in which the scope of clausal operators is variable, as in (6) from Tauya (Papua New Guinea), and second, there are cases in which not all clausal operators are shared across the clauses, as in (7a) from Wambule (Sino-Tibetan) and (7b) from Korafe (Papua New Guinea).

- (6) *Tepau-fe-pa yate fitau-a-nae?* Tauya (MacDonald 1990)  
 break-TRANS-SS go throw-2-Q  
 a. ‘Did you break it and go away?’, or  
 b. ‘You broke it and did you go away?’, or  
 c. ‘Did you break it before going away?’

- (7) a. Wambule (Opgenort 2004) [Sino-Tibetan; cited in Bickel (2010:67)]  
*Na hep ja:ma-k tyaj iskul di-ŋ-m.*  
 previously grain eat.1sg-PAST-SEQ from.now school move-1sg-ASS  
 ‘I ate cooked grain before, and now I will go to school.’  
 b. Korafe (Farr 1999) [Papua New Guinea; cited in Bickel (2010:66)]  
*Mut-eno er-ira-re.*  
 give.1sg-SEQ-REALIS.1sgDS IPFV-go.DUR.PRES.3sgIND-CUR.REL  
 ‘I gave it, and he is currently going.’

The problem that both Foley and Bickel see in (6) is the apparently variable scope of the illocutionary force operator: it seems to have scope over

both clauses, yielding the reading in (6a), or only over the second clause, yielding the interpretation in (6b), or only over the first clause, yielding (6c). They interpret this as evidence against the notion of cosubordination, because it involves obligatory sharing of operators at the level of juncture, and here the operator sharing is variable and optional. The structure would be cosubordination on the (6a) reading but not on the (6b) or (6c) interpretations. Foley argues further that the examples in (7) are problematic, because not all clausal operators are shared across the two clauses: in (7a) there is one illocutionary force operator on the second clause, but the tense operator in the first clause has scope only over it, and the second clause gets a non-past interpretation; in (7b) there is again a single illocutionary force operator, but the first clause is interpreted as past tense, due to the sequential-realis affix, while the second clause is marked for present tense.

As noted above, Foley assumes the model of clause structure and the notion of cosubordination presented in Foley & Van Valin (1984), ignoring all subsequent work. In the 1984 version of the theory, no formalism of any kind had been developed, and in the informal representations used at that time, operator sharing was all or nothing. His proposed solution exploits an idiosyncratic feature of Lexical-Functional Grammar [LFG], namely the distinction between IP and S, where 'IP' contains grammatical categories like tense and illocutionary force and 'S' is a 'small clause' containing the predicate and its arguments. He claims that the contrast between coordination in e.g. (4a) and cosubordination in e.g. (4c) is a function of what is linked, not a difference in linkage relations. Hence there is coordination in (4a) between IPs but in (4c) between Ss. According to this analysis, what RRG calls 'cosubordination' is just coordination of Ss under one or more IP nodes, each reflecting a different grammatical category, and therefore cosubordination is not a distinct linkage type.

This alternative analysis of the phenomena which motivate the postulation of cosubordination does not call the notion of cosubordination into question. To begin with, it is limited to clause-level linkages; it does not apply to cosubordination at the core level, as in (2a), or at the nuclear level, as in (1a,b).<sup>4</sup> One would have to postulate something like a VP-level IP and a V-level IP in order to deal with these examples, and that is not an option in LFG. The result is a situation in which cosubordination is a linkage relation at sub-clausal levels but the analogous phenomena at the clause-level are handled in terms of a special type of coordination involv-

---

<sup>4</sup> See Bolmeyer & Van Valin (2017) for discussion of the importance of cosubordination in core junctures in relation to the Macro-Event Property.

ing Ss instead of IPs. It is difficult to see why this is an improvement over an analysis in which operator-sharing constructions are given a unified treatment at the clause level as well as sub-clausal levels.

Bickel (2010) starts out by saying that categories like ‘coordination’, ‘subordination’ and ‘cosubordination’ are too broad and not fine-grained enough to capture the diversity found in clause-linkage constructions. He decomposes the different constructions into a set of 11 features, each with a range of values. For example, ‘T[ense]-mark[ing]’ in dependent clauses has the values: **OK** (allowed), **Banned** (not allowed), **Harmonic** (allowed but subject to constraints based on the tense or status choice in the main clause). He then performs a statistical analysis to see if the features cluster into well-defined categories like coordination, subordination and cosubordination. He argues that this is not the case: there is tentative evidence for a specific prototype of subordination, but none for coordination and cosubordination, which seem to form a continuum. Given that coordination and cosubordination share the crucial feature of being a flat, i.e. non-embedded, structure, it appears that Bickel’s results reflect the salience of embedding as a feature of complex sentences. This is an interesting result, as there has been some debate within RRG as to which of the two defining features of nexus,  $[\pm \text{dependent}]$  and  $[\pm \text{embedded}]$ , is more basic. Van Valin (1993b) proposed that  $[\pm \text{embedded}]$  is the more basic feature, setting subordination off from coordination and cosubordination, which are then distinguished by  $[\pm \text{dependent}]$ . In Van Valin & LaPolla (1997), on the other hand,  $[\pm \text{dependent}]$  was taken as the basic distinction, with coordination being  $[- \text{dependent}]$  and the other two bring  $[+ \text{dependent}]$ ; subordination and cosubordination were then differentiated by  $[\pm \text{embedded}]$ . Bickel’s results support the 1993 analysis, not the 1997 one. This is summarized in Figure 2.

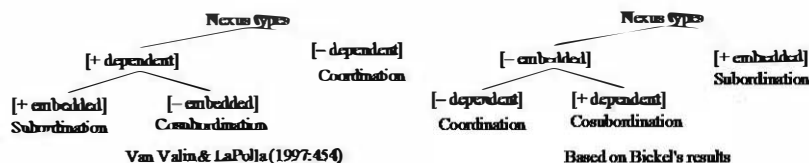


Figure 2: Nexus types

Bickel’s results can, thus, be seen as evidence of the significance of embedding in the structure of complex sentences and not as evidence against the validity of the notion of cosubordination.



#### 4. Re-examining cosubordination

Cosubordination was first proposed as a linkage type in Olson (1981), and it was further developed in Foley & Van Valin (1984). As noted above, it was assumed at that time that all of the operators at a given level of the clause must be shared in cosubordination, and that is the case in all of the examples presented in Foley & Van Valin (1984). Foley (2010:29) explicitly states that all operators must be shared.

In the decade after the publication of Foley & Van Valin (1984), however, work on complex sentences in a variety of languages, e.g. Mandarin Chinese (Tao 1986, Hansell 1993), Nootka (Jacobsen 1993), Japanese (Hasegawa 1992, 1996), and Turkish (Watters 1993), made clear that not all operators must be shared at the level of juncture. Rather, at least one must be shared, and the more that are shared, the tighter the link between the units. “[I]n a cosubordinate linkage at a given level of juncture, the linked units are dependent upon the matrix unit for expression of one or more of the operators for that level.” (Van Valin 1993b:112; see also Van Valin & LaPolla 1997:455, Van Valin 2005:201) In clausal junctures, illocutionary force, the outermost operator, must be shared; other clausal operators such as status and tense may or may not be shared. This can be seen clearly in the contrast between the Korafe example in (7b) and the Amele example in (8). The Amele example has the structure in Figure 3a; the Korafe sentence has the structure in Figure 3b.

- (8) *Ho busale-ce-b dana age ŋo-ig-a fo?* Amele (Roberts 1988)  
 pig run.out-DS-3sg man 3pl hit-3pl-TPAST Q [Papua New Guinea]  
 ‘Did the pig run out and did the men kill it?’  
 (\*‘The pig ran out and did the men kill it?’)

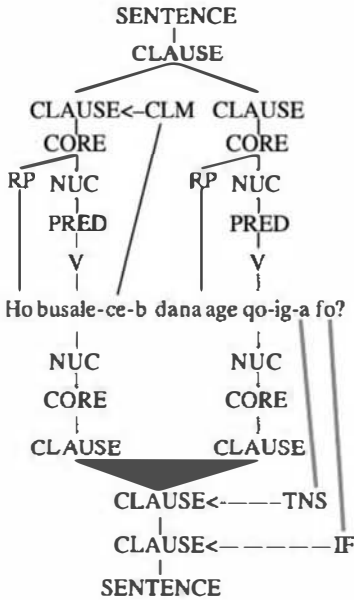


Figure 3a: Structure of (8)

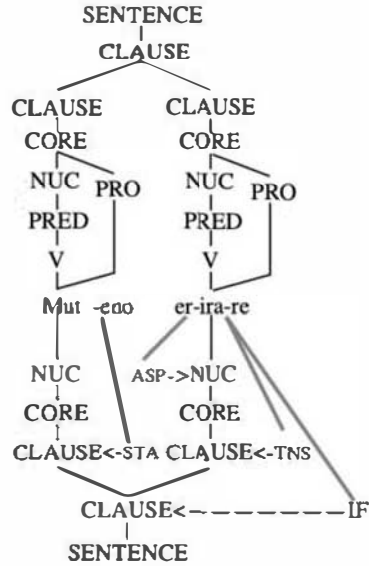


Figure 3b: Structure of (7b)

In Amele, both tense and illocutionary force must be shared in clausal cosubordination, and this is explicitly represented by having the operators modify the superordinate clause nodes. The motivation for having ‘duplicate’ clause nodes is that it is necessary to represent the scope of each operator separately, since they may or may not be shared. Korafe is just such an example: there are three clausal operators (status, tense and illocutionary force), and only illocutionary force is shared, with each clause having independent tense and status marking. Such a situation could not be captured in terms of the 1984 version of the LSC in Figure 1, but it can be readily expressed in terms of the RRG multiple projection representation. In sub-clausal junctures, at least one operator at the level of juncture must be shared; which operator that will be depends on the inventory of core and nuclear operators in the language.

When there is operator sharing in cosubordination, it must be obligatory, and in light of this requirement both Foley and Bickel point to the Tanya example in (6) as being extremely problematic for the concept of cosubordination. At first glance, it does indeed appear to be a counterexample to this requirement, but if we invoke another aspect of the theory not available in 1984, a solution readily presents itself. Foley himself (2010:47) points to the solution: there is only one IF operator (Figure 4) but the focus vs. presupposition division of the sentence varies (Figures 5a-c), which is represented in the focus structure projection in terms of the actual focus domain.

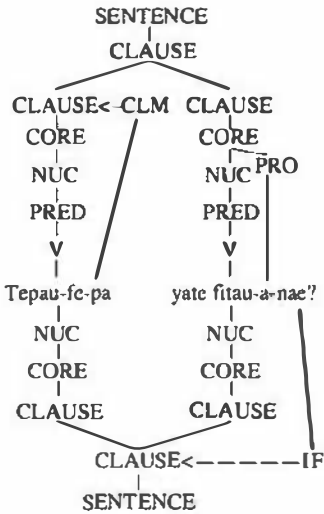


Figure 4: Tanya--Structure of (6)

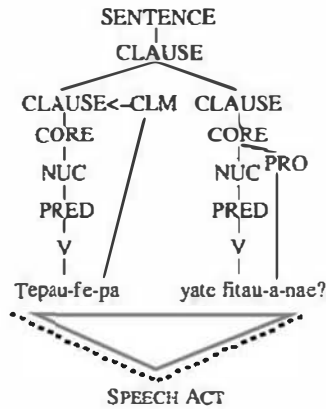


Figure 5a: AFD for (6a)

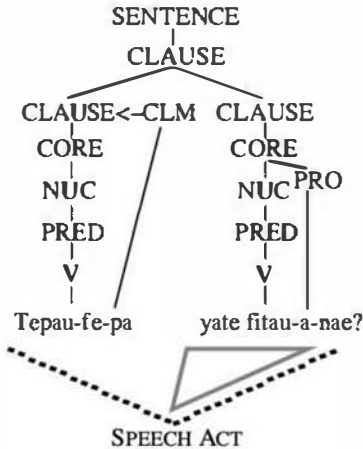


Figure 5b: AFD for (6b)

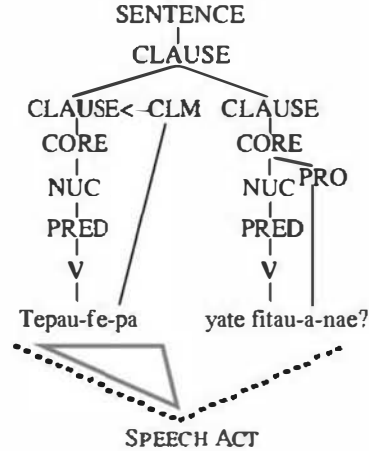


Figure 5c: AFD for (6c)

In Figure 4 the constituent and operator projections are given, showing that the illocutionary force operator *nae* ‘interrogative’ has scope over both clauses. This establishes that the potential focus domain is the entire sentence; it is represented by the dotted black line in the focus structure projection in Figures 5a-c. What varies is the actual focus domain [AFD], where the focus of the question lies. The reading in (6a), represented in Figure 5a, has both clauses within the AFD, as represented by the grey triangle. The one in (6b), on the other hand, reflects the AFD being limited to the second clause, the first one being presupposed; this is shown in Figure 5b. The most revealing interpretation is the one in (6c), in which the AFD includes the first clause but not the second, as given in Figure 5c. This reading is crucial evidence in favor of a cosubordination analysis, because if this were a typical coordinate construction, it would be impossible for an illocutionary force marker in the second clause to skip over the clause it occurs in and have just the first clause in its scope.<sup>5</sup> Rather, the AFD includes only the first clause, the second one being presupposed, and the scope of the question operator is the entire sentence. A complete analysis of the Tauya construction involves all three projections of the LSC. Thus, the Tauya example in (6) turns out to be strong evidence in favor of

<sup>5</sup> As Foley (2010:47) notes, it would be impossible to represent such a situation using the conjoined Ss under IP analysis.

a cosubordinate analysis and against a coordinate analysis.

Another example of variation in operator scope cited by Bickel (2010:61) is given in (9) from Belhare.

- (9) *Kimm-e n-ta-ch-u ki mun n-dny-chi.* Belhare  
 house-LOC 3ns-reach-1l-3sg SEQ 3dl 3ns-chat.NPST-1l [Tibeto-  
 a. ‘They will reach home and chat,’ or Burman, Nepal]  
 b. ‘When they reach home, they’ll chat,’ or  
 c. ‘They reached home and now they will chat.’

Bickel notes “the scope [of the main clause tense marker-RVV] is primarily limited to the main clause, but can optionally be extended into the dependent clause” (2010:61). This seems to be a case of optional rather than obligatory operator sharing, but it is less of a problem than it appears. As mentioned earlier, in clausal cosubordination illocutionary force, the outermost operator at the clause level, must be shared across the units, and that is the case in (9), which is a statement. Tense may, as in (8), or may not, as in (7b), be shared, but the construction is still clausal cosubordination, due to the shared illocutionary force. The variation in the interpretation of the tense in (9) parallels the variation in the interpretation of the AFD in (6), and it is tempting to offer a similar analysis. It is, however, difficult to see why the interpretation of tense should be tied to variation in the AFD, since tense and focus are rather different notions and belong to distinct projections of the clause. What would be problematic would be variation like this in sub-clausal operators, e.g. variability in the interpretation of the scope of aspect marking in a nuclear juncture like (1a). I am aware of no such examples; this kind of variability seems to be found in clausal junctures only and only with operators other than illocutionary force. Accordingly, the revision proposed in Van Valin (2005:205), that in some languages cosubordination is characterized in terms of possible rather than obligatory operator sharing, is unnecessary.

## 5. Conclusion

Foley (2010) and Bickel (2010) raise important questions about the validity of the notion of cosubordination as a nexus relation in complex sentences. It has been argued that these questions can be answered satisfactorily within the contemporary version of RRG, based on the LSC and the projection grammar representation of it, on the post-1984 conception of cosubordination, and including the information structure component. Cosubordination has been an integral part of the description of clause link-

age in numerous languages; in addition to those mentioned earlier, they include Yaqui (Guerrero-Valenzuela 2006), Q'eqchi' Mayan (Kochelman 2003), Chechen (Good 2003), Kwaza (van der Voort 2004), and Kikuyu (Kihara 2017). It remains a valid and valuable concept in the analysis of complex sentences.<sup>6</sup>

## References

- Bickel, Balthasar. 2010. Capturing particulars and universals in clause linkage: A multivariate analysis. In Bril, ed., 51-101.
- Bohnemeyer, Jürgen and Van Valin, Robert. 2017. The Macro-Event Property and the layered structure of the clause. *Studies in Language* 41:142-197.
- Bril, Isabelle, ed. 2010. *Clause Linking and Clause Hierarchy*. Amsterdam: Benjamins.
- Farr, Cynthia. 1999. *The Interface between Syntax and Discourse in Korafe, a Papuan Language of Papua New Guinea*. Canberra: Pacific Linguistics.
- Foley, William. 2010. Clause linkage and nexus in Papuan languages. In Bril, ed., 27-50.
- Foley, William & Van Valin, Robert. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Good, Jeff. 2003. Clause Combining in Chechen. *Studies in Language* 27: 113-170.
- Guerrero-Valenzuela, Lilián. 2006. *The Structure and Function of Yaqui Complementation*. *Studies in Native American Linguistics* 54. Munich: Lincom Europa.
- Hansell, Mark. 1993. Serial verbs and complement constructions in Mandarin: A clause linkage analysis. In Van Valin, ed., 197-233.
- Hasegawa, Yoko. 1992. *Syntax, semantics, and pragmatics of TE-linkage*. Berkeley: University of California, Berkeley (PhD dissertation).
- . 1996. *A Study of Japanese Clause Linkage: The Connective TE in Japanese*. Stanford: CSLI.
- Jacobsen, William H. 1993. Subordination and cosubordination in Nootka: Clause combining in a polysynthetic verb-initial language. In Van Va-

---

<sup>6</sup>This research was supported in part by the German Research Foundation as part of CRC 991 'The structure of representations in language, cognition and science'. This paper was presented at the 2015 International Conference on Role and Reference Grammar at the Heinrich Heine University Düsseldorf.

- lin, ed., 235-74
- Johnson, Mark. 1987. A new approach to clause structure in Role and Reference Grammar. *Davis Working Papers in Linguistics* 2:55-59.
- Kihara, C. P. 2017. *Aspects of Gikūyū (Kikuyu) Complex Sentences: A Role and Reference Grammar Analysis*. Düsseldorf: Heinrich Heine University (PhD dissertation).
- Kochelman, Paul. 2003. The Interclausal Relations Hierarchy in Q'eqchi' Maya. *International Journal of American Linguistics* 69:25-48.
- MacDonald, Lorna. 1990. *A Grammar of Tauya*. Berlin: Mouton de Gruyter.
- Olson, Michael. 1981. *Barai Clause Junctures: Toward a Functional Theory of Interclausal Relations*. Canberra: Australian National University (PhD dissertation).
- Opgenort, Jean. 2004. *A Grammar of Wambule*. Leiden: Brill.
- Roberts, John. 1988. Amele switch-reference and the theory of grammar. *Linguistic Inquiry* 19:45-64.
- Tao, Liang. 1986. Clause linkage and zero anaphora in Mandarin Chinese. *Davis Working Papers in Linguistics* 1:36-102.
- van der Voort, Hein. 2004. *A Grammar of Kwaza*. (Mouton Grammar Library 30). Berlin: Walter de Gruyter.
- Van Valin, Robert. 1987. Recent developments in Role and Reference Grammar: The layered structure of the clause and juncture. (Editor's introduction) *Davis Working Papers in Linguistics* 2:1-5.
- . 1990. Layered syntax in Role and Reference Grammar. M. Bolkestein, J. Nuyts and C. Vet (eds.), *Layers and Levels of Representation in Language Theory: A Functional View*, 193-231. Amsterdam: John Benjamins.
- . ed. 1993a. *Advances in Role and Reference Grammar*. Amsterdam: Benjamins.
- . 1993b. A synopsis of Role and Reference Grammar. In Van Valin, ed., 1-164.
- . 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert & LaPolla, Randy. 1997. *Syntax: Structure, Meaning & Function*. Cambridge: Cambridge University Press.
- Watters, James K. 1993. An investigation of Turkish clause linkage. In Van Valin, ed., 535-60.

**PART 4.**  
**INFORMATION STRUCTURE**





# A ROLE AND REFERENCE GRAMMAR ACCOUNT OF ASPECTS OF THE INFORMATION STRUCTURE-SYNTAX INTERFACE IN TAGALOG

ANJA LATROUITE<sup>1</sup>

& ROBERT D. VAN VALIN, JR.<sup>1,2</sup>

HEINRICH HEINE UNIVERSITY DÜSSELDORF<sup>1</sup>

UNIVERSITY AT BUFFALO, THE STATE UNIVERSITY  
OF NEW YORK<sup>2</sup>

## Abstract

Tagalog, an Austronesian language spoken in the Philippines, exhibits a number of inversion constructions which show divergence from the default predicate-initial word order. Marked word order is assumed to signal a marked packaging of information and licensed by a certain discourse context. In this paper, we are interested in the syntax-information structure interface of the different inversion constructions and will address specific questions raised by constructions which may have more than one information-structural function and possibly also more than one syntactic base analysis.

## Keywords

Tagalog, inversion, displacement, voice, information structure

## 1. Introduction<sup>1</sup>

The interplay between discourse-pragmatics and morphosyntax is a topic of central importance in Role and Reference Grammar [RRG], and in this paper we investigate aspects of this interaction in Tagalog, an Austronesian language spoken in the Philippines. It has several inversion constructions which show divergence from the default predicate-initial word order. Marked word order is assumed to signal a marked packaging of information and is licensed by a certain discourse context. The analysis will focus on the syntax-information structure interface of the different inversion constructions and will address specific questions raised by constructions which may have more than one information-structural function and possibly also more than one syntactic analysis.

The discussion will proceed as follows. Section 2 presents the basic facts about Tagalog morphosyntax that are relevant to this discussion. The next section introduces the inventory of inversion constructions found in the language. In section 4 the location of the inverted element in the clause is investigated. The following section argues that there are four different structural positions that the inverted (displaced) element can appear in. Section 6 examines a puzzling set of data and proposes a solution for it, and the final section summarizes our findings.

## 2. Some basic facts about Tagalog

The default word order in Tagalog is predicate/nucleus-initial with variable ordering of the core arguments following it. A striking characteristic of Tagalog is its symmetrical, multiple-voice system. As shown in (1) a verb like *sulat* 'to write' can take a number of voice affixes which signal the thematic role of the argument which receives the nominative case marker *ang*. In (1a) it is the actor argument, in (1b) the theme argument and in (1c) the recipient argument. While there is a certain preference for the nominative marked argument to appear in sentence-final position, other

---

<sup>1</sup> This research was supported by the German Research Foundation as part of CRC 991 'The structure of representations in language, cognition and science'. This paper was presented at the Thirteenth International Conference on Austronesian Linguistics, Academia Sinica, Taipei, Taiwan, and at the 2015 International Conference on Role and Reference Grammar, Heinrich Heine University Düsseldorf. We would like to thank Reyal Panotes Palmero for sharing her native speaker intuitions with us.

word orders have also been observed (cf. Sauppe 2016). So far there are no studies on what drives non-default post-verbal argument phrase order, but the relative length of a given phrase may be one of the factors that play a role.<sup>2</sup>

- (1) a. *S<um>ulat ng liham sa lalake ang babae.*  
 <AV.RLS>write GEN letter DAT man N●M woman  
 ‘The woman wrote a letter to a/the man.’
- b. *S<in>ulat ng babae sa lalake ang liham.*  
 <UV.PFV>write GEN woman DAT man N●M letter  
 ‘A/the woman wrote the letter to a/the man.’
- c. *S<in>ulat-an ng babae ng liham ang lalake.*  
 <UV.PFV>write-L V GEN woman GEN letter N●M man  
 ‘A/the woman wrote the man a letter.’

Tagalog has a number of second position clitics that have been investigated since their discussion in Bloomfield (1917) in works such as Schachter & Otanes (1972), Schachter (1973), Kroeger (1993, 1998), Billings & Konopasky (2004), Billings (2005). Nominative pronouns like *siya* in example (2) usually appear in the second position of the sentence and for the most part behave like clitics, even if they are not monosyllabic and could possibly be stressed. The same holds for non-pronominal particles like *ba*, which are true clitics.

- (2) *T<um>a~tawa ba siya kay Maria ngayon.*  
 <AV.RLS>IPFV~laugh Q 3sgN●M DAT.PN now  
 ‘Is he laughing at Maria now?’

As Figure 1 (next page) shows, clitics in RRG do not necessarily all have the same status with respect to the constituent structure. Pronominal clitics like *siya* (3sg) are part of the constituent structure projection, while an illocutionary force clitic like *ba* (question particle) is part of the operator

<sup>2</sup> Abbreviations: ADV ‘adverb’, ASP ‘aspect’, AV ‘actor voice’, CLM ‘clause-linkage marker’, CPR● ‘clitic pronoun’, DAT ‘dative’, GEN ‘genitive’, IF ‘illocutionary force’, INV ‘inversion marker’, IPFV ‘imperfective’, IRR ‘irrealis’, LDP ‘left-detached position’, LNK ‘linker’, LV ‘locative voice’, NPI ‘negative polarity item’, N●M ‘nominative’, NUC ‘nucleus’, ●BL ‘oblique’, PFV ‘perfective’, PL/pl ‘plural’, PN ‘personal proper name’, PrCS ‘precore slot’, PRED ‘predicate’, Q ‘interrogative’, RLS ‘realis’, RP ‘reference phrase’, sg ‘singular’, S & ● ‘Schachter & Otanes 1972’, TNS ‘tense’, UV ‘undergoer voice’.

projection. The RRG representation is in line with a suggestion that has been put forward by Kaufmann (2010: 12), who noticed that pronominal clitics differ from non-pronominal clitics, in that pronominal clitics may not penetrate complex fronted phrases (3a), while non-pronominal clitics may (3b).

- (3) a. [Sa=dalawa=ba=ng malaki=ng palabas]=sila li~litaw?  
 QBL=two=QM=LNK big=LNK show=3P.NOM AV~appear  
 'Will they appear in two big shows?'  
 b. [Saun=pa=ba=ng panadero]=kayo b<um>i~bili ng tinápay?  
 where=still=QM=LNK baker=2P.NOM <AV.RLS>IPFV~buy GEN=bread  
 'From which other baker do you buy bread?'

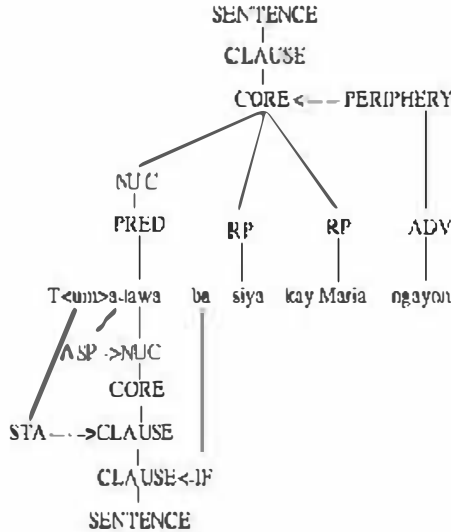


Figure 1. The layered structure of (2)

Kaufmann (2010: 12) refers to Halpern's (1995) theory of clitic placement, in which oblique clitics undergo prosodic inversion with the following prosodic word. These data are important as clitic position is mentioned as an important test for syntactic structure in RRG. However, if we distinguish between clitics that attach to the prosodic word and clitics that attach to syntactic units, then tests with the latter, but not necessarily with the former will be a good indicator for constituent structure and the positions

of elements in a tree. In the context of the clitic test it is equally important to note that some clitics may also appear freely, as noted in Table 1.

Table 1: Tagalog non-pronominal clitics and free counterparts

	CLITIC	FREE
aspect	= <i>na</i> ‘already’	∅
	= <i>pa</i> ‘still’	∅
focus	= <i>din</i> ‘also’	∅
	= <i>man</i> ‘even’	∅
	= <i>naman</i> ‘switch topic’	( <i>naman</i> )
	= <i>!à</i> ‘emphasis’	∅
	= <i>!a!</i> ‘only’ = <i>!áma!</i>	∅, <i>!áma!</i>
	= <i>talaga</i> ‘emphasis’	<i>talaga</i>
politeness	= <i>pò</i> , = <i>hò</i> ‘politeness’	∅
mood	= <i>pala</i> ‘surprise’	∅
	= <i>yátà</i> ‘perhaps’	∅
	= <i>sána</i> ‘hopefully’	<i>sána</i>
	= <i>náwa</i> ‘hopefully’	<i>náwa</i>
	= <i>ba</i> ‘question marker’ (= <i>baga</i> )	∅..( <i>baga</i> )
	= <i>daw</i> reported speech	∅

### 3. The inventory of inversion constructions in Tagalog

Schachter and Otanes (1972) mention a number of non-verb-initial constructions, so-called inversion constructions, that we will quickly introduce here. The first has been labeled ‘nominative inversion’, which may be a bit of a misnomer. In this construction the argument that would be in the *ang*-marked nominative in the default sentence pattern, as in (4a) and (9c), may

appear in the predicate position, while the deverbal argument phrase is marked by the reference phrase [RP] marker *ang* and appears in the default position for arguments. Note that Tagalog does not have auxiliaries. The structure in (4b) is basically an equative structure with specificational semantics (cf. Latrouite 2019).

(4) **Nominative inversion**

- a. *T<um>a-tawa siya.*  
 <AV.RLS>IPFV~laugh 3sgNOM  
 ‘He was laughing.’
- b. *Siya ang t<um>a-tawa.*  
 3sgNOM NOM <AV.RLS>IPFV~laugh  
 ‘He was the one who was laughing, It was he who was laughing.’

Another inversion construction is called ‘adjunct inversion’, ‘oblique inversion’, or ‘emphatic inversion’ (Schachter & Otnes 1972:496-98) and is exemplified in (5). Adverbials, adjuncts and oblique arguments may be moved to the sentence-initial position and attract both types of clitics in this case. The translation suggests a cleft-like semantics with narrow contrastive focus on the oblique phrase.

(5) **Adjunct/oblique inversion**

- a. *T<um>awa siya sa kaniya kahapon.*  
 <AV.RLS>laugh 3sgNOM DAT 3sgOBL yesterday  
 ‘She laughed at him yesterday.’
- b. *Kahapon siya t<um>awa sa kaniya.*  
 yesterday 3sgNOM <AV.RLS>laugh DAT 3sgOBL  
 ‘It was yesterday she laughed at him.’
- b’. *Sa kaniya siya t<um>awa kahapon.*  
 DAT 3sgOBL 3sgNOM <AV.RLS>laugh yesterday  
 ‘It was at him she laughed yesterday.’

More mysterious is the *ay*-inversion of arguments (6c) and oblique elements (6b). Schachter & Otnes (1972) note that the *ay*-inverted element may be topical or focal. Latrouite & Riestler (2018) observe that *ay*-inverted core arguments are often contrastive topics in the sense of Büring (2007), that is, topical with respect to a local subquestion under discussion, but focal with respect to the more global question under discussion.

(6) **Ay-inversion** (cf. S & ●:485-93)

- a. *Pu~punta kami bukas.*  
IPFV~go 1plNOM tomorrow  
'We'll go tomorrow.'
- b. *Bukas ay pu~punta kami.*  
tomorrow INV IPFV~go 1plNOM  
'Tomorrow, we'll go.'
- c. *Kami ay pu~punta bukas.*  
1plNOM INV IPFV~go tomorrow  
'(As for) Us, we'll go tomorrow.'

Another inversion construction has been labeled the 'contrastive inversion' construction and seems to be functionally similar to adjunct inversion. Formally, however, the construction resembles *ay*-inversion without the *ay*-particle: the adverbial phrase appears sentence initially followed by a pause instead of *ay* and does not attract pronominal clitics. The example (7b) shows two parallel sentences, in which the adverbial and the verb phrases are contrasted, i.e. a typical contrastive topic-contrastive focus construction.

(7) **Contrastive inversion** (cf. S & ●:493-96)

- a. *Mag-pa~pahinga kami bukas.*  
AV.IRR-IPFV~rest 1plNOM tomorrow  
*Dapat kami-ng mag-trabaho ngayon.*  
must 1plNOM-LNK AV.IRR-work today  
'We will rest tomorrow. We've got to work today.'
- b. *Bukas, mag-pa~pahinga kami.*  
tomorrow AV.IRR-IPFV~rest 1plNOM  
*Ngayon, dapat kami-ng mag-trabaho.*  
today must 1plNOM-LNK AV.IRR-work  
'Tomorrow, we will rest. Today, we've got to work.'

For the sake of completeness, we also mention 'adverbial inversion', which we do not have a lot to say about in this paper. Manner adverbials may appear sentence-initially before the predicate they specify, like in (8b), or sentence-finally, like in (8a). The position of the adverb is said to influence the scope. The relation to information structure is therefore not a direct one and warrants further exploration.



(8) **Adverbial inversion** ('non-emphatic inversion' [S & ●:498-500])

- a. *L<um>a~lapit ang bagyo sa Manila nang mabilis.*  
 <AV.RLS> IPFV~approach NOM storm DAT ADV rapidly  
 'The storm is rapidly approaching Manila.'
- b. *Mabilis na l<um>a~lapit ang bagyo sa Manila.*  
 rapidly LNK <AV.RLS>IPFV~approach NOM storm DAT  
 'Rapidly, the storm is approaching Manila.'

**4. Position of the inverted element: inside or outside of the clause?**

The different inversion constructions introduced above behave differently with respect to the placement of pronominal clitics; some inverted elements attract them, others do not. Since the pronominal clitics must appear in second position in the clause, the inverted elements that attract them must be in initial position in the clause. Accordingly, the inverted XPs that do not attract them cannot be in initial position in the clause. Is there any other evidence supporting this analysis?

**4.1 Negation**

A good test for ascertaining whether an inverted element is part of the clause is negation. Negation is typically characterized in terms of a contrast between 'internal' vs. 'external' negation. RRG treats negation as an operator which can modify any layer of the clause. Thus 'internal' negation is differentiated into nuclear negation (scope only over the nucleus) vs. core negation (scope over the core or a subpart of the core), and 'external' negation is clausal (scope over the entire clause). Negation is expressed primarily by *hindi* 'not', which normally occurs right before the material being negated. The following sentences show that the inverted element can be negated in adjunct inversion (9b), nominative inversion (9c) and adverbial inversion (9d). However, it is not considered grammatical for the *ay*-inversion (9e) and the structurally rather similar contrastive inversion (9f).

(9) **Negation of the inverted element:** (S & ●:499-500)

- a. *Hindi p<um>unta si Juan doon.* Default pattern  
 NEG <AV.RLS>go NOM.PN there  
 'Juan didn't go there.'

- b. *Hindi doon p<um>unta si Juan.* **Adjunct inversion**  
 NEG there <AV.RLS>go N●M.PN  
 ‘It wasn’t there that Juan went.’
- b’. *Doon hindi p<um>unta si Juan.*  
 there NEG <AV.RLS>go N●M.PN  
 ‘It was there that Juan **didn’t** go.’
- c. *Hindi si Juan ang p<um>unta doon.* **Nominative inversion**  
 NEG N●M.PN N●M <AV.RLS>go there  
 ‘It wasn’t Juan who went there,’ ‘The one who went there wasn’t Juan.’
- d. *Hindi palagi-ng nag-su~sumbbrero si Juan.* **Adv. inversion**  
 NEG always-LNK AV.RLS-IPFV~wear.hat N●M.PN  
 ‘Juan **doesn’t** always wear a hat.’ [Unambiguous]<sup>3</sup>
- e. *\*Hindi bukas ay pu~punta kami.* **Ay-inversion**  
 NEG tomorrow INV IPFV~go 1plN●M  
 ‘Not tomorrow, we’ll go.’
- f. *\*Hindi bukas, mag-pa~pahinga kami.* **Contrastive inversion**  
 NEG tomorrow AV.IRR- IPFV~rest 1plN●M  
*Ngayon, dapat kami-ng mag-pa~pahinga.*  
 today must 1plN●M-LNK AV.IRR-IPFV~rest  
 ‘Not tomorrow, we will rest. Today, we we must rest.’

## 4.2 Parameters of variation

Summing up the different morphosyntactic properties of the inversion constructions in question, we can state that:

- (i) *Ay*-inversion may, and contrastive inversion must, have a pause after the inverted element, whereas the others cannot have a pause.
- (ii) In adjunct/oblique (emphatic) and adverbial inversion, clitics are attracted to the displaced element but not in *ay*- or contrastive inversion.

<sup>3</sup> In sentence-final position the scope of the adverbial is ambiguous.

(i) *Hindi nag-su~sumbbrero si Juang palagi.*  
 NEG AV.RLS-IPFV~wear.hat N●M.PN Juan.LNK always  
 ‘Juan **doesn’t** always/always **doesn’t** wear a hat.’ [Ambiguous]  
 (Schachter & Oatnes 1972:500)

- (iii) The displaced element can be negated in nominative, adjunct (emphatic), and adverbial inversion, but not in *ay*- or contrastive inversion.

The obvious question raised by the data is then how to best capture and explain the observed differences. Addressing this question entails addressing the question whether all elements occurring in initial position are in the same structural position. If they are not, how many different initial positions do we have to distinguish and what are their characteristics? This question was already addressed by Nagaya (2007) for a couple of constructions. The main thrust of his paper is to argue (1) the clause-initial position is always part of the actual focus domain, and (2) the distinction between sentence and clause, which is crucial to the RRG analysis of syntactic structure but which is not recognized in some other approaches, is strongly supported by the Tagalog data. While we largely agree with Nagaya's structural analysis, an important focus in this paper is on *ay*-inversion, which he only mentions in a footnote. Section 5 introduces a puzzle with respect to negation in the *ay*-inversion construction (cf. Latrouite & Riester 2018, Nuhn 2019).

With respect to the interface between information structure and syntax, RRG is not a theory in which information structural features like focus and topic project syntactic structure; rather the basic idea is that information structure is a level of its own and can be mapped onto the constituent structure. If constructions like the *ay*-construction can have more than one information-structural interpretation, we need to think about how to capture and represent this. Do we wish to think about a certain construction as underspecified for its information-structural interpretation, with context enriching the interpretation so that the resulting reading can be derived? Or do we assume that there are two (or more) syntactic templates for a given inversion type and that context helps select which template is chosen? In the following we are going to address these questions.

## 5. Four different structural positions

Our claim is that the constructions in (4)-(9) involve four different structural positions.

## 5.1 Clause-initial element in nucleus position

In the construction in (4b), the initial position is the normal nucleus, following Nagaya (2007:349). The nominative expression is the predicate in the nucleus, and the core argument resembles a relative clause (see also Kroeger 1993) in that it contains a verbal predicate, the voice of which must be the same as that of the default patterns with the ‘inverted’ element in its canonical position. In other words, only the nominative argument of a verb can appear in a nominative inversion construction. In (1’) below the examples from (1) are presented as nominative inversion constructions.

- (1’) a. *Ang babae ang s<um>ulat ng liham sa lalake.*  
 NOM woman NOM <AV.RLS>write GEN letter DAT man  
 ‘The one who wrote a letter to a/the man was the woman.’
- b. *Ang liham ang s<in>ulat ng babae sa lalake.*  
 NOM letter NOM <UV.PFV>write GEN woman DAT man  
 ‘What the/a woman wrote to a/the man was the letter.’
- c. *Ang lalake ang s<in>ulat-an ng babae ng liham.*  
 NOM man NOM <UV.PFV>write-LV GEN woman GEN letter  
 ‘The one a/the woman wrote a letter to was the man.’

While there is no relative pronoun and while, synchronically speaking, *ang* does not have the formal or semantic properties of a (demonstrative) pronoun to justify analyzing it as ‘this one (who is running)’, Reid (2002) and others have argued that historically *ang* can be traced back to a demonstrative pronoun plus linker. The plausibility of this reconstruction is augmented by the fact that in spoken Tagalog the demonstrative pronoun *iyon* + linker (= *yung*) is often used to mark the nominative argument instead of *ang* (cf. Nagaya 2011), and in this usage it is showing signs of slowly losing its high degree of referentiality.

Since the predicate reference phrase (RP) is in the core-internal nucleus position, it can be negated, like any other nucleus, hence the grammaticality of (9c) is expected. The information-structural effect according to Nagaya (2007) is narrow focus on the nominative argument often combined with an exhaustive and contrastive reading. Functionally, the construction therefore resembles the English cleft and one of its information-structural interpretations (cf. Declerck 1988). Formally the construction lacks an expletive (like *it*-clefts) or *wh*-pronoun (like *wh*-clefts), an auxiliary and the relative pronoun. The RRG representation of the specificational predication in Figure 2 shows that we are still dealing with a bi-clausal

structure (cf. Latrouite 2019), where the predicatively used RP *siya* ‘he’ in the predicate slot provides the value for the variable expressed by the argument phrase, which consists of a verb in actor voice that gets an Actor-denoting reading when following the case particle *ang*.

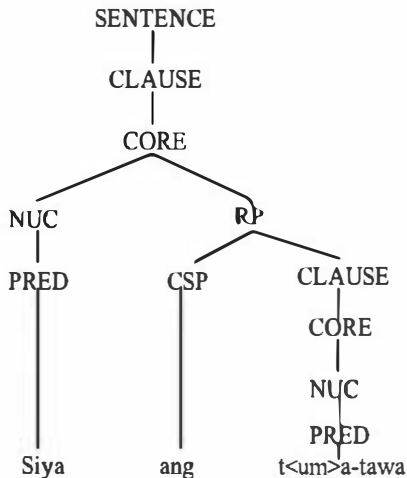


Figure 2: Structure of nominative inversion in (4b)

As mentioned earlier, there are no functional projections in RRG. Functional elements tend to be part of the operator projection or to be assigned via rules like the case particle (CSP) *ang*. The particle *ang* is a reference phrase marker and as such a referentiality operator within the RP. It has, furthermore, case marker functionality, i.e. if there is more than one argument in a predicate-initial default sentence, then only one of the arguments may be marked by *ang* and that is the argument signaled on the voiced verb. The particle is also the default RP marker in equational sentences without voiced verbs. We are not going into great detail of the RP structure in Tagalog in this paper. The crucial thing is that the voice affix, which is a clausal status marker (active-affirmative) clearly shows that the RP contains a clause. Clauses can be negated, so it is possible for the negation marker *hindi* to appear right before the predicate *tumatawa* ‘laughing’ (*Siya ang hindi tumatawa*. ‘He is the one not laughing.’) *Ang* as a definiteness operator inside the RP in its operator projection is not reflected in Figure 2.

## 5.2. Clause-initial element in the PrCS: adjunct inversion

The adjunct/oblique (emphatic) inversion in (5) complements the construction we have just discussed, because it targets adjuncts (5b), adjunct question words (12), and oblique core arguments (5c), exclusively. The clause-initial phrase is in the pre-core slot (PrCS) in these constructions. The characteristics of the PrCS are:

1. It precedes the core.
2. The contrastive focus reading on the displaced phrase makes it the focus of the assertion, and this requires that it be within the scope of the IF operator, hence within the clause.
3. Pronominal clitics which must occur in second position in the clause follow the displaced phrase, as in (5b, b'), not the nucleus as in (5a); hence it must be clause-initial.
4. No pause can follow the displaced phrase.

Given that the PrCS phrase is the focus of the assertion, it can be negated, hence the grammaticality of (9b). In terms of information structure the reading we get tends to be narrow focus (Nagaya 2007:353). There is no corpus study as of yet on this construction so that we cannot add more in terms of its information-structural function. The structure is represented in Figure 3. The representation shows that the construction is monocausal with two clausal operators operating over it, the illocutionary force operator (IF), expressed by the sentence structure and the lack of a question particle, and the status operator (STA), which is reflected in the verb; it is not negated and thus affirmative-active (actor) voice.

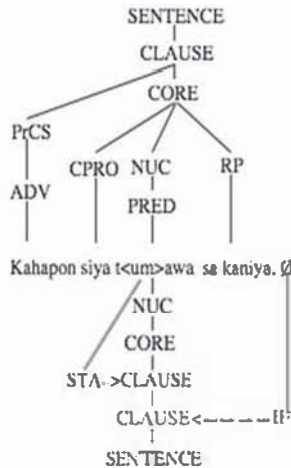


Figure 3: The structure of (5b)

### 5.3 Clause-initial element in a core-level periphery: adverbial inversion

The displaced adverbial in non-emphatic inversion in (8) is in clause-internal position, because it attracts the second position pronominal clitics, as shown in (10), and it cannot be followed by a pause.

- (10)a. *P<in>asok nila ang silid nang bigla* (S & O:498)  
 <UV.PFV>enter 3plGEN NOM room LNK suddenly  
 ‘They suddenly entered the room’, ‘They entered the room suddenly.’
- b. *Bigla nila-ng p<in>asok ang silid.*  
 suddenly 3plGEN-LNK <UV.PFV>enter NOM room  
 ‘Suddenly, they entered the room.’

These facts suggest the hypothesis that the adverbial is in the PrCS. Schachter & Otnes (1972:498) suggest that “non-emphatic inversion has no effect upon the meaning of the sentence.” This is certainly not true, as already indicated by the English translation. Non-emphatic inversion only has an evaluative reading, while the default sentence allows for a pure manner reading of the adverb as well as an evaluative reading. As there is

no information-structure related difference, we assume that the adverbial is in a core-level periphery which precedes rather than follows the core.<sup>4</sup> This assumption is confirmed by (11b), in which adjunct (emphatic) displacement co-occurs with adverbial (non-emphatic) displacement; the PrCS is filled by the displaced adjunct, not the displaced adverbial.

- (11)a. *Nag-trabaho ang lahat nang mabilis kanina-ng umaga.* (S & O:499)  
 AV.RLS-work NOM all LNK quickly this-LNK morning  
 ‘Everyone worked quickly this morning.’
- b. *Kanina-ng umaga mabilis na nag-trabaho ang lahat.*  
 this-LNK morning quickly LNK AV.RLS-work NOM all  
 ‘It was this morning that everyone worked quickly.’

Because the adverbial is in the clause, it can be negated, as shown in (9d).

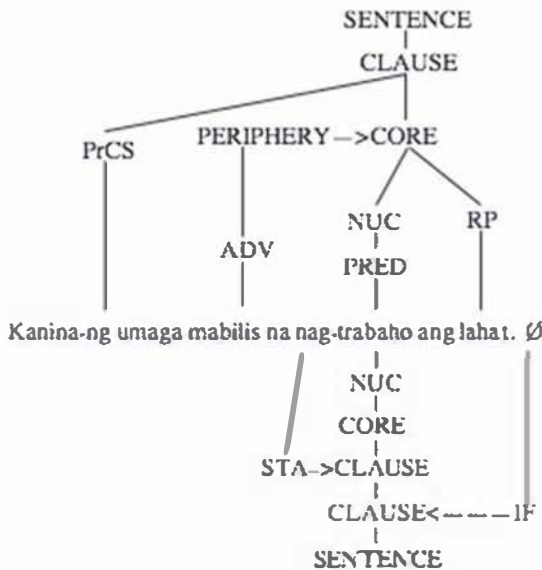


Figure 4: The structure of (11b)

<sup>4</sup> Analyzing certain preposed adjuncts as being in the core-level periphery rather than the PrCS or the LDP was first proposed in Bentley (2008).



## 5.4 Sentence-initial element in the left-detached position: *ay*-inversion

As already mentioned, in *ay*-inversion and contrastive inversion in (6) and (7), the sentence-initial phrase may be a nominative RP, an oblique core argument or an adjunct. In this construction, we have evidence that it is placed within the LDP, as also suggested by Nagaya (2007:361 ff). One bit of evidence is that an *ay*-marked phrase always precedes elements in the PrCS, whenever they cooccur as in (12).

- (12) *Si May ay kailan ba ba-balik dito?*  
 NOM.PN INV when Q IPFV~return here  
 ‘As for May, when will [she] come here?’

Another bit of evidence is that WH-expressions in a question can never be marked by *ay*, nor is the *ay*-construction a good candidate for marking the information focus in the answer to a narrow argument question (Kaufman 2005, Dery 2007) like ‘Who will come back?’ (cf. 13b). Nuhn (2019) suggests that a different construction is preferred in such cases. Our consultant indeed prefers a construction termed ‘reversed *ang*-inversion’ (Nuhn 2019), exemplified in (13e), over a default verb-initial structure (13d) and a nominative inversion construction (13c).

- (13) a. *Sino ang ba-balik bukas?*  
 who.NOM NOM IPFV~come.back tomorrow  
 ‘Who is the one coming back tomorrow?’  
 b. #*Si May ay babalik.*  
 NOM.PN INV  
 c. ??*Si May ang babalik.*  
 d. ?*Babalik si May.*  
 e. *Ang babalik ay si May.*  
 NOM INV  
 ‘May is the one who came back.’

The semantic difference between the four answers concerns the presuppositions. (11b) has been said to be out because the referent is marked as topical here via the marker *ay*. Latrouite & Riester (2018) note that the *ay*-inversion is not only used to mark givenness topics, but also in partial answers to a given question under discussion, i.e. in order to signal a contrastive topic in the sense of Büring (2007). The English ‘as for’ translation of the *ay*-construction in (12) is a typical contrastive topic

construction: *As for Mary, she will come back tomorrow, I do not know about John, Peter, Harry, Susan coming back.*, etc. By choosing the answer in (13b) the speaker, thus, expresses either that Mary is given, which would be clearly wrong, or s/he expresses that s/he assumes or is aware that the questioner is interested in other people on top of Mary. If such an answer is marked as awkward by native speakers, it is because they need to accommodate quite a bit to make this answer felicitous. The answer in (13c) is not out, but also very marked. In Latrouite & Riester (2018) it is noted that the nominative-inversion construction is often chosen in contexts in which the choice of the referent in question is not expected, either based on the immediate common ground or based on prototypical scenarios associated with the general common ground. So once again, the answer seems to be rather marked and requiring some additional assumptions on the part of the answerer. The structure of the answer in (13d) does not reflect any special information packaging in terms of focus-background structure. The verb-initial structure is a typical choice for athetic sentence (all-new) or a predicate focus construction (Cf. Latrouite & Riester 2018). We assume that this is the reason why (13e) is the preferred choice here. In this construction the givenness of the predicate phrase is signaled and the focal argument is marked as such. The label ‘reversed *ang*-inversion’ not only signals the reversed syntactic order of the constituents but also the reversed information-structural packaging.

A third piece of evidence for the LDP position of the *ay*-marked phrase lies in the fact that it may be accompanied by or replaced by a pause (S & ●: 489, Nagaya 2007: 365-6), as in (14). According to S & ●, the use of *ay* reflects a more formal style. The initial XP in contrastive displacement must be followed by a pause (S & ●: 493). The notion contrastive displacement already points to the contrastive topic usage we mentioned above as well as to the external status of the phrase.

- (14) *Si May, kailan ba (siya) ba~balik dito?*  
 NOM.PN            when Q (3sgNOM) IPFV~return here  
 ‘As for May, when will (she) come here?’

As (14) shows, the potentially *ay*-marked phrase does not attract second-position clitics, neither pronominal nor non-pronominal clitics. Rather the clitics *ba* and *siya* attach to the PrCS containing *kailan* ‘when’, as would be expected. The inverted argument may be taken up by a resumptive pronoun, as in (14). D. Kaufmann (pc.) pointed out that this holds only, if the argument is followed by a pause, not by *ay*. In our corpus, we did not have examples with resumptive pronouns for either construction. Therefore, it

is still an open question when and under which conditions a resumptive pronoun may appear. Crucially, there is never a resumptive pronoun for the inverted argument in the nominative inversion construction.

Another important point is that there can be more than one *ay*-marked phrase in a sentence (S&O, Kaufman 2005); they may occur in any order, as in (15). Contrastive and *ay*-marked displacement can co-occur in a single sentence, but the order is not free: the *ay*-marked phrase must follow the contrastive phrase, as in (16).

- (15) a. *Pu~punta kami bukas.*  
 IPFV~go 1plNOM tomorrow  
 ‘We’ll go tomorrow.’  
 b. *Kami’y bukas ay pupunta.*  
 ‘As for us, tomorrow, we’ll go.’  
 b’. *Bukas ay kami’y pupunta.*  
 ‘Tomorrow, as for us, we’ll go.’

- (16) a. *Bukas, kami’y mag-pa-pahinga.*  
 tomorrow 1plNOM.INV AV-IPFV~rest  
 ‘Tomorrow, as for us, we’ll rest.’  
 b. *Kami, bukas ay mag-pa-pahinga.*  
 1plNOM tomorrow INV AV-IPFV~rest  
 ‘We, tomorrow, we’ll rest.’

Finally, possessor raising to the LDP is possible, as shown in (17).

- (17) *Si Jose ay na-matay ang asawa.*  
 NOM.PN INV UV-die NOM wife  
 ‘As for Jose, [his] wife died.’

An LDP phrase is clause-external and therefore outside of the scope of the IF-operator, and accordingly it cannot be negated, hence the ungrammaticality of (9e,f).

As already mentioned apart from the contrastive topic usage that we mostly get for core arguments, there is also a frame-setting topic usage of the marker *ay*, e.g. (16b).<sup>5</sup>

<sup>5</sup> See Nuhn (2020) for a detailed investigation of the information-structural properties of *ay*-inversion.

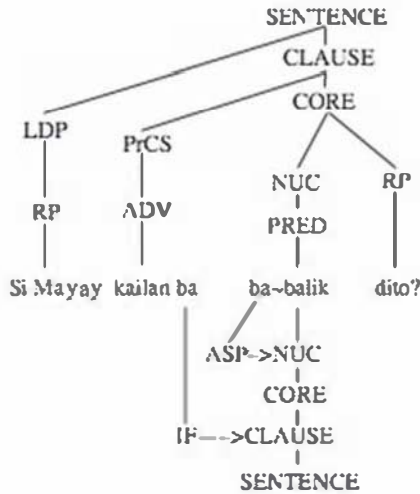


Figure 5: The structure of (14)

### 5.5 Interim summary

Table 2 sums up what we have said so far about the constructions.

Table 2: Interim summary of inversion constructions and their functions

Construction	Location of XP	Function
Nominative inversion	Nucleus	Argument narrow focus
Adjunct inversion	PrCS	Adjunct/oblique narrow focus
Ay-inversion	LDP	(Frame-setting) Topic
Contrastive inversion	LDP	Contrastive topic
Adverbial inversion	PeripheryCORE	(Frame-setting) evaluative usage

We have identified four structural positions which can be the first element in either a clause or a sentence: the default nucleus position, as in (2) and

(4b), Figures 1 and 2, respectively; the pre-core slot, as in (5b), (11b), and (14), Figures 3, 4 and 5, respectively; the left-detached position, as in (5b) and (14), Figures 3 and 5, respectively; and the core-level periphery, as in (11b), Figure 4. These are structural positions which can, in principle, serve a variety of information-structural functions. It was mentioned earlier that Nagaya (2007) argued that the clause-initial position in a Tagalog clause is always part of the actual focus domain. We may have identified two exceptions to this generalization, namely adverbial inversion and reversed *ang*-inversion. In cases of adverbial inversion, e.g. (8b), (9a) (10b), the adverbial can be first in the clause, as in these examples, and it's not clear that it is within the actual focus domain of the utterance, especially given its (frame setting) evaluative interpretation. Further investigation is necessary here. The situation is clearer with reversed *ang*-inversion, as in (13e): the focus in (13e) is *si May*, the answer to the question in (13a), and actual focus domain does not include *ang babalik* 'the one who returned', which is part of the immediate common ground because of the question. This construction requires further investigation, both to confirm its information-structural properties and to uncover its syntactic structure. Analyzing *ang-babalik* as being in the LDP with the following clause consisting entirely of *si May*, does not strike one as a plausible analysis. This suggests that *ay* is a multifunctional inversion marker.

## 6. A puzzle

An interesting puzzle remains, and it is not unrelated to the issue that the previous section concluded with.

### 6.1 The issue: *ay*-iuverson and negation

The inverted phrase in *ay*-inversion appears to be in the LDP, which is outside the scope of the IF operator, and this accounts for why it cannot be negated with *hindi* 'not', as in (9e). However, there are cases in which negative polarity items (NPI) occur in the *ay*-marked phrase, and the licensing negative occurs in the following clause, as in (18) and (19) (S & ●:492). (18a) shows a predicate initial sentence containing a particle *ni* that induces the reading 'even'.

- (18) a. *Hindi nag-dala si Rosa ni lapis/anuman.*  
 NEG AV.RLS-bring N●M.PN even pencil/anything  
 'Rosa did not bring even a pencil/anything at all.'

- b. *Ni lapis/anuman ay hindi nag-dala si Rosa.*  
 even pencil/anything INV NEG AV.RLS-bring NOM.PN  
 ‘Rosa did not bring even a pencil/anything at all.’

Note that the *ay*-inverted example contains the same focus sensitive particle meaning *even*. Focus sensitive particles like *even* and *also* stress that the focal element is not exclusive and exhaustive with respect to the assertion. For this reason, *ay*-inversion rather than *ang*-inversion is observed with these particles. Interestingly, with these particles, a focal undergoer may appear in-situ, i.e. sentence-finally or *ay*-inverted. A focal actor, however, may not be realized in-situ; it must occur in a displaced position. This observation goes well with the idea that the default status for undergoers is focal, and for actors it is topical.

- (19) a. \**Hindi ma-bu~buhat ni ang Tatay ito.*  
 NEG able.UV-IPFV~lift even NOM father this.NOM  
 ‘Not even Father can lift this.’  
 b. *Ni ang Tatay ay hindi ma-bu~buhat ito.*  
 even NOM father INV NEG able.UV-IPFV~lift this.NOM  
 ‘Not even Father can lift this.’

NPIs must occur within the scope of negation, and it has long been known that the scope of negation falls on the focus of an utterance (hence the use of negation as a test for focus in the information-structure literature.<sup>6</sup> The puzzle is: how can a NPI occur in the LDP, which is outside of the assertable and therefore the deniable part of a sentence? This should not be possible. Indeed, how can a NPI occur in a position which cannot be directly negated? All of the morphosyntactic evidence points unambiguously to the inverted *ay*-marked phrase being in the LDP, yet in these forms it has properties which are associated with the asserted part of the sentence, a part which does not include the LDP.

## 6.2 A possible solution

As a first step toward possible solutions to these puzzles, it is necessary to question two assumptions inherent in the discussion in the previous section. The first is that the only option for handling displaced constituents is the LDP. The second is that all cases of *ay*-inversion are basically the

---

<sup>6</sup> See Erteschik-Shir (2007) for an overview.

same construction, i.e. *ay* always marks a topical phrase in the LDP. (18) and (19) are clearly cases in which an element within the scope of the IF operator, i.e. inside the potential focus domain, is marked by *ay*.

In all of the standard cases involving *ay*, the inversion is optional, subject to information-structural constraints. The inversion in (19b), however, is not optional: when the actor is marked by the NPI *ni*, it is *obligatorily displaced*. As noted above, this is true of focal actors in general. This suggests strongly that the *ni-ay* construction should be treated as distinct from the optional *ay*-LDP construction. If this is correct, then it is no longer necessary to assume that the *ay*-RP in (18b) and (19b) is in the LDP. If it's not in the LDP, then what position is it in? In this structure the negative *hindi* has wide scope, and this is represented by it being a modifier of the superordinate clause node in the operator projection. IF also has scope over the entire clause, as (20) shows.

- (20) *Ni ang Tatay ay hindi ba ma-bu~buhat ito?*  
 even N●M father INV NEG Q able.UV-IPFV~lift this.N●M  
 'Couldn't even Father lift this?'

Thus, all of the evidence points to the *ay*-marked RP as being in the PrCS. This structure accounts for the problematic properties of the *ni-ay* construction, including the somewhat contrastive interpretation of the *ni*-RP.

There is a complication, however. Consider (21) from Nuhn (2020).

- (21) *Bakit kahit si Pedro ay hindi ma-bu~buhat ito?*  
 why even N●M.PN INV NEG able.UV-IPFV~lift this.N●M  
 'Why couldn't even Pedro lift this?'  
 'Why is it that even Pedro couldn't lift this?'

We have argued that in the *ni-ay* construction, the *ay*-marked RP is in the PrCS, and we have analyzed adjunct-inversion as also involving the PrCS. Given that the PrCS cannot be doubly filled, *bakit* 'why' and *si Pedro* cannot both be in the PrCS. Moreover, there would seem to be no other structural position available for *bakit* 'why', since as a question word it cannot occur in the LDP, which is outside of the scope of the IF operator over the clause. There is, however, another possibility, which was introduced earlier in a slightly different form. We argued that in adverbial inversion the displaced adverbial occurs in the core-level periphery (see Figure 4) which precedes the core. The adverbials discussed there, as well as the adjunct question words meaning 'where', 'when' and 'how', are all core-level elements. *Bakit*, on the other hand, questions the reason(s) for

some state of affairs, and as argued in Van Valin (2005:194-5) reason expressions, e.g., clauses introduced by *because* or *because of* RP PPs in English, occur in the *clause-level* periphery. It is therefore possible to apply the idea behind adverbial inversion on the core-level to *bakit* on the clause-level and analyze *bakit* in (21) as being clause-initial in the clause-level periphery. It precedes the PrCS, as (21) shows, and follows the LDP constituent *at si Pedro* ‘and [as for] Pedro’, as in (22) (Reyal Panotes Palmero, p.c.).

- (22) *At si Pedro, bakit naman hindi niya kaya-ng buhat-in ito?*  
 and NOM.PN why OPP NEG 3sgGEN afford-LNK lift-UV this.NOM  
 ‘And Pedro, why can’t he (in contrast to others) lift this?’

This predicts that the core-level adjunct question words mentioned above cannot occur in this position, and Nuhn (2020) confirms that this is the case. There are thus five positions we have identified as being involved in the constructions examined in this paper: LDP – Clause-Periphery – PrCS – Core-Periphery – Nucleus.

Coming back to our question with respect to the relationship between information structure and constituent structure, our answer with respect to the *ay*-construction is clearly that we assume two different constructional schemas, one for frame-setting topics and one for the contrastive topic function. However, such an approach may not be suitable for the different IS-readings that are found for the nominative inversion (Latrouite 2019). This leads to the insight that while some IS-notions in Tagalog may be considered as structure-building, others are possibly best viewed as purely pragmatic (cf. Krifka and Musan 2012). In other words, there can be syntactic structures that are specialized for specific functions, e.g. the English PP-initial presentational construction, while others, e.g. the English RP-(AUX)-V-RP-(RP/PP) default pattern, are compatible with a range of information-structural interpretations.



## 7. Final summary

Table 3: Final summary of inversion constructions and their functions

Construction	Location of XP	Function
Nominative inversion	Nucleus	Argument narrow focus
Adjunct inversion	PrCS	Adjunct/oblique narrow focus
<i>Ay</i> -inversion	LDP	(Frame-setting) Topic
<i>Ni-ay</i> construction	PrCS	Contrastive focus
Contrastive inversion	LDP	Contrastive topic
Adverbial inversion	Periphery <sub>CORE</sub>	(Frame-setting) evaluative usage
<i>bakit</i> ‘why’ question	Periphery <sub>CLAUSE</sub>	Reason question
Reversed <i>ang</i> -inver.	Nucleus[?]	Completive narrow focus

We have identified five structural positions that are relevant to information-structural functions. Some of the positions identified (nucleus, core-level and clause-level peripheries) are structural notions not associated with specific information-structural functions, whereas others (PrCS, LDP) seem to have such a function or range of functions. The syntactic structure of the reversed *ang*-inversion construction (Nuhn 2019) is not clear, and in Table 3 the location of the inverted XP is stated as the nucleus, which is based on a simplistic analysis, in the absence of a more detailed investigation.

It might well be objected, in view of the multitude of distinctions that the Cartographic approach (Rizzi 1997) postulates, is five enough? The answer seems to be ‘yes’. Bentley (2008) presents an RRG account of the Italian left-periphery in terms of the RRG distinctions, and Shimojo (2011) does the same for Japanese.

## References

- Bentley, Delia. 2008. The interplay of focus structure and syntax: Evidence from two sister languages. In Van Valin, Robert (ed.), *Investigations of the Syntax-Semantics-Pragmatics Interface*, 263-284. Amsterdam: John Benjamins.
- Billings, Loren A. 2005. Ordering clitics and postverbal R-expression in Tagalog: A unified analysis? In Carnie, Andrew, Heidi Harley, and Sheila Ann Dooley (eds.), *Verb First: On the Syntax of Verb-initial Languages*, 303-39. (Linguistik Aktuell, 73.) Amsterdam: Benjamins.
- Billings, Loren and Abigail Konopasky. 2002. Morphology's role in ordering verb-adjacent clitics. In Mikkelsen, Line and Christopher Potts (eds.) *The proceedings of the Twenty-first West Coast Conference on Formal Linguistics*, 29-42. Somerville, MA: Cascadia Press.
- Billings, Loren and Abigail Konopasky. 2003. Reassessing the role of syntax inside the morphological word: verb-adjacent clitics in Tagalog and Bulgarian. In Anastasia Riehl, and Thess Savella (eds.), *Proceedings of AFLA 9*, Ithaca, NY, 2002. Cornell Working Papers in Linguistics 19.
- Billings, Loren, and Daniel Kaufman. 2004. Towards a typology of Austronesian pronominal clisis. *ZAS Papers in Linguistics* 34, 15-29. (Proceedings of AFLA 11.)
- Büring, Daniel. 2007. Intonation, semantics and information structure. In Ramchand, Gillian and Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, 445-74. Oxford: Oxford University Press.
- Declerck, R. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Dordrecht: Foris.
- Dery, Jeruen. 2007. Pragmatic focus and word order variation in Tagalog. *Language and linguistics*, 8:373-402.
- É. Kiss, K. 2006. Focus as predication. In Molnar, V. and S. Winkler, (eds.), *The architecture of focus*, 169-93. Berlin: Mouton de Gruyter.
- Erteschik-Shir, Nomi. 2007. *Information structure: the syntax-discourse interface*. Oxford: Oxford University Press.
- Foley, William and Robert Van Valin, Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford: CSLI.
- Kaufman, Daniel. 2005. Aspects of pragmatic focus in Tagalog. In: Arka, I.W. and M. Ross (eds.), *The many faces of Austronesian voice systems*, . Canberra: Pacific Linguistics.
- Kaufman, Daniel. 2010. *The morphosyntax of clitics*. Ithaca: Cornell University (PhD dissertation).

- Krifka, Manfred and Renate Musan. 2012. Information structure: overview and linguistic issues. In Krifka, M. and R. Musan (eds), *The expression of information structure*, 1-44. Berlin: De Gruyter.
- Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI.
- Kroeger, Paul. 1998. Clitics and clause structure in Tagalog. In Maria Lourdes Bautista (ed.), *Pagtanaw*, 1-44. Manila: DSLU Press.
- Kuno, S. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.
- Latrouite, Anja. 2019. Specificational predication and cleft-like constructions in Tagalog and Japanese. To appear.
- Latrouite, Anja and Katalin Balogh. 2016. Approaching the morphosyntax-discourse interface: A closer look at Tagalog and Hungarian. Talk at the Cognitive Structures Conference, Heinrich-Heine-Universität Düsseldorf.
- Latrouite, Anja and Arndt Riester. 2018. The role of information structure for morphosyntactic choices in Tagalog. In Riesberg, Sonja and Asako Shiohara (eds.), *Perspectives on information Structure in Austronesian languages. Studies in diversity linguistics*, 247-284. Berlin: Language Science Press.
- Nagaya, N. 2007. Information structure and constituent order in Tagalog. *Language and linguistics* 8:343-72.
- Nagaya, Naonori. 2011. Rise and fall of referentiality: articles in Philippine languages. In Foong Ha Yap, Karen Grunow-Härsta and Janick Wrona (eds), *Nominalization in Asian Languages: Diachronic and typological Perspectives*, 589-626. Amsterdam: John Benjamins.
- Nuhn, Patrick. 2019. Reversed *ang*-inversion and narrow focus marking in Tagalog. *Studies in Language* 43(2):469-497.
- Nuhn, Patrick. 2020. *Aspects of information structure and morphosyntax in Tagalog*. Düsseldorf: Heinrich Heine Universität (PhD dissertation).
- Ramos, Teresita. 1971. *Tagalog structures*. Honolulu: University of Hawaii Press.
- Reid, Laurence. 2002. Determiners, nouns or what? Problems in the analysis of some commonly occurring forms in Philippine languages. *Oceanic Linguistics* 41(2):295-309.
- Rizzi, Luigi 1997. The fine structure of the left periphery. In L. Haegeman (ed), *Elements of grammar*, 281-337. Dordrecht: Kluwer.
- Sauppe, S. 2016. Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences. *Frontiers in Psychology*, 7(95). doi: 10.3389/fpsyg.2016.00095

- Schachter, P. & F. Otanes. 1972. *Tagalog reference grammar*. Berkeley: University of California Press.
- Schachter, Paul. 1973. Constraints on clitic order in Tagalog. In: A. B. Gonzalez (ed.), *Parangal kay Cecilio Lopez*, 214–231. Manila: Linguistic Society of the Philippines.
- Shimojo, Mitsuaki. 2011. The left periphery and focus structure in Japanese. In W. Nakamura (ed.), *New perspectives on Role and Reference Grammar*, 266–293. Newcastle: Cambridge Scholars Publishing.
- Van Valin, Robert, Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert, Jr. & Randy LaPolla. 1997. *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Wedgwood, D. 2003. Predication and information structure. Edinburgh: University of Edinburgh (doctoral dissertation).
- Wolff, John U., Maria Theresa C. Centeno, Der-Hwa V. Rau. 1991. *Pilipino through Self-Instruction*. (1. Edition). Ithaca: Cornell University Press.



# WHY EVE SHOULDN'T EAT THE SNAKE: AN INFORMED ANSWER FROM FOCUS STRUCTURE AND REFERENCE TRACKING IN BIBLICAL HEBREW

NICOLAI WINTHER-NIELSEN  
DANSK BIBEL INSTITUT, COPENHAGEN  
VU UNIVERSITY AMSTERDAM

## Abstract

This paper presents a solution for the analysis of the information structure in the ancient language Biblical Hebrew. Using Role and Reference Grammar (RRG), it deals with mapping of lexico-grammatical structures onto tracking of referents in texts. A program developed for reference tracking by Eep Talstra for the database of the Eep Talstra Center for Bible and Computer ran into an interesting problem. In the conversation between Eve and the snake in Gen 2:25-3:5, it miscalculated the coreference as eating the snake rather than the fruit. The solution is enhanced reference tracking with an RRG focus structure analysis and apply common ground management theory formulated by Malte Zinuner mann and others. Furthermore, the hierarchical structure of the text must play a role in an account of the activation of referents.

## Keywords

RRG, focus structure, reference tracking, Biblical Hebrew, ETCBC

## 1. Introduction

When we in oral encounters are confronted with urgent warnings like *Don't Eat It!*, the message easily comes across through tone, pointing or other cues in the interlocution. It helps when the warning is accompanied by a symbol for danger of poisoning, sound alerts or other means of prohibiting dangers of consumption. This is different for written communica-

tion, when we don't have the intended impact unless we can decode it from specific context information that disambiguates the function of the utterance.

The point of departure for this paper is the challenge of decoding the communicative function of written information, illustrated by the first story narrated in the Hebrew Bible, where Eve got herself into a serious communicative challenge. The paper will use her communicative problem as an illustration of the crucial role of focal points and reference in communication.<sup>1</sup> As readers of the story we are not only in Eve's shoes and have to struggle with divine conversation through the mouth of a persuasive snake (Winther-Nielsen 2003), but as linguists working with Biblical Hebrew we also have to interpret the shifting word orders in written texts without access to pitch or pointing. How can a reader know for sure what the precise ordering of the reference phrases are, and how divergent structural patterns mark specific functions for pragmatic impact? Will linguists even have ways to translate the focal and referential structures into modern languages, unless they master the basic mechanisms involved in packaging of information? And if not, are they willing to guess from intuition or insist on tradition?

Lambrecht's (1994) monograph on information structure has set the scene for all serious linguistic work on the interrelationship between pragmatic reasoning and linguistic expressions in languages. Role and Reference Grammar (RRG) was the first theory to include his approach into its handling of pragmatic functions. Even before the publication of Lambrecht's seminal work, it was included by Van Valin (1993), and it took its prominent position in the first standard version of RRG (Van Valin and LaPolla 1997) and the update of the theory (Van Valin 2005). Contemporary RRG is still working on what I from now on will call focus structure (FS), exploring how linguistic structure and function help language users introduce and maintain referents in written communication. This paper will suggest that for Biblical Hebrew we now have resources that can help us put FS on a firmer footing by, when we combine FS with a viable reference tracking (RT) mechanism.

---

<sup>1</sup> This paper has been thoroughly revised following the first presentation at the RRG Conference on August 2, 2015, at the Heinrich Heine University in Düsseldorf. It emerged out of my stay as Researcher in Residence at the ETCBC at the Vrije Universiteit in Amsterdam. I would like to thank Professor Wido van Peursen for this invitation, and Professor Eep Talstra for including me in his research.

RRG claims to be based on communicative adequacy, and this should work also for written communication without interlocutors available to manipulate the felicity of the communicative exchange. Biblical Hebrew is in this regard an ideal testing ground, because not only is it typologically very different from the English evidence usually produced for FS, but the Hebrew data are also accessible for natural language processing through a fully annotated corpus. Because the common ground between interlocutors may be elusive in historical contexts, the ultimate test of FS and RT is how language users can calculate the pragmatic states of referents and decode the mental representation of an entity as active, accessible or inactive as required by (Lambrecht (1994: 49). And, as we shall see, for Eve it turned out that she needs much more than simple parsing rules and in effect she has to rely on a robust theory of both IS and RT in order to get it right.

The paper is structured as follows: In section 2, I refer to new research on corpus-driven RT and explain why the snake is not the right referent in Gen 3:3 and what RRG syntax can offer instead. In section 3, I introduce IS, and I include a similar case from Gen 3:11-12. Section 4 looks at the larger context in Gen 2:25-3:3, and compares RRG's discourse representation with the new corpus-driven RT. In section 5, I add my own proposal for a pseudo-code to deal with focus articulation in BH.

## 2. Corpus-driven RT facing the snake in Gen 3:3

The present proposal to enrich RT through an RRG-based use of FS emerges out of decades of computational linguistic work on Biblical Hebrew in a unique, linguistically annotated corpus. For more than 40 years, the Eep Talstra Center for Bible and Computer (ETCBC) at the VU University in Amsterdam has developed applications for syntactic analysis and produced linguistic data for grammar and discourse studies. This database was introduced to the RRG community and to linguists without training in Biblical Hebrew through the Role-Lexical Module (Winther-Nielsen 2008; 2009; Winther-Nielsen et al 2009), exemplifying parsing, logical structure, templates and focal structures through the database.

Current work on PT by the founder and former director of the ETCBC, Eep Talstra (2016), and elsewhere, shows an impressive potential in an intelligent tutoring system for semi-automated tracking of referents in Hebrew Bible. Talstra is programming lexical and syntactic rules to track the referents through the texts, as illustrated in example (1) which he worked on in March and April of 2015 to produce data on chapter 3 in the Book of Genesis. The first of three programs picks up every instance of a participant referent, e.g. *mimm-ennû* is the 30<sup>th</sup> referent occurring in Genesis 3 (PRef: 30). It performs the first referential registration based on



agreement and cross-clausal syntactic and lexical identification. A second program then searches for an identical referent mentioned earlier within the text in order to gather sets of similar participants, i.e., the 30<sup>th</sup> referent is now classified as part of the 17<sup>th</sup> set, consisting of the 3MSg suffixes (PSet: 17). The goal of a third and final program is then to calculate how referents referred are introduced in reference phrases and tracked through clitic or independent pronouns, e.g., the reference is now pinned down to a final set of actors glossed as "HE" (PAct: 12).

- (1) Participant tracking from line 11 in Gen 3:3:<sup>3</sup>  
 [tō̄ <Ng>] [tō̄-xl-ū {29} <Pr>] [mimm-ennū {30} <Co>]  
 not IMPF-eat-you (2MPI) from-SUFF-3MMsg  
 PRef: 29 [T>KLW <Pr>] PSet: 16= 2pm= PAct: 11= 2pm"Y●UPlmas"  
 PRef: 30 [MMNW :sfx] PSet: 17= 3sm= PAct: 12= 3sm"HE"

Talstra's three reference tracking programs are technically very sophisticated and constantly under development in order to be able to deal with challenges encountered when working with new texts. His program development is not our concern here, but rather how useful such data output are for linguistic research and particularly for the challenge of FS. Unfortunately, there is a snake in this otherwise perfect computational paradise! Automated participant tracking cannot deliver a completely infallible output, but beyond this unrealistic phantom the really interesting question is how far algorithms can take us, and at what point human understanding has to step in, and why. What Talstra has achieved is through gradual trial and error experiments to get closer to the point where a mechanical tracking of participants hits the ceiling, working text by text and improving the programming to explore the limits of purely structural data. Mistakes in the output are often trivial, but sometimes they give crucial feedback on the real limits and challenges of a rule-governed RT and they illustrate what we need to solve through other approaches. These mistakes help linguists to explore conditions for human cognition and what role encyclopedic knowledge should play in reading and interpretation.

Our example (1) from Genesis 3:3 is a rather amusing example of what can go wrong in very complex natural language processing. Talstra's research program in March 2015 tracked the 2<sup>nd</sup> person plural (PRef: 29) and identified them with the addressees Adam and Eve. It succeeded, because

<sup>3</sup> The representation of the text is slightly adapted from [http://bibleol.3bmoodle.dk/text/show\\_text/ETCBC4-translit/Genesis/3/3](http://bibleol.3bmoodle.dk/text/show_text/ETCBC4-translit/Genesis/3/3).

there are no other plural addressees in the preceding text. However, the program had problems calculating the referent for the 3<sup>rd</sup> masculine singular clitic *mimm-ennû* ‘from-him/it (3MSg suffix)’ (PRef: 30). The actor “HE” (PAct: 12) would seem to be coreferential with the only preceding male participant, the snake. Given the drift of the text, a message *Don’t eat from the snake!!!* would obviously miss the point, and no doubt fatally so.

One solution is of course to add the semantic information that the masculine gender also can refer to inanimate referents, so the snake may not be the sole referent available. However, our goal is rather to solve the RT challenge through the FS solution which RRG offers through its theory of the layered structure of the clause and the two positions in front (and after) the core of the clause (Van Valin 2005: 5-6).<sup>4</sup> A Pre-Core Slot (PrCS) is in front of the core of the clause, yet it is still a part of the clause. A Left-Detached Position is external to the clause and functions on the sentence level; however, from now on we will call it the Pre-Detached Position (PrDP), because in a language written from right to left this makes far more sense (Jensen 2016). Applied to our snake-eating problem, example (1) is the core of a clause preceded by ‘But-from-the.fruit.of.the-tree’ (Gen 3:3a) in example (2). This is a PrDP construction even if it is an oblique core argument of the lexical verb ‘eat from/of some fruit’, because in the core of the clause the prepositional phrase (PP) *mimm-ennû* ‘from-him/it (3MSg suffix)’ is a resumptive pronoun that refers back to the fruit rather than the snake. The PP is the coreferential copy (Pavey 2010: 282).

- (2) The preceding Pre-Detached Position (PrDP) in Gen 3:3:  
 $\hat{u}=mip=p^{\prime}ri$                        $h\bar{a}=\zeta\bar{e}\zeta$                        $\rho^{\alpha}\zeta er$                        $b^{\beta}=t\hat{o}x-$   
 ‘but=FR●M=FRUIT[-of] THE-TREE    which    in=midst[-of]

$h\alpha g=g\bar{a}n$   
 the=garden’

On this assumption it should be easy enough to disambiguate the reference of the pronoun because it is explicitly expressed by means of PrDP construction.

<sup>4</sup> For Biblical Hebrew, see Winther-Nielsen (1995: 42-43; 2009: 27-30).

### 3. Introducing FS for Gen 3:3 and 3:12

Even if a simple structural explanation can solve the problem of the eating of the snake in Gen 3:3, this does not explain the functions of the referential structure in many other cases. Distance to previous clauses is far less important than position in the hierarchical structure of the text, and this crucial aspect of pragmatic focusing must be handled through a robust theory of FS, hence we focus on that now.

Lambrecht's (1994: 221-235) canonical version of FS is still the standard taxonomy for focus types in RRG (Van Valin 2005: 69-73). To illustrate the point of FS, we can use the simple piece of evidence presented by Lambrecht (2000: 614) in example (3). Essentially, FS is defined as the formal means to express pragmatic assertion, i.e., to convey new information to an addressee. Lambrecht illustrates three focus categories by three different answers in English to the question *Why didn't Mary come to work today?*

- |     |  |                      |
|-----|--|----------------------|
| (3) | a. <i>She</i> <sub>Topic</sub> <i>had an ACCIDENT</i> <sub>Focus</sub>     | Predicate Focus (PF) |
|     | b. <i>Her HUSBAND</i> <sub>Focus</sub> <i>is to blame</i> <sub>Topic</sub> | Narrow Focus (NF)    |
|     | c. <i>Her HUSBAND IS SICK</i> <sub>Focus</sub>                             | Sentence Focus (SF)  |

In the first example, Mary is treated as predictable information because she is conscious in the mind of the language user and then also identifiable as a referent, and this topic-comment structure is an almost default type of expression with focus on the predicate. In the second example, the opposite holds. Focus is now on the "subject" rather than the predicate, and it is fronted as a constituent in focus. Note that in contrast to Lambrecht, RRG does NOT use Lambrecht's term Argument Focus, because the entity singled out for focus need not be an obligatory argument of the verb, but can also be an adjunct (Van Valin 2005: 71 n. 3). Here the predicate is treated as predictable and presupposed, while the identity of a referent is determined. In the third example, all information is unpredictable and "all new". This three-way taxonomy of presenting all new information, talking about this new topic, and then identifying something new for this topic in some context can be illustrated by the following mini-story: *There's A BEE ON YOUR HAND* (SF). *It's looking HUNGRY ...* (PF). *It's your DONUT that it wants to eat* (NF) (Pavey 2010: 275).

With these three choices for focus structure in mind we can now look for an informal answer to Eve's problem with processing 'eat from it' within its complex linguistic context. There is a tendency in languages like the Mayan VOS language Tzotzil that the PrCS is used for focus articula-

tion and the PrDP has topic function (Van Valin 2005: 7). This is also the case in Japanese according to Shimojo (2011), and it is even stated as a fact by Pavey that “(d)etached positions contain topical material” (2010: 285). According to such textbook evidence, the PrDP *but from the fruit of the tree [which is in the garden] (said God)* (3:3a-c, example (1) and (2)) would be the construction used for topical material. For the clause to have narrow focus, it would have to front the resumptive *mimmennū* prepositional and clitic pronoun: *from IT you shouldn’t eat*, thus using the PrCS slot. On this account, Eve would have been well served had she paid attention to the marking of topical information.

To prove this point, we could look at another example. Another PrDP is used later on in the Adam and Eve story when God confronts Adam and holds him accountable. God uses the diagnostic question word which helps us identify a narrow focus filling the PrCS slot: *mī higgīd ləxā kī fērōm ʔāttā<sup>h</sup>* ‘WH<sub>Focus</sub> told to-you<sub>Topic</sub> that naked [ARE] you<sub>Topic</sub>’ (Gen 3:11). In answer to this question, Adam in example (4) blames the woman as the culprit. In his answer he introduces *hā=ʔiššāh* ‘the woman’ as a topical referent in the PrDP position. In this case the writer does place the resumptive pronoun in the PrCS slot in order to refer to the women in narrow focus through the fronted resumptive independent pronoun *hī?* ‘she’. The rest of the clause is then treated as presupposed and topical.<sup>5</sup>

(4) Completive Narrow Focus in Gen 3:12

<i>hā=ʔiššāh</i>	<i>ʔāšer</i>	<i>nātat-tā<sup>h</sup></i>	<i>šimmād-i</i>
DEF=woman	who	place-PERF/2MSg	with-1Sg
[ [PrDP] <sub>Topic</sub>		(Relative clause ) ]	
<i>hī?</i>	<i>nāt-n-ā<sup>h</sup></i>	<i>=ll-i</i>	<i>min=hā-šēš</i>
SHE	give-3FSg	to-1Sg	from=DEF-tree
( [PrCS] <sub>Focus</sub>		[CORE ] <sub>Topics</sub> ) Clause	

This kind of narrow focus on the NP filling the answer slot of WH-question is very common and hence also integrated as completive focus by Van Valin (2005: 72-73). However, beyond this type and the contrastive focus type, the textbook only casually refers to the FS literature for “further divisions”. This suggests that RRG calls for further work on FS in

<sup>5</sup> For the analysis of the *nātan* in Gen 3:12 and in all of Genesis, see Winther-Nielsen 2017 (365) *et passim*. The text is adapted from [http://bibleol.3bmoodle.dk/text/show\\_text/ETCBC4-translit/Genesis/3/12](http://bibleol.3bmoodle.dk/text/show_text/ETCBC4-translit/Genesis/3/12)

order to benefit from current research in the area which as a vibrant and productive field on its own continues to explore the interplay between mental activation and referential identifiability. A significant work by Erteschik-Shir (2007:67-69, 71) acknowledges how the pairing of the multidimensional projectionist architecture of RRG with Lambrecht's information structure theory has triggered a stimulating development of both approaches as well as "interesting cross-linguistic research" (71). She herself develops the idea of focus structure as a filing system (42-47). A year later, Krifka (2008) published an influential paper on focus structure as a strategy for common ground-management. This explains how the mutually shared knowledge of the interlocutors in a discourse is negotiated by hierarchies of open questions in need to be settled. More recently Matic, Van Gijn, and Van Valin (2014) have argued that linguistic reference calls for an additional separate theory of RT. They combine focus structure and reference tracking as two aspects of the common ground between interlocutors. They use focus to assert how possible worlds are updated in response to controversial, and therefore assertion-worthy, aspects of their common ground. Reference tracking is then a matter of "both interlocutors ascribing same referent value to an expression" (2014: 2).

The most promising new FS solution is proposed by Zimmerman and Ormea (2011). Developing the idea of common ground updating, they propose that opting for an analysis in which a focus structure imposes an ordering on the set of possible worlds, and among those we find the privileged possible worlds (PPW). Those worlds are not defined by content in the sense that they need to be part of the context set at the time of utterance, but they rather "constitute an interpretive background against which the proposition expressed is evaluated." These PPWs provide "a more restricted search space that makes it easier for the hearer of an utterance ... to evaluate, and eventually accept or reject, the actual proposition expressed" (1654). Hartmann and Zimmermann (2009: 1340-1342; cf also Zimmerman and Ormea 2011: 1663) builds on prior work by Rooth on how grammar assigns focus to a constituent based on a set of alternatives, operating with 4 main focus types that all concern how an element is introduced into the common ground through focus (Zimmerman and Ormea 2011: 1663). The selection from a simple set of possible focus expressions can be illustrated for the set of colours {blue, red, green, pink, . . .} as the PPWs, and to each selection I add the short-hand terms  $New_{Foc}$ ,  $Cor_{Foc}$ ,  $Sel_{Foc}$ , and  $Con_{Foc}$  for ease of reference:

- (5) Focus selection from privileged possible worlds (PPWs)
- i. **new-information:** no alternatives explicitly introduced in the preceding discourse  
*Which color did Peter paint his bicycle? He painted it* [New<sub>Foc</sub>]
  - ii. **correctively:** competitors have been explicitly mentioned in the preceding discourse  
*Peter painted his bicycle red. No, he painted it* [Cor<sub>Foc</sub>]
  - iii. **selectively:** choice from a restricted subset explicitly mentioned in the preceding context  
*Did Peter paint his bicycle red or blue? He painted it* [Sel<sub>Foc</sub>]
  - iv. **contrastively:** juxtaposed to one or more elements in the preceding discourse, “where Y, Z, . . . are of the same syntactic category and denote into the same semantic word field as X”  
*Paul painted his bicycle<sub>i</sub>* [Con<sub>Foc</sub>], *and Peter painted it<sub>i</sub>* [Con<sub>Foc</sub>]

In many ways, this new proposal gleans the fruits from the Functional Grammar tradition of Simon Dik, and it offers a transparent calculation of focus-types that can help implementation in an RRG focus structure analysis. This work is also typologically relevant as demonstrated by the analyses of information structure for African languages by Güldemann, Zerbian, and Zimmermann (2015). Even more to the point, the PPWs to a large extent explain how to use the notion of potential focus domain in RRG focus structure theory. Finally, the PPWs can give more structure to the search for an independent RT solution called for by Matić, Van Gijn, and Van Valin (2014). Last, but not least, it might help us use the unique data on Biblical Hebrew RT coming out of the work of Talstra and the team of the ETCBC, and this will be our next step to take.

#### 4. Reformulating RT for an RRG analysis of Gen 2:25-3:3

In order to use the unique data for Biblical Hebrew available through the corpus-based RT research of Talstra, we will again turn to the snake problem that Eve faced in Gen 3:3 and now include the larger prior context for the utterance, which is Gen 2:25-3:3. The goal is first and foremost to put the canonical RRG approach to the test and then explore to what extent the new reference tracking research can help us develop a more robust solution for Biblical Hebrew.

Because Van Valin (1993) at a very early stage adapted FS for the RRG theory, it was possible for Winther-Nielsen (1995) to use the ETCBC database and test the mechanisms of RRG on the Book of Joshua from the Hebrew Bible. In this work, the discourse topicality terminology of Dik

was used as a useful supplement for an RRG analysis (Winther-Nielsen 1995: 66-67). Unfortunately, this research did not get much attention within the guild of Biblical Hebrew linguistics. The dissertation of Den Braber (2010) clearly demonstrates that scholars of the Hebrew Bible by and large look for historical and diachronic explanations of the text on the one hand, and for contemporary relevance on the other hand, so the market for FS really wasn't there. Later on Hebrew Bible linguistics moved towards cognitive linguistic solutions and this was applied for corpus-based research (Talstra and van der Merwe 2002; 2004). However, information structure research has continued to make an important impact on the research in a number of dissertations, such as Heimerdinger (1999), Shimasaki (2002), Floor (2004), Lunn (2006), Moshavi (2010), Kummerow (2011), Pang (2011), and Westbury (2014).

The *status questionis* in Biblical Hebrew linguistics is by now captured in the reference work of van der Merwe et al (2017) which will no doubt become an influential resource to guide students and teachers of Biblical Hebrew information structure. It is therefore useful to use this work as a point of reference on previous FS work on Biblical Hebrew. To begin with the PrDP, the account of its topic function is relatively clear in the account of van der Merwe et al (2017: 515-516). It is used for topic-announcing, so that a referent is made available to play a role in the following discourse. The functions are (re)activation of weakly accessible referents into primary or secondary topics, but also for comparison and contrast, and to activate as topic for immediate focus.

The overview of the function of PrCS in this reference grammar is more challenging because it includes many examples that relies on interpretation of specific contexts (497-509). This of course will go far beyond the goal set for this paper. However, it may be possible to provisionally to tie their interpretative distinctions into the Zimmerman-Otega approach.

As for the PrCS, we would by default assume that this position is used for marked focus, but secondarily we also need to consider "topic (re)activation" as a possible alternative. In other words, a topical function for this slot must be argued on other grounds that could be specified in clear rules for criteria to select from.

Most of the cases in table (1) are contrastive or new foci. The constituent focus interpretation is the narrow focus of RRG.

Table 1 Topic (re-)activation (based on van der Merwe et al (2017: §47.2.1 (1)))

			PPW
(a) Select from set	<i>a</i> Comparison	<i>Abram lived in Cana 'an: but Lot ...</i>	Con <sub>Foc</sub>
	<i>b</i> Contrast	<i>Elkhana: went ... but Hanna didn't</i> (1 Sam 2:21-22)	Con <sub>Foc</sub>
	<i>d</i> Simultaneity	<i>For seven days ... but on the eight</i> (Ex 22:29)	Con <sub>Foc</sub>
(b) Topic-shift	<i>a</i> New entities	<i>And the long dress...</i> (2 Sam 13:19)	New <sub>Foc</sub>
	<i>b</i> New persons	<i>And to Sara he said:...</i> (Gen 13:12)	New <sub>Foc</sub>
(c) Summary	End of story	<i>Moses ... and the Israelites stroke them</i> (Jos 12:6)	New <sub>Foc</sub>

The cases in table (2) are either contrastive or new foci. Depending on the context they would be either have selective or corrective functions.

Table 2 Constituent focus (based on van der Merwe et al (2017: §47.2.1 (2)))

			PPW
(a) Identity	<i>a</i> identify (person, entity, manner)	<i>Who should ... Juda should</i> (Judg 1:1-2)	Self <sub>Foc</sub>
(b) Modify	<u>LIMIT</u> to <i>c</i> person	<i>GOD ... revealed himself...</i> (Ex 5:3)	Cor <sub>Foc</sub> ?
	<i>e</i> purpose	Purpose clause (Gen 5:7)	
	<i>f</i> quality	<i>In the cleanness of my heart.</i> (Gen 20:5)	
	<i>g</i> quantity	<i>ALL THE PEOPLE ... all soldiers died</i> (Jos 5:4)	
	<u>FUNCTION</u> to <i>i</i> specify	<i>SONS OF JACOB ... plundered ... their SHEEP, CATTLE ...</i> (Gen 34:27-28)	
	<i>j</i> supplant	<i>The planted X ... Y they reaped</i> (Jer 12:13)	



	<i>k</i> expand	<i>All men ... they circumcised</i> (Gen 17:27)	
	<u>Focus Particle</u> <i>l</i> expand	<i>Also THE WILD ANIMALS I gave</i> (Jer 27:6)	
	<i>n</i> replace	<i>No, but rather</i> ● <i>NY...</i> (Gen 19:2)	
(c) Deictic Referent	<i>o</i> cataphoric	<i>BY THIS ...</i> (Gen 19:2)	Self <sub>Foc</sub>
	<i>p</i> anaphoric	<i>THESE EIGHT ...</i> (Gen 22:21-23)	Self <sub>Foc</sub>
(d) ●aths	<i>a</i> (probably)	<i>I swear</i> (Gen 21:23-24)	?

The examples given by van der Merwe et al (2017) on sentence focus appear to indicate that sentence focus pertains to presentative clauses or background information, implying that sentence focus will introduce segments in the beginning of larger texts or smaller embedded texts. Most of the cases in table (3) are new foci.

<u>Introduce:</u> (a) Report	Direct speech	<i>A man came to us: "</i> ... (2 Kings 1:6)	New <sub>Foc</sub>
(b) Episode	Narrative	<i>Benhadad mobilized his army..</i> (1 Kings 20:1)	New <sub>Foc</sub>
(c) Back-ground/ Flashback	<i>a</i> New episode	<i>Saraj... bore him no children... She had an Egyptian servant</i> (Gen 16:1-2)	New <sub>Foc</sub>
	<i>c</i> Explanation	<i>She had brought them on the roof ...</i> (Josh 2:6)	Self <sub>Foc</sub> ?
(d) Temporal	Event	<i>While he was still talking, Rachel came ...</i> (Gen 29:9)	Con <sub>Foc</sub>

Beyond this, the reference grammar also assumes that fronting in Biblical Hebrew may have specialized functions. Among those are what Erteschik-Shir (2017: 16-17) calls stage topics, stipulating that every sentence focus must have at least one topic. These cases in table (4) would probably work as new foci. In this group there are also poetic uses of fronting.

Table 4 Other fronting (based on van der Merwe et al (2017: §47.2.1 (4)-(6)))

(4)	Grounding in time/place	<i>AT THAT TIME TIME ABIJA ... GOT SICK</i> (2 Kings 1:6)	NewFoc
(5)	Simultaneity	<i>WHEN THEY CAME to Zuf, Saul said ...</i> (1 Sam 9:5)	NewFoc
(6)	Chiasm in poetry		?

Before I implement an FS analysis based on the of PPWs along the lines suggested by Zimmermann and others, I will first describe how canonical RRG handles RT and then elaborate on the corpus-driven RT developed by Talstra for the ETCBC. For more than a decade, the canonical version of RRG has treated context through the discourse representation theory (DRT) of Kemp and Reyle (1993). Following DRT, Van Valin (2005: 174) introduces a new set of logical terms *s*, *t*, *u*, *v* to formalize discourse structure (Figure 5.32). For derivation of focus structure, he uses the more traditional logical structure terms *x*, *y*, *z* and an informal specification of semantic representation (Figure 5.32). Furthermore, Van Valin exemplifies the derivation of completive and contrastive narrow focus (Figure 5.34). This results in the inclusion of two boxes for presupposition and assertion in the already extremely complex three-dimensional projections of linking (2005: 173 Figure 5.35).

I illustrate RRG's discourse representation in a crudely simplified way in figure 1. From a reading of the early chapters of the Book of Genesis in a translation, the reader can gain a first impression of what is at stake. First, we need to include situational knowledge on the instruction given by God to Adam in Gen 2:16-17. The story presupposes a permission to eat from all trees except for a prohibition to eat from the tree of knowledge of good and evil in the midst of the garden. This requires a contextual distinction between ( $y_1$ ) and ( $y_2$ ) and selection between the two. Second, there is a conversational turn by Eve to correct the snake. This is the background for a marked contrastive PrDP to select between all trees and the one central tree which is singled out as topical.

Shimojo (2011) to my knowledge was one of the first to use DRT for his clarification of the missing verb in the numeral quantifier construction in Japanese, but it is noteworthy that Shimojo's (2011: 288-291) later work does not use DRT for linking. This testifies to the fact that it may very well be too cumbersome to specify all linguistic data to this depth,

and I assume this explains why DRT has not figured prominently in RRG research over the last decade.

Snake (Gen 3:1-3):  Adam and Eve can eat from all fruit of the trees	Presupposition  x, y Adam, Eve (x) Fruit of all tree (y <sub>1</sub> ) X(y <sub>2</sub> ) x eat y	Assertion  x, y Adam, Eve (x) Fruit of tree (y <sub>1</sub> ) Fruit of tree in midst of garden (y <sub>2</sub> ) x eat y	⇒ Eve (Gen 3:3)  ... but from the fruit of the tree ... do not eat from it Contrastive Topic
God (Gen 3:11-12):  Who gave Adam the idea to eat from the fruit of the trees?	Presupposition  x, y, z X(x) Fruit of tree (y) Adam (z) x give y to z	Assertion  x, y, z X(x) From fruit of tree (y) Adam (z) x give y to z	⇒ Adam (Gen 3:12)  The women .... she gave me from the fruit of the tree Completive NF

Figure 1. Presupposition and Assertion in DRT format

Now let me introduce the challenge of RT for the snake problem from the perspective of the semi-automated corpus-driven analysis of RT in the output from the program developed by Talstra. I have inserted the output of the analysis of participant actors (cf PAct=11 and PAct=12 in example (1)) for all of Gen 2:25-2:3 in the transliteration of column 2 of table 5 as raised numbers. I can now explain the snake problem as follows: The program does not identify the reference of *mimmennū*<sup>[12]</sup> 'it' with *mippəri ʕēs-haggān*<sup>[3]</sup>, 'from-fruit.of tree.of—the-garden (3:2b) because in that case Eve could not eat any fruit at all, and she would starve to death. The program also avoids to identify the antecedent with the snake's casual reference to *mikkōl ʕēs haggān*<sup>[6]</sup> '(eat) from all trees of the garden' (3:1e).

Table 5. Participant Reference and RRG analysis of Gen 2:25-3:3

	Text with Participant Actors	TT	Cl Type	Focus
<b>Gen 2</b>	<i>wayyih<sup>2</sup>yū</i> <i>š<sup>2</sup>rēhem</i> <i>š<sup>2</sup>rūmmim</i> and.they.were      the.two <sub>ACS-T</sub> naked,	N	WayX	PF>SF NEW <sub>Foc</sub>
<b>25a</b>	<i>hāʔādām</i> <i>w<sup>2</sup>ʔišō</i> the.MAN      and his.WOMAN <sub>ACS-T</sub>			
<b>25b</b>	<i>w<sup>2</sup>lōʔ</i> <i>yitbōšāšū:</i> and-not they.were.ashamed	ND	WxY●	PF
<b>Gen 3</b>	<i>w<sup>2</sup>hannāhās<sup>[1]</sup></i> <i>hāyā<sup>h</sup></i> <i>š<sup>2</sup>arūm</i> but-the-SNAKE <sub>BNA</sub> it.was cunning	N	WXQt	NF CON <sub>Foc</sub>
<b>1a</b>	<i>mikkōl</i> <i>hāyyat</i> <i>hāsšāde<sup>[28]</sup></i> more.than.all      animals-of      the.field <sub>ACS-T</sub>			
<b>1b</b>	<i>ʔāšer</i> <i>šāsā<sup>[2]</sup></i> <i>YHWH</i> <i>ʔelōhīm<sup>[2]</sup></i> which he.made      Yahweh God <sub>ACS-T</sub>	N	xQtX	PF
<b>1c</b>	<i>wāy<sup>2</sup>ōmer<sup>[1]</sup></i> <i>ʔel-hāʔiššā<sup>[3]</sup></i> and-it.said      to-the.woman <sub>ACC-T</sub> ●F	N	Way●	PF
<b>1d</b>	<i>ʔaf</i> <i>kī-</i> <i>ʔāmar<sup>[4]</sup></i> <i>ʔelōhīm<sup>[4]</sup></i> <sub>ACC-T</sub> ●F True      that he.said      God	NQ	xQtX	PF
<b>1e</b>	<i>lōʔ</i> <i>tōʔxlū<sup>[5]</sup></i> <i>mikkōl</i> not you.eat      FROM-ALL.OF	NQ	xYq●	SF NEW <sub>Foc</sub>
	<i>šēs</i> <i>haggān<sup>[6]</sup></i> TREE.OF      THE-GARDEN <sub>BNA</sub>			
<b>2a</b>	<i>wattōʔmer<sup>[3]</sup></i> <i>hāʔiššā<sup>[3]</sup></i> And-she.said      the.woman	N	WayX	PF
	<i>ʔel-hannāhās<sup>[1]</sup></i> to-the-snake <sub>ACC-T</sub> ●F			
<b>2b</b>	<i>mipp<sup>2</sup>ri</i> <i>šēs-</i> from-FRUIT.of      TREE.of	NQ	xYq●	NF SEL <sub>Foc</sub>
	<i>haggān<sup>[8]</sup></i> <i>nōʔxel<sup>[3]</sup>:</i> the-GARDEN <sub>BNA</sub> we.eat			
<b>3a</b>	<i>ūmipp<sup>2</sup>ri</i> <i>hāšēs<sup>[9]</sup></i> but-from-fruit.of      the-tree <sub>BNA</sub>	NQ	CPen	PrPD CON <sub>Top</sub>
<b>3b</b>	<i>ʔāšer</i> <i>b<sup>2</sup>ōx-</i> <i>haggān<sup>[7]</sup></i> which in-midst.of-      the-garden	NQ	NmCl	PF
<b>3c</b>	<i>ʔāmar<sup>[4]</sup></i> <i>ʔelōhīm<sup>[4]</sup></i> he.said      God <sub>ACC-T</sub> ●F	N	ZQtX	PF
<b>3d</b>	<i>lōʔ</i> <i>tōʔxlū<sup>[11]</sup></i> <i>mimmennū<sup>[12]</sup></i> not      you.eat      from-it <sub>ACV</sub>	NQ	xYq●	PF
<b>3e</b>	<i>w<sup>2</sup>lōʔ</i> <i>tigg<sup>2</sup>ʔū<sup>[11]</sup></i> <i>bō<sup>[12]</sup></i> and-not      you.touch on-it <sub>ACV</sub>	NQ	WxY●	PF
<b>3f</b>	<i>pen-</i> <i>t<sup>2</sup>mutūn<sup>[11]</sup>:</i> in.order.not-      you-die	NQ	xYq●	PF

This corpus-driven RT gradually evolves from sophisticated research. A semi-automated tracking of referents disambiguates RT in the majority of clauses, and it is therefore a strong solution for a corpus linguistic disambiguation of referents. The RT output could no doubt assist a researcher who is trying to do an analysis in the DRT-format in figure 1 and it could determine presupposition and assertion for “Adam, Eve (x)”. However, our interest is rather to be able to determine the FS type and how we can select from alternative PPWs, and this calls for a final step on the analysis presented in the remaining part of the table 5.<sup>6</sup>

## 5. Improving activation analysis for corpus-driven IS

The final step in our demonstration of a new solution is to illustrate the most likely assignment of focus in the natural language context we have worked with from Gen 2:25-3:3.

In RRG, grammaticalization of FS is included in its projection of the clause as a cognitive model of context (2005: 182 Figure 5.40). This model is calculated through “the activation level of the referents” (2005: 79). Van Valin reduces the activation categories to five: (1) active in direct mention, (2) accessible as recognizable from world knowledge or environment, (3) Inactive, but mentioned earlier, (4) brand new anchored to accessible or earlier mentioned, (5) brand new unanchored, never previously mentioned nor accessible. Lambrecht’s (1994: 109) original version had more detail and we will therefore use the model proposed by Van Valin and La Polla (1997: 201) and by Pavay (2008: 308), here illustrated in figure 2.

---

\* The text in the display has been copied from

[http://bibleol.3bmoodle.dk/text/show\\_text/ETCBC4-translit/Genesis/2/25](http://bibleol.3bmoodle.dk/text/show_text/ETCBC4-translit/Genesis/2/25) &

[http://bibleol.3bmoodle.dk/text/show\\_text/ETCBC4-translit/Genesis/3/1/3](http://bibleol.3bmoodle.dk/text/show_text/ETCBC4-translit/Genesis/3/1/3)

The special textual and grammatical sigla used in the third and fourth column are:

TT=Text type; Codes: N=narrative text. ND=narrative discursive (a non-narrative direct address to the reader), NQ=quotation in narrative.

Cl type=clause type. WayX=*wayyitol* verb conjugation before subject, Way0=*wayyitol* verb without Subject; ZQtX=clause-initial *qatal* conjugation followed by subject, xQtX=a non-subject element before a *qatal* followed by subject, WXQt=a coordination before a fronted subject followed by *qatal* xY#0=a non-subject element before a *yiqtol* conjugation without a subject NmCl=verbless clause (or nominal clause)

CPen=casus pendens (the PrPD).

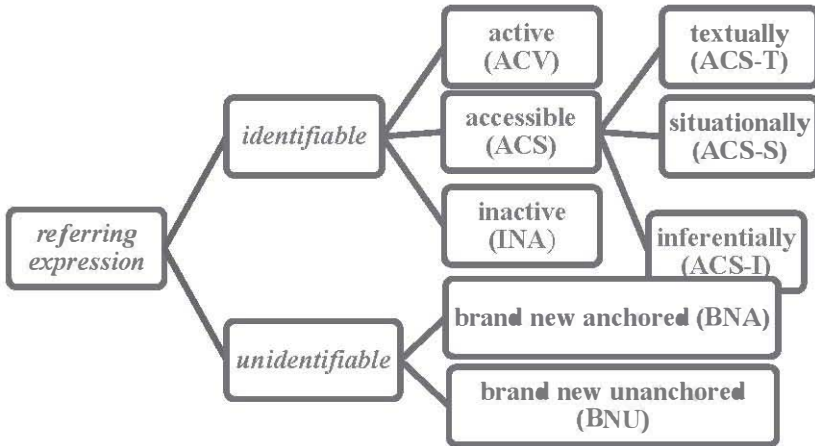


Figure 2. Identifiability and activation of referring expressions

Activation analysis is crucial, because Lambrecht claimed that active and accessible referents are topical (Lambrecht 1994: 102). Others will align referential expressions on a continuum from topical active to focal brand new. However, activation level cannot stand alone, because pragmatic functions also involve “at which discourse level they occur (textual hierarchy), and how prominent they are in subsequent discourse (relevance)” (Winther-Nielsen 1995: 63). The RT analyses based on the ETCBC corpus also suggest that we need to work with activation spans within hierarchically organized segments of discourse. It is of course not the case that preceding common ground is simply the preceding clause, but rather how clauses are positioned within the network of a particular preceding segment.

Furthermore, the research literature indicates that nominal full referential phrases can be licensed by cases of disagreement, negotiation of the identity of a referent or for expression of overt contempt (1995: 68). This calls for a much richer analysis in the sense that known and identifiable entities can very well be expressed as focus when there is a negotiation of alternatives among sets in the common ground, as set out for the four types in the Hartmann-Zimmermann taxonomy. The analysis of the Joshua corpus by Winther-Nielsen (1995) proved that demarcation of discourse segments was another primary factor involved in the encoding of full referential phrases.

The strength of the database is then to allow the researcher to do an analysis within a segment defined by the hierarchical structure of the preceding contexts. Within this context it is possible to explore whether a referential phrase is expressed because it is accessible textually from prior mention in the narrative. The alternative is that it can be mentioned in order to introduce a quote as part of the mechanism to keep track of who says what in monologues and in conversational exchanges in the narratives. The latter explains most of the cases with explicit speaker and hearer expressions in PF.

This approach is used for the analysis of FS displayed in table 1. The analysis is made manually, based on observation and interpretation of focus structure, and then mapped in column 5. When appropriate I specify which focus type is used.

The ETCBC database provides some very useful help for the linguist to decide on the function of word order variation: Does a clause start with a conjunction (W)? Is there a subject (X) or a non-subject (x) *before* the verb? Is there a subject (X) or no subject (Ø) *after* the verb? Which form of the verb conjugation is used? BH has three finite verbal forms, infinites and verbless clauses. The use of these four-letter morpho-syntactic labels are explained in example (6).

- (6) Morpho-syntactic labels in the ETCBC database
- *Wayyiqtol*: a clause-initial narrative chaining form; WayX with PSA (“subject”) vs. WayØ without
  - *Yiqtol*: Non-perfective (and non-past in most cases): Label has one or more of the Yqt letters
  - *Qatal*: Multivalent, Perfective in Narrative – Imperfective in Quotation: Label from Qtl letters
  - Other sigla refer to imperative clauses (ImpØ), verbless clauses (NmCl for “nominal clause”), etc.

Information on text type (TT) in column 4 enables the linguist to track referential phrases (RP) within clearly delimited segments and to work with a linguistic demarcation of subsegments.

I will therefore reformulate the activation analysis exemplified for the segment in Gen 2:25-3:3 into a pseudo-code in table 6. This stepwise procedure for analysis is a first preliminary version of an algorithm that needs to be refined through analysis of much broader stretches of text from the Hebrew Bible, and its role at this stage is simply to illustrate the mechanisms involved in an analysis.

Table 6. Pseudo-code for Activation Analysis of Gen 2:25-3:3

For every new clause fragment:

If N● RP, but head-marked person-gender-number morphology coreference: assign active (AVC) and assume PF (2:25b, 3:3d.e.f)

If RP is identifiable

If RP REACTIVATED in a position after the verb:

If the verb is a communicative predicate, then assume that reference to any textually accessible speaker or hearer is expressed as part of communicative dynamics in the story in order to introduce or highlight interlocutors (3:1d, 2a, 3c): assign PF

If the RP is in a subordinated clause, then the RP serves disambiguation of reference (3:1b, 3b): assign PF

Else, if this RP can be new focus in the beginning of a quote, assign PF

Else, if the RP is topical, assume it is active (ACV)

If RP is unidentifiable

If X is a brand new entity, but anchored in the story, and the RP is placed in the PrCS: assign NF and explore whether contrastive, selective or corrective focus.

Else, RP is new information: assign SF

We can now spell out this analysis in plain linguistic language in order to explain the context for our snake problem. The fronted PrDP in Gen 3:3a discussed in example (2) is part of a contrastive pair initiated by the preceding PrDP *from-fruit.of tree.of-the-garden we-eat* (3:2b). This pair part is an initial affirmation by the woman of God's permission to eat from the trees in general. It is introduced as a brand new anchored entity in focus within the domain of the conversation. The speaker Eve must assume this information to be totally new to the snake. Both fronted entities have had no PAct status before now, but Eve is now asserting these referents for the first time in the conversation.

At the beginning we find another contrastive focus expressed with literary wordplay for the couple which are *ʕarimmim* 'naked' and without any shame caused by guilt (2:25) and the snake which is *ʕarim* 'sly'. The referent *šimê-hem* 'two-theirs' is quite unusual, and the possessive suffix anchors this referent to the previous episode. The clause type label WayX indicates that the RP is placed in its default position after the verb and therefore is unmarked. The first mention of the nakedness introduces the two as topical and then refers to them in PF focus using a post-core slot



(PoCS). This is then followed by a reference to the snake placed in the PrCS for contrastive focus (3:1). The snake is a brand new entity, but anchored in the reference to all animals in general.

Space will not permit us to go into further detail for Biblical Hebrew at this point. The snake problem and its immediate context just illustrate the kind of work that needs to be done in order to apply focus structure analysis for Biblical Hebrew. It should serve as a basis for projects with a much broader scope than this limited case. Talstra continues to develop his RT analyses for many other chapters from the Hebrew Bible and such output from semi-automated analysis could be used for developing more refined pseudo-code for activation analysis. The case at hand not only provides a basis for more solid research into IS, but it also illustrates a new direction for and a much firmer ground for doing corpus-driven FS studies.

## 6. Conclusion

Despite the problem with the snake and many other linguistic challenges encountered in a pragmatic analysis of RT and IS of Gen 2:25-3:3 and 3:12, it seems that paradise is certainly not lost, and that corpus-driven RT and IS can bring us much closer to an informed understanding of the function of variation in referential phrases.

We have seen how RT can be handled in the layered clause structure of Gen 3:3 (section 1), and how focus structure can be explained for Gen 3:11-12 (section 2). We then presented the RT programs developed by Talstra as an important new solution to replace the representation of discourse in the current version of RRG (section 3) and in the final part I presented the outline of a new pseudo-code for analysis of IS in BH, and tested it on Gen 2:25-3:3 (section 4).

The analysis presented has a limited scope. The next step is to bring much more evidence into play. Eventually we can hope that it will be possible for programmers to develop a user-friendly interface for semi-automated corpus-driven analysis of FS based on the new RT data coming out of research on the ETCBC corpus of the Hebrew Bible. It would track all referents and provide outputs of linguistic expression within their domains. Similar corpora are currently being built for other ancient texts on the Bible in other languages, but linguistically annotated corpora could also be built for modern languages.

## References

- Braber, Marieke E. J. den. 2010. *Built from many stones: An analysis of N. Winther-Nielsen and A. G. Auld on Joshua with focus on Joshua 5:1–6:26*. Amsterdam: 2VM.
- Erteschik-Shir, Nomi. 2007. *Information structure: The syntax–discourse interface*. Oxford: University Press.
- Floor, Sebastian Jonathan. 2004. *From information structure, topic and focus, to theme in Biblical Hebrew narrative*. Stellenbosch: University of Stellenbosch. (Doctoral dissertation).
- Jensen, Christian Højgaard. 2016. Participant-reference shifts in Zech. 1:1–6. An assessment of diachronic and synchronic approaches. *HIPHIL Novum* 3(1): 25–46. <https://hiphil.org/index.php/hiphil/article/view/8>
- Güldemann, Tom, Zerbian, Sabine & Zimmermann, Malte. 2015. Variation in information structure with special reference to Africa. *Annual Linguistic Review* 1:155–178.
- Hartmann, Katharina & Zimmermann, Malte. 2009. Morphological focus marking in Gùrùntùm (West Chadic). *Lingua* 119 (9): 1340–1365.
- Heimerdinger, Jan M. 1999. *Topic, focus and foreground in Ancient Hebrew narratives*. Library of Hebrew Bible Old Testament Studies 295. Sheffield: Sheffield Academic.
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55: 243–276.
- Kummerow, David. 2011. *Functional–typological and constructional linguistic studies on Tiberian Hebrew: Anaphora, deixis, and the verbal system*. Sydney: Australian College of Theology. (Doctoral dissertation).
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge Studies in Linguistics 71. Cambridge: Cambridge University Press.
- . 2000. When subjects behave like objects. An analysis of sentence-focus constructions across languages. *Studies in Language* 24(3): 611–682.
- Lum, Nicholas P. 2006. *Word-order variation in Biblical poetry*. Paternoster Biblical Monographs. Milton Keynes: Paternoster.
- Matić, Dejan, Van Gijn, Rik & Van Valin Jr., R. D. 2014. Information structure and reference tracking in complex sentences: An overview. In van Gijn, Rik, Hammond, Jeremy, Matić, Dejan, van Putten, Saskia & Galucio, Ana Vilacy (eds.), *Information structure and reference tracking in complex sentences*, 1–42. Amsterdam: Benjamins.

- Moshavi, Adina. 2010. *Word order in the Biblical Hebrew finite clause*, Linguistic Studies in Ancient West Semitic. Winona Lake, IN: Eisenbrauns.
- Pang, Philip K. 2012. *Fulfillment hermeneutics in the Books of Kings*. Dallas: Dallas Theological Seminary. (Doctoral dissertation).
- Pavey, Emma L. 2010. *The structure of language: An introduction to grammatical analysis*. Cambridge: Cambridge University Press.
- Shimasaki, Katsuomi. 2002. *Focus structure in Biblical Hebrew: A study of word order and information structure*. Bethesda, MD.: CDL Press.
- . 2013. Information structure: Biblical Hebrew. In Khan, Geoffrey (ed.), *Encyclopedia of Hebrew Language and Linguistics*, 279–83. Leiden: Brill.
- Shimojo, Mitsuaki. 2011. The left periphery and focus structure in Japanese. In Wataru Nakamura (ed.), *New perspectives in Role and Reference Grammar*. Cambridge Scholars Publishing, 266-293.
- Talstra, Eep. 2016. Approaching the mountain of Exodus 19: thou shalt explore syntax first. *HIPHIL Novum* vol 3(1): 1-24  
<https://hiphil.org/index.php/hiphil/article/view/9>.
- . 2017. The Hebrew Bible and the computer. The poet and the engineer in dialogue. *Journal of Humanities & Arts Computing: A Journal of Digital Humanities* 1(1): 49–59.
- Talstra, Eep, & van der Merwe, Christo H. J. 2002. Analysis, retrieval and the demand for more data. Integrating the results of a formal textlinguistic and cognitive based pragmatic approach to the analysis of Deut. 4:1-40. In Cook, Johan (ed.), *Bible and computer. Proceedings of Association Internationale Bible et informatique, "From alpha to byte"*, University of Stellenbosch, 17-21 July, 43–78. Leiden: Brill.
- . 2004. Biblical Hebrew word order. The interface of information structure and formal features. *Zeitschrift für Althebraistik* 15/16: 68-107.
- Van der Merwe, Christo. 2013. Fronting: Biblical Hebrew. In Khan, Geoffrey (ed.), *Encyclopedia of Hebrew Language and Linguistics*, 931-935. Leiden: Brill.
- Van der Merwe, Christo H. J. & Naudé, Jacobus A. & Kroeze, Jan H. 2017. *Biblical Hebrew reference grammar: Second edition*. London: Bloomsbury T & T Clark.
- Van Valin Jr., Robert D., 1993. A synopsis of Role and Reference Grammar. In Van Valin, Robert Jr. (ed.), *Advances om Role and Reference Grammar*, 1-164. Amsterdam and Philadelphia: John Benjamins.
- . 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

- Van Valin Jr., Robert D. & LaPolla, Randy J. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
- Westbury, Joshua R. 2014. *Left dislocation in Biblical Hebrew: A Cognitive Linguistic account*. Stellenbosch: Stellenbosch University. (Doctoral dissertation).
- Winther-Nielsen, Nicolai. 1995. *A functional discourse grammar of Joshua: A Computer-assisted rhetorical analysis*. Stockholm: Alquist & Wiksell.
- . 2003. Reading for the creator's discourse: Speech acts and transcendence in Genesis 1-3. In Ohrstrom, Peter (ed.), *Time, Reality, and Transcendence in Rational Perspective*, 39-73. Aalborg: Aalborg University Press
- . 2008. A role-lexical module (RLM) for Biblical Hebrew: A mapping tool for RRG and WordNet. In Robert Van Valin (ed.), *Investigations of the Syntax-Semantics-Pragmatics Interface*, 455-78. Amsterdam and Philadelphia: John Benjamins.
- . 2009. Biblical Hebrew parsing on display: The role-lexical module (RLM) as a tool for role and reference grammar. *SEE-J Hiphil* 6:1-51 [<http://www.see-j.net/index.php/hiphil/article/view/78>].
- . 2017. Corpus-driven valence: Give and the meaning of נתן nātan in Genesis. In Moshavi, Adina & Notarius, Tania, *Advances in Biblical Hebrew linguistics: Data, methods, and analyses*, 363-385. Eisenbrauns: Winona Lake.
- Zimmermann, Malte & Onea, Edgar. 2011. Focus marking and focus interpretation. *Lingua* 21.1: 1651-1670.



# LANGUAGE INDEX

- Afrikaans 177, 184-187  
Aghem, 215, 218-220  
Amele 243, 248, 249, 254  
Babungo, 219, 236  
Bafut, 219  
Bamunka, 213, 215-237  
Barai 242, 254  
Basque 177, 185-187  
Belhare 252  
Bengali, 35  
Bulgarian 177, 180  
Celtic, 177  
Chechen 253  
Chinese, Mandarin 72, 242, 248, 254  
Dutch 170, 173, 175, 177-182, 184, 186, 187  
Dyirbal, 193  
English, 65, 68, 69, 71, 136, 137, 143, 144, 159, 162, 163, 191, 193, 194, 196, 206, 207, 218, 222, 229, 267, 270, 272, 279,  
French, 36, 177, 181-184, 186, 187,  
Georgian, 14  
German 177-182, 184, 186, 187,  
Germanic 177-180  
Hebrew, Biblical 285-289, 293-296, 304-307  
Hindi, 35  
Icelandic 177, 179-182, 184, 186, 187  
Irish, 36, 177, 185-187,  
Italian 177, 183, 184, 186, 187,  
Japanese, 15, 72, 89-107, 248, 253, 291, 297, 306  
Kikuyu 253, 254  
Korafe 245, 248, 249, 253  
Kwaza 253, 254  
Lithuanian 180  
Mopan, 72  
Mparntwe Arrernte 11  
Ngaanyatjarra, 126  
Noni, 215  
Nootka 248, 253  
Norwegian 177, 179-182, 184, 186, 187  
●bang, 219  
Persian, 32-57  
Pitjantjatjara, 113, 116-118, 123, 124, 126, 127, 129, 130-133  
Portuguese 177, 181, 182, 184-187  
Q'eqchi' 253  
Romance, 177, 181, 183  
Romanian 177, 183, 184, 186, 187  
Russian 136, 137, 142, 145-147, 154, 158, 159, 162-164, 169, 170, 173, 175-180  
Serbian, 180  
Slavic 177, 180-182, 184, 186, 187  
Spanish, 36, 177, 183, 184, 186, 187, 191-193, 196, 198, 205-207  
Swahili, 72  
Tagalog, 257-262, 266-268, 276, 279, 281-283  
Taiwan Sign Language, 63-80, 84  
Tauya 245, 250, 251, 254  
Turkish, 72, 248, 254  
Tzotzil, 290  
Urdu, 35  
Wambule 245, 254  
Yankunytjatjara, 113, 130, 132, 133  
Yaqui 253



# SUBJECT INDEX

- acti**o**nality 174, 175, 178, 182, 183, 186
- acti**v**ation, 285, 292, 294, 295, 300-304
- act**o**r 98-102, 139, 140, 143, 144, 146, 150, 158, 160, 171, 172, 176, 180, 190-193, 196, 198, 200-203, 205-208,
- act**o**r-und**e**rgoer hierarchy 79, 100, 102, 143, 196
- act**u**al focus do**m**ain 242, 246, 250-252, 266, 276
- ad**j**ective 213-215, 218, 220-230, 235-237
- di**m**ensional 20-22
- ad**j**unct 95, 136, 140, 142, 145, 154
- ad**p**osition 142, 143, 145, 163, 193
- non-predicative 142
- predicative 142
- see also postpositions, prepositions
- ad**p**osition assignment rules 193, 194, 204, 206
- English 143, 162, 194-196
- Spanish 204-206
- ad**v**erb 145, 157
- ag**r**ee**m**ent 217, 221, 223, 225, 227, 232, 234, 235
- Aktionsart 3-27, 54, 177
- accomplishment 4-27
- achievement 4, 7-8, 23
- act**i**ve accomplishment 9-17, 22-24, 27
- act**i**vity 4-5, 7-10, 12, 17-18, 23-25
- causative accomplishment 9-10, 14-17
- agency 5, 7-9, 14
- animacy 173, 177, 217, 232, 234
- appositi**v**es 215, 235
- arg**u**ment, core 39, 43, 44, 54, 92, 96, 99, 102, 106-109, 135, 142-148, 157, 162, 163, 222,
- di**r**ect 114, 117, 127, 146-148, 193
- obli**q**ue 114, 117, 122-125, 129, 130, 135-137, 191, 192, 199, 200, 203, 204
- arg**u**ment-ad**j**unct 142, 143, 154
- as**p**ect, nominal 214, 234
- classifiers 224, 233
- attribute-value matrix 21, 26
- autonomy 173-175, 180, 183
- case (see also prepositions)
- accusative 42, 99, 100, 102, 136, 140, 145, 146, 156, 193
- comitative 196
- dat**i**ve 92, 95, 96, 99, 100, 103, 104, 106-109, 114, 117, 122-125, 129, 130, 135-137, 144-147, 163
- ergative 116, 118, 119, 132
- instrumental 135-148, 150, 154, 158, 159, 163, 164, 169, 171, 191, 207
- nomina**t**ive 99, 100, 146, 147, 193
- case marking 92, 96, 99, 109
- case assignment rules 146, 147
- causal explanati**o**n 16-17
- causati**o**n 10, 25
- causal chain 172
- causal embed**d**ing 173
- causee 171



- causative  
   accomplishment, see  
     accomplishment, causative  
   morphology 10, 14-15  
   paraphrase 10, 14, 16-17  
   structure 24  
   verb, see verb, causative
- change  
   incremental 11-14, 19-20, 23-24  
   ofstate 6, 8, 10, 14, 20, 22-23,  
     25
- clause linkage 241, 242, 244, 253,  
 254
- cleft 262, 267, 283
- clinic 259-265, 269, 270, 273, 281,  
 282
- comparative construction 63-87
- completeness constraint 150, 151,  
 155
- compositionality 49, 50, 55
- co-composition 148, 149
- conjunction 243, 244
- contrastive 263, 265, 274, 275, 280  
   focus 262, 263, 269, 280  
   topic 262, 263, 272-274, 279
- controllability 174, 183
- core 242-246, 249, 258, 262, 264,  
 267, 269-274, 276, 278-280
- degree achievement 7, 11-12, 19-20
- demonstrative, 214-218, 230, 231,  
 235-237
- dependence 243, 247, 248, 252
- Discourse Representation Theory,  
 297, 298, 300
- displacement 257, 271, 273, 274
- embedding 232, 236 244, 247
- frame-based representation 21-22,  
 26-27
- focus 216, 219, 221, 258, 261-263,  
 266-269, 272, 273, 275- 283  
   narrow 262, 290, 291, 294, 297-  
     299,303  
   predicate 273, 290, 299, 302,  
     303  
   sentence 290, 299, 303
- focus structure, 285, 286-294, 297,  
     300, 302, 304, 306
- gender 217, 232, 233
- genitive construction 219, 222, 230
- Generative Lexicon 148
- Generative Semantics 3, 5
- gerund 244
- iconicity, 213-215, 218, 229, 230,  
 234-237
- idioms 49-51,
- illocutionary force 227, 242, 245-  
 249, 251, 252, 259, 269, 274,  
 276, 278
- incorporation  
   argument 141  
   morphological 32  
   pseudo 32, 38-40, 52, 53, 56
- information structure 257, 258, 263,  
 266, 269, 279, 281-283, 285,  
 286, 292-294, 305,306
- Instrument-Subject Alternation 169-  
 187
- interval semantics 5-6
- inversion 257-270, 272, 274-280,  
 282
- juncture 36, 241-244, 246, 248, 249,  
 252,254
- clausal 242, 244, 248, 249, 252
- core 141, 142, 157, 242-246,  
     249
- nuclear 242
- see also juncture-nexus  
     combinations
- juncture-nexus combinations 234
- layered structure  
   adpositional phrase 222
- clause 226-228, 242, 244, 245,  
     249, 251, 252, 289
- noun phrase 234
- reference phrase 221-225, 230-  
     233
- left-detached position 259, 271-280,  
 289
- Lexical-Functional Grammar 246
- light verb construction 32, 34, 36,  
 40, 41, 45, 50-53

- linking 79-83,
  - semantics to syntax 79,
  - syntax to semantics
- logical structure 3-4, 8-10, 14-16, 23-27, 96, 137, 143, 145, 170-177, 182, 190, 196, 199-203, 201-203, 206, 207
- macroroles 98, 99, 116, 122-124, 131, 132, 172, 176, 196
  - non-macrorole arguments 135, 144-148, 150, 158, 162, 163, 190-195
  - see also actor, undergoer
- metonymic clipping 172
- modifier phrase 221, 235
- motion 155
  - caused motion 16-17
  - directed, motion to goal 9, 12, 14-17, 23-24, 27
- natural event condition 176, 177
- negation 264, 266, 268, 274, 277
- negative polarity item, 259, 276-278
- Neo-Davidsonian representation 25
- nexus 36, 241-244, 247, 252, 253
  - coordination 241-247
  - cosubordination 241-253
  - subordination 241-244, 247
- noun classes, 216-221, 225, 232 (see also gender)
- noun phrase, 213-222, 225, 229-237 (see also reference phrase)
  - arguments, 219, 222
  - associative, 217-230, 235, 237
  - NP coordination, 234
  - NP cosubordination, 234
  - NP operators, 234, 237
- nucleus 242, 244, 245, 258, 259, 264, 267, 269, 275, 280
- numeral 213-215, 218, 229-237
- operators 92, 97-101, 103, 109, 243-252
- periphery 129, 270, 271, 276, 278, 280, 282, 283
- possession 113, 118, 217, 219
- postpositions, 95, 102, 103, 105, 109,
  - potential focus domain 251
  - pre-core slot 259, 269-280, 289-291, 294, 303, 304
  - pre-detached position 289, 291, 294, 297, 303
  - predicate 92-97, 99-101, 104, 109, 219, 222, 228
  - prepositions 103,
    - comitative 143, 144, 159-163,
    - instrumental 159-163, 170-172,
  - Principle of Domain Integrity, 236, 237
  - Principle of Head Proximity, 236
  - Principle of Scope, 235, 236
  - privileged syntactic argument 172, 176, 190-194, 198, 200, 205
  - selection principles 193
  - process 11-13, 19-20, 23, 27
  - qualia 148, 150-152, 155
  - reference phrase, 213, 215, 220-222, 229, 234-237 (see also noun phrase)
    - arguments, 235
    - associative, 223, 225, 229
  - reference tracking, 285-289, 292, 293, 297-301, 304, 305
  - referential locus 63, 73, 74, 77, 78, 80, 81, 84
  - relative clause 214
  - result state 11, 19
  - satellites, 214, 229, 230, 235-237
  - scale 20, 22, 24
  - semantic decomposition 3-27
  - sign languages 64, 67, 85, 86
  - signing space, 63-67, 70, 73-84
  - stage 19-24
  - telicity 4, 11
  - tense 242, 243, 246-249, 252
  - thematic relation 171
    - agent 171, 175 (see also agency)
    - effector 171, 172, 175, 176
    - force 171
    - implement 173
    - instrument 136, 138, 140, 144, 159, 162, 169-178, 186, 187
    - location 144

- topic
  - givenness topic 272
  - see also contrastive topic
- undergoer 98-102, 139-144, 146, 150, 158, 172, 190-193, 200-208
- verbs
  - causative 15
  - heavy 34, 35, 54
  - implicative 7
  - light 32, 34, 35, 39, 56
  - of consumption 10, 17, 24
  - of directed perception 18
  - of motion, see motion
- voice 257-259, 267-269, 281
  - antipassive 123, 139, 141, 158
  - passive 42, 43, 94, 139, 156, 158, 160
- word classes 213, 215, 220, 229, 235