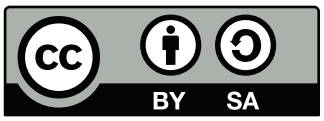
Applying and Expanding Role and Reference Grammar

Rolf Kailuweit, Lisann Künkel, Eva Staudinger

NIHIN Studies

FRIAS
FREIBURG INSTITUTE FOR ADVANCED STUDIES
ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG



Website links in this book

All website links mentioned in this book were operational at the time of publishing. Publications available on the Role and Reference Grammar website are referenced as "available on RRG website", and no links are provided. At the time of publication, the RRG website was accessible at http://www.acsu.buffalo.edu/~rrgpage/rrg.html.

List of abbreviations

ABS	absolutive case	**DET**	determiner
ACC	accusative case	**DIR**	directional
ADJ	adjective	**DM**	discourse marker
ADV	adverb	**DL**	dual
AFFIR	affirmative	**DOB**	direct object
ALL	allative case	**DOM**	differential object marking
ANTI	antipassive	**ERG**	ergative case
ARG	argument	**EVQ**	event quantification
ART	article	**EXL**	exclusive
ASP	aspect	**F**	feminine
AUX	auxiliary verb	**FOC**	focus
CH	case hierarchy	**FUT**	future
CLD	clitic left dislocation	**GEN**	genitive case
CLI	clitic	**IF**	illocutionary force
CLM	clause linkage marker	**INT**	intensifier
COM	comitative	**IO**	indirect object
COMP	comparative	**IMP**	imperative
COMPL	completive	**IPFV**	imperfective aspect
COND	conditional	**IMPS**	impersonal
CONJ	conjunction	**INAN**	inanimate
CONT	continuous	**INC**	inclusive
COP	copula	**INE**	inessive case
DAT	dative case	**INF**	infinitive
DCA	domain of case assignment	**INGR**	ingressive
DCM	differential case marking	**INS**	instrument/instrumental
DEIC	deictic	**INT**	intensifier
DEM	demonstrative	**KP**	kontrast position
DES	desiderative	**LDP**	left detached position

LOC	locative
LS	logical structure
LSC	layered structure of the clause
LSNP	layered structure of the noun phrase
LWDP	leftward detached position
M	masculine
MR	macrorole
N	noun
NEG	negation
NEUT	neuter
NFUT	non-future
NM	noun marker
NMR	non-macrorole argument
NOM	nominative case
NPFP	noun phrase final position
NPIP	noun phrase initial position
NPST	non-past
NSO	non-specific object
NUC	nucleus
O	object
OBL	oblique
P	preposition
PART	partitive case
PASS	passive
PAST/PST	past tense
PASTC	past continuous
PFV	perfective aspect
PL	plural
POSS	possessive
POT	potential
PrCS	pre-core slot
PRED	predicate
PREP	preposition
PRES/PRS	present tense
PRES.ACT	present active participle
PRES.PASS	present passive participle
PRET	preterite
PRO	pronoun
PROC	process
PROG	progressive aspect
PRT	particle
PRV	preverb
PSA	privileged syntactic argument
PSRR	principle of the separation of reference and role
PSTP/PPLE	past participle
PST.PASS	past passive participle
PTCP	participle
PURP	purpose
Q	question word/particle
QD	quantitative determinacy
QTY	quantity
RDP	reduplication
REFL	reflexive
REL	relativizer
RES	result
RP	reference phrase
S	subject
SG	singular
TNS	tense
TOP	topic
U	undergoer
V	verb
VBLZ	verbalizer
VER	version
VN	verbal noun

Contents

Acknowledgements

An edited volume is always the product of a joint effort. We would therefore like to thank all those who supported us in one way or another. First of all, we are grateful to the Freiburg Institute of Advanced Studies (FRIAS), and in particular its director Bernd Kortmann, for their generous support of the present volume. We are also indebted to the anonymous peer reviewers and all those who have helped us to prepare the manuscript, especially Chelsea Brooks, Hannah Davidson, Christa Hohendahl, Anabel Martín, Carmen Pietropaolo and Vanessa Tölke. Finally, we would like to thank Christina Meuser, who guided us through the publishing process.

Introduction

Eva Staudinger[α] & Rolf Kailuweit[β]
Albert-Ludwigs-Universität Freiburg[α]
Heinrich-Heine-Universität Düsseldorf[β]

Role and Reference Grammar (RRG) is a theory adopting a functionalist approach to language. As such, RRG focuses on the analysis of grammatical structures in relation to their semantic and communicative functions. What further characterises RRG is its commitment to typological adequacy (Van Valin & La Polla 1997; Van Valin 2005). Due to its typological bias, RRG is "traditionally" concerned with the description of general characteristics of linguistic systems and their interaction with semantics and pragmatics. The application of RRG in the context of increasingly detailed analyses of linguistic phenomena, however, leads to proposals for further elaborations of various components of the theory. The present volume gathers a series of contributions that apply and expand RRG in different respects.

The book comprises 16 articles divided into five thematic parts: Syntactic structure (layered structure of the clause and operators, Part I), *Aktionsart* (Part II), Macroroles, linking and case assignment (Part III), Complex Sentences (Part IV) and Computational Implementation (Part V).

Part I comprises two papers on different aspects of syntactic structure. In RRG, clause structure is represented in the layered structure of the clause (LSC), which includes different levels of analysis: the constituent and the operator projections. While elements with referential content are part of the constituent projection, non-referential morphemes such as tense, mood, aspect, definiteness, negation etc. are represented in the operator projection. The constituent structure of sentences is represented by means of syntactic templates. Some aspects of syntax, such as the distinction between the nucleus, core and periphery, are considered to be universal (Van Valin & La Polla 1997: 27), but the syntactic inventory is also subject to cross-linguistic variation (e.g. not all languages have a post-core slot and a corresponding syntactic template, cf. Van Valin & La Polla 1997: 37).

Farhad Moezzipour examines to which extent left-dislocation in Persian can be described by means of the syntactic templates provided in RRG. Persian has double possessor con-

structions with two left-dislocated possessors. RRG proposes a left-detached position (LDP) for left-dislocated pre-clausal elements. However, the LDP can only host one element, which poses a problem for the description of Persian double possessor constructions. In a similar vein, neither the LDP nor the pre-core slot (PrCS) position prove apt to adequately capture the status of the left periphery in Persian clitic left-dislocation constructions. Farhad Moezzipour proposes a more detailed distinction of constituents in the left periphery within RRG and provides a detailed pragmatically motivated justification for this.

Brian Nolan examines the encoding of negation in Irish and its representation in the layered structure of the clause (LSC) and the layered structure of the noun phrase (LSNP). Nolan proposes a taxonomy that represents the different cases of negation found in Modern Irish, and analyses their different scope levels. Modern Irish features nuclear, core and clausal negation. In addition to the forms of negation provided at the different levels, Nolan examines the ways in which negation interacts with indefiniteness and quantification.

Part II consists of three papers on *Aktionsart*. In RRG, predicates are classified according to six different *Aktionsart* classes, each of which has a causative counterpart. Robert D. Van Valin Jr. discusses the logical structure (LS) of the *Aktionsart* class of active accomplishments. He revises the representation proposed in Van Valin (2005: 44–45) arguing that it did not adequately capture the incremental nature of many active accomplishments (e.g. *write a poem*) and their underlying activity predicates. In addition, a distinction is drawn between motion and non-motion active accomplishments. Motion active accomplishments can occur with indications of both the distance covered and the goal reached (e.g. *Sally ran two miles to the park*), and the LS should be able to accommodate both of these specifications.

Ranko Matasović analyses aspectual morphology in Slavic. Aspect is usually treated as an inflectional category that is represented in the operator projection. However, Matasović's anaylsis shows that aspectual alternations in Slavic are derivational rather than inflectional in nature, and thus should be captured in terms of lexical rules. They change the *Aktionsart* of the verb and alter the lexical representation. Two additional *Aktionsart* classes previously not included in the RRG *Aktionsart* system are proposed for the description of the Croatian case.

While it is tacitly assumed that all predicates pertain to one of the *Aktionsart* classes proposed in RRG, Lisann Künkel shows that for some predicates *Aktionsart* classification is not as straightforward as it appears. In her analysis of verbs of precipitation in Spanish, she suggests that these predicates cannot be considered prototypical instances of any of the *Aktionsart* classes. She proposes to classify them as non-prototypical activities, whose linking behaviour nevertheless follows the general linking procedures proposed for Spanish.

Part III contains six articles dealing with macroroles, linking and case assignment. Linking is a major topic in RRG. RRG proposes a bidirectional linking algorithm that maps the

semantic to the syntactic representation and vice versa (Van Valin & LaPolla 1997: 340–349, Van Valin 2005: 136–150). Linking from semantics to syntax involves the assignment of case to the macrorole arguments of the predicate. Although RRG relies primarily on general principles (e.g. the linking algorithm, the actor-undergoer hierarchy or the privileged syntactic argument (PSA) selection hierarchy) in order to account for syntactic structures, it nevertheless recognises the importance of constructions (Van Valin 2005: 131–132). Syntactic phenomena such as passive or question formation cannot be derived solely from general principles, but require additional specifications, which are rendered in language-specific constructional schemas. A wide range of linguistic phenomena such as German passive formation (Diedrichsen 2009), Romance anti-causative alternations (Kailuweit 2011) and Irish light verb constructions (Nolan 2014) were analysed using constructional accounts within RRG.

Wataru Nakamura's paper provides an RRG account of dative case assignment in Russian clauses with the semi-predicatives (SemP) *odin* ("alone") and *sam* ("oneself"). Depending on the syntactic environment, SemPs receive different case marking. While actor control constructions only admit nominative case marking, constructions without an argument controller only admit dative case. The case marking found with undergoer control constructions is of particular interest, because they admit both dative and accusative case marking for SemPs. Nakamura provides an explanation for the case marking differences taking into account the properties of the respective constructions, and analyses the respective case assignment procedures within the framework of RRG.

Koen Van Hooste investigates partitive marking in Finnish, which is governed by a complex interplay of factors. Following previous analyses of the Finnish partitive, he examines different contexts in which partitive marking in Finnish occurs, and analyses how to best capture its functions within the framework of RRG. It is shown that partitive marking of objects in Finnish is not only an issue of case assignment, but that also the operator level has to be taken into account. In addition, a constructional schema handling partitive marking of subjects is proposed.

Anna Riccio's paper analyses result-oriented constructions in Italian (e.g. *cortare i capelli corti* – 'cut the hair short'). She discusses in which respect these constructions can be considered resultatives and provides an analysis of their semantic and syntactic structure. Result-oriented adjectival phrases (APs) are treated as a second predicate nucleus that specifies, intensifies or modifies the degree of the action specified by the verb. Her proposal also includes an analysis of the juncture-nexus type and remarks on the linking procedure for result-oriented APs.

Rolf Kailuweit critically reviews the RRG macrorole assignment rules. Macrorole assignment in RRG is based on several heterogeneous criteria: *Aktionsart*, number of arguments

(one or two-place states), syntactic order (leftmost vs. rightmost argument in LS), and even time sequence of processing (assign actor first). However, the macrorole assignment mechanisms fail to predict the linking properties of several verb classes, such as three-place states *(envy)*, non-causative object experiencer verbs *(interest)* or transitive versus intransitive verbs of motion *(run two miles, enter, reach, climb Mont Blanc, climb up Mont Blanc)*. Kailuweit proposes the activity hierarchy as an alternative (cf. Kailuweit 2013), and presents an elaboration of this model in the present volume. Starting from a set of two actor features (c = causation/control, m = mental) and one undergoer feature (r = resultative), the activity hierarchy brings into play semantic criteria that have dominated the discussion on generalised semantic roles in the last decades. It allows for a sufficient number of activity degrees to describe macrorole assignment in the mentioned verbs classes.

Sergio Ibáñez examines the behaviour of commercial event verbs (e.g. *comprar* 'buy', *vender* 'sell') in Spanish. He provides a quantitative corpus-based analysis of the constructional possibilities of these verbs and shows that they are attested with a range of different participant constellations. Commercial event verbs involve up to four participants (the seller, the buyer, the money, the goods). While usually the representation of the logical structure (LS) of a predicate only includes those participants considered to be (direct) core arguments (i.e. the arguments most relevant for linking), Ibañez argues that the number of arguments specified in the predicate LS should coincide with the number of participants specified in the cognitive frame. The selection of participants to be realised and their syntactic coding are then a part of the linking process.

Eva Staudinger provides a syntactic description of light verb constructions (LVCs) in French and Spanish focusing on the semi-compositional predicates *faire peur* ('to scare, to frighten'; lit. 'make fear') and *dar miedo* ('to scare'; lit. 'give fear'). She provides a constructional account that explains the different degrees of syntactic flexibility of the French LVC and its Spanish counterpart. It is argued that constructional schemas for the description of semi-compositional LVCs have to conflate lexical and language specific grammatical information, but that there is no need to stipulate predicate-specific syntactic rules in order to explain their behaviour, given that LVCs rely on general patterns of the grammatical systems they pertain to. Furthermore, it is suggested that due to the different properties of LVCs in different languages, different constructional approaches are required.

Part IV deals with complex sentences. The unmarked pattern for constructing complex sentences (Van Valin 2005: 183–224) consists of the juncture of nuclei with nuclei, cores with cores, clauses with clauses, or sentences with sentences. The syntactic relationships between the units are called nexus relations. Traditionally, only two nexus relations are recognised, coordination and subordination. Following Olson's (1981) analysis of sentence linkage in

Barai (a Papuan language), RRG postulates a third nexus type: cosubordination, which is essentially close, dependent coordination. The dependency affects the operators, that is, in cosubordination, the units compulsorily share one or more operators at the level of juncture.

Complex semantic relations form a continuum that ranges from strong semantic cohesion, where the propositional units linked in the complex structure express facets of a single act or event, to weak semantic cohesion where the units linked express discrete acts or events. The syntactic linking relations are hierarchically ordered by the strength of the syntactic bond between the units, being strongest in nuclear cosubordination and weakest in sentence-level coordination. Basically, RRG assumes an iconic relationship for the interaction of the two hierarchies: the closer the semantic relationship between two statement units, the stronger the syntactic connection with which they are linked (Interclausal Relations Hierarchy).

Lilián Guerrero explores whether Yaqui syntactic linkages are compatible (or not) with the cross-linguistic predictions made by the Interclausal Relations Hierarchy. In Yaqui, purpose, reason and temporal relations are expressed by multiple syntactic structures. The Interclausal Relations Hierarchy (Van Valin 2005: 209) predicts a semantic scale where purpose outranks reason and reason outranks temporally ordered events. However, Yaqui does not conform to the expected rankings. In this language, simultaneous and sequential clauses show a higher degree of syntactic bonding in comparison to reason clauses, but a less tight linkage when compared with purpose structures. Yaqui provides evidence for re-examining the ranking of the Interclausal Semantic Hierarchy. Guerrero suggests a revision, such as purpose outranking simultaneous and sequential relations, and then reason outranking inferred temporally ordered events. Further studies may verify the cross-linguistic relevance of the proposed semantic scale.

Soteria Svorou shows that Modern Greek V1 *ke* V2 constructions exhibit juncture-nexus relations of core cosubordination and present a closely knit semantic integration of the lexical templates of the two verbs. The analysis could be extended to English data as well. The constructions examined follow the iconicity principle, confirming the Interclausal Relations Hierarchy, in that core cosubordination is interfaced with semantically integrated conjunction of verbs. The language specific properties of the constructions concerning syntax, morphology, semantics and pragmatics are represented in a constructional schema. The author notices an intricate interaction of discourse structures with some Modern Greek V1 *ke* V2 constructions. The integrated $V_{pi\gamma eno/pao}$ *ke* V_2 construction codes a particular understanding of surprise on the part of the speaker in discourse settings of contrast, with a preceding adverbial clause (e.g. 'although', 'instead'), a contrastive particle, in a rhetorical question or exclamatory statement that further research may describe in a more systematic manner.

Luis París deals with the directed motion subtype (DM) of the Spanish Gerund Construction (SGC) consisting of a directed motion matrix verb adverbially modified by a gerund

phrase (GP) without the mediation of any explicit linking marker. París concludes that DM comprises an ad-nuclear subordination linkage. The analysis of DM has consequences for the understanding of the Interclausal Relations Hierarchy. First, it is shown that an adverbial modifier that holds a manner relation to the matrix verb might behave like a complement. This fact seems central for RRG since the linking algorithm projects lexically required information into the core whereas non-lexically required elements occupy the periphery of the core. The examined DM constructions suggest that this crucial distinction should not be bluntly projected to the entire syntactic distribution of adjuncts. Adjuncts are characterised by variable linking. Second, the syntactic realisation of adjuncts might be sensitive to the semantic class of the matrix verb. The difference in the syntactic realisation of DM and non-DM meanings appears to be motivated by the semantic class of the matrix verb.

Part V presents computational approaches within RRG. It comprises one paper on an RRG based learning tool, and a paper addressing issues arising in a principled formalisation of representations and procedures underlying the RRG framework. Computational approaches require a detailed description of the phenomena to be analysed and a stringent formalisation of the analytical models used. They put linguistic theory to the test.

Judith Gottschalk presents a model for a semantic parser to be used within a computer-supported vocabulary learning tool for Biblical Hebrew. The idea is to automatically disambiguate the meaning of a verb by analysing the syntactic and semantic context in which it appears, and to provide learners with an enriched representation of lexical entries. It is proposed to integrate an ontology in the implementation of the syntax to semantics linking algorithm in order to allow for a semantically informed syntactic analysis.

The paper by Rainer Osswald and Laura Kallmeyer highlights that a principled formalisation of a theory of grammar like RRG requires rigid analyses. Their outline of a formalisation of RRG draws a clear distinction between declarative and procedural elements and puts emphasis on syntactic and semantic compositionality. They propose a modular description of syntactic templates in which more complex structures are derived from basic templates by means of the procedures of substitution, adjunction, and wrapping, and they provide a brief sketch of the linking procedure using the English resultative construction as an example.

References

Diedrichsen, Elke. 2009. Towards a reconsideration of Constructional Schemas in RRG: Are all constructions driven by “constructions”? [Available on RRG website]

Kailuweit, Rolf. 2011. Romance Anticausatives: A Constructionist RRG Approach, In *New Perspectives in Role and Reference Grammar*, W. Nakamura (ed.), 104–133. Newcastle upon Tyne: Cambridge Scholars Publishing.

—. 2013. Radical Role and Reference Grammar (RRRG): A sketch for remodelling the Syntax-Semantics-Interface. In B. Nolan & E. Diedrichsen (eds), 103–141. *Linking Constructions into Functional Linguistics. The Role of Constructions in Grammar.* Amsterdam: John Benjamins.

Nolan, Brian. 2014. Complex predicates and light verb constructions in Modern Irish. *Revista Española de Lingüística Aplicada* 27(1): 140–167.

Olson, Michael L. 1981. *Barai clause junctures: toward a functional theory of interclausal relations.* Unpublished Ph.D dissertation, Australian National University.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface.* Cambridge: CUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax, Structure, Meaning and Function.* Cambridge: CUP.

Toward Extending the Scope of the Left Periphery in RRG – Evidence from the Clitic Left-Dislocation in Persian

Farhad Moezzipour[α] & Haleh Cheraghi[β]
Shahrood University of Technology[α]
Hakim Sabzevari University[β]

Abstract

In this paper, we are concerned with a re-examination of the pragmatically motivated left-peripheral positions of Role and Reference Grammar, including the pre-core slot (PrCS) and the left-detached position (LDP), utilizing the Persian clitic left-dislocation construction in order to show that the current status of the left periphery in RRG cannot adequately account for an array of information-structural properties of left-dislocated NPs in the left periphery of the Persian core.

1 Introduction

In the present paper, I am going to explore the information-structural implications of the Persian clitic left-dislocation for expanding the domain of the core periphery in RRG as it is presumed that the PrCS and LDP cannot transparently represent how the discourse-functional properties of left-dislocated elements are mapped syntactically onto a corresponding structural position in the left periphery. I will argue that the conception of the LDP is informationally static in the sense that it fails to capture the topicality gradience in relation to the double external possession construction, where the possessors are linearly arranged in the left periphery in order of discourse saliency. Hence, it is suggested that the replacement of the LDP with the conception of the Leftward Detached Position (LWDP) enables us, on the one hand, to syntactically project two left-dislocated constituents akin to double external possession in the Persian language in LDP_1 and LDP_2 anchored by the superordinate node of the LWDP and demonstrate, on the other hand, the pragmatically higher saliency of the first external possessor in comparison with the second one. I will also argue for the existence of another position

in the left periphery referred to as kontrast position (KP), which is allocated to a contrastive focus that selects the positive value of the feature [±exhaustive] and [±contrastive] as opposed to an exhaustive focus merely with the positive value of the features [±exhaustive] residing in the PrCS.

The organization of this paper is as follows. I will present a discourse-pragmatic sketch of clitic left-dislocation constructions in Section 2, maintaining that the role a left-dislocated constituent plays in the left periphery can vary from a primary topic and secondary topic to a contrastive focus. Section 3 is devoted to the RRG treatment of double external possession introducing the internal structure of the LWDP hosting a sequence of two external possessors together with integrating a Centering-based model of discourse to capture the notion of discourse saliency in the language of RRG. Section 4 deals with the RRG treatment of contrastive clitic left-dislocation presenting the core-external position KP. Section 5 concludes the paper.

2 The Persian clitic left-dislocation: A discourse-pragmatic sketch

Clitic left-dislocation (CLD) throughout this paper is understood as the syntactic displacement of an NP from an oblique or possessive phrase conditioned by the lodgment of a resumptive clitic in the canonical position of the moved element. The dislocation from an unmarked discourse-pragmatic perspective is intended to establish an *as-for-x* or aboutness relationship between the referent of the dislocated constituent and the proposition expressed in the clause. The topical status of the dislocated NP is demonstrated by the appearance of *ra*, which is essentially governed by a syntactic constraint, viz. extraction from a possessive construction necessitates *ra*-marking of the dislocated possessor NP insofar as it is not in the subject position (Dabir-Moghaddam 1992: 562). Left-dislocated possessors are subsumed under the category of *external possession constructions*.

Following Payne & Barshi (1999: 3), an external possession construction in Persian is characterized as a construction in which a semantic possessor-possessum relation is expressed by coding the possessor as a core grammatical relation of the verb and in a constituent separate from that which contains the possessum. However, an external possession construction does not neatly fit into this definition because the possessor is not coded as a direct core argument. Rather, it is the possessum that is affected by the state of affairs, which is expressed by the verb. The external position of the possessor is attested by the removal of the *Ezafe*/genitive linker *-e* in canonical possessive constructions. The sentence in (1a) represents a canonical possessive construction, whereas one in (1b) represents an external possession construction.[1]

[1] Since *ra* is assumed to have multiple functions, it is glossed as RA in this paper. The genitive linker is also glossed as EZ, which is the abbreviated form for *Ezafe*.

(1) a. Ketab-e qæzæl-o peyda kærd-æm.
book-EZ Qazal-RA find do.PST-1SG
'I found Qazal's book.'

b. Qæzæl-o, ketab-eš-o peyda kærd-æm.
Qazal-RA book-CLI.3SG-RA find do.PST-1SG
'As for Qazal, I found her book.'

The diachronic survey of *ra* from Old Persian to Contemporary Modern Persian reveals the fact that there has been a process of grammaticalization according to which *ra* that used to appear with oblique NPs for the expression of purpose and goal has been grammaticalized to denote accusative objects (Hopper & Traugott 2003: 166–168). Assuming this, Dabir-Moghaddam (1992: 564) believes that the syntactic stabilization of *ra* as a direct object marker in the Modern Persian period has led to its pragmatic function as a secondary topic maker irrespective of the grammatical function of the *ra*-marked NP; whether it co-occurs with direct objects, left-dislocated obliques, or external possessors.

The examples in (2)-(3) display left-dislocation of an oblique object and a possessor from a prepositional phrase and a possessive NP. The external possessor in (3) does not appear with *ra* because left-dislocation has applied to a possessive NP in the subject position. Another point relates to the default interpretation of the left-dislocated NPs, which is topical. For this reason, the *as-for-x* phrase is inserted in the glosses. Nonetheless, we will see that left-dislocation can serve as a contrastive focusing tool depending on the discourse context to which it applies.

(2) Qæzæl-o, ketab-o be-š dad-æm.
Qazal-RA book-RA to-CLI.3SG give.PST-1SG
'As for Qazal, I gave her the book.'

(3) Qæzæl, ketab-eš pare šod-e.
Qazal, book-CLI.3SG torn become.PST-PTCP
'As for Qazal, her book became torn.'

2.1 Double external possession and the non-uniqueness of topicality

Left-dislocation is also applicable to a sequence of two possessors embedded within a complex possessive NP in Persian, where the two external possessors, while residing in the left periphery, are marked by *ra*. As a result of this, two pronominal clitics which represent the original position of the two external possessors appear in the sentence. Dabir-Moghaddam (2005: 263) as the first to make this observation claims that the multiple occurrences of *ra* with NPs in double external possession indicate that the notion of secondary topic is scalar.

The example in (4) demonstrates a sentence whose left periphery hosts two external possessors.

(4) Bæčče-ro pirhæn-ešh-o, dogme-šh-o be-duz.
kid-RA shirt-CLI.3SG-RA button-CLI.3SG-RA IMP-sew.2SG
'As for the kid's shirt, sew its button.' (Dabir-Moghaddam 2005: 129)

It is worthy of note that the non-uniqueness of topic has been cross-linguistically attested, and it is conceivable, as Dalrymple & Nikolaeva (2011: 53) explicitly point out, to have "more than one referent under discussion at the time of utterance, so that the utterance simultaneously increases the addressee's knowledge about both of them". However, the existence of multiple topical entities at the utterance time seems to be beyond the cognitive limitations of the human mind. For instance, Lambrecht (1981, 2001) argues that the number of pre-verbal topics in Non-standard French cannot exceed two. One example in support of his claim deals with non-canonical possessives in dislocation constructions. In (5a), the syntactically independent external possessor *mon frère* 'my brother' and the possessum *sa voiture* 'his car' cannot be preceded by a left-dislocated free pronoun *moi* 'me', but if they are restored to the traditional possessive construction of French, the result becomes grammatical and pragmatically acceptable. In (5b), a canonical genitive phrase *la voiture de mon frère* 'my brother's vehicle' is a canonical possessive, and the number of topical entities including the dislocated pronoun equals two.[2]

(5) a. * Moi mon frère sa voiture ne m'intéresse pas.
b. Moi la voiture de mon frère ne m'intéresse pas.
(Lambrecht 1981: 74)

The contention regarding multiple topic constructions has a remarkable history in Generative Grammar. Rizzi (1997) acknowledges topic recursion in the left periphery in Italian. He proposes that there are two topic fields in the left periphery that are assigned two different syntactic positions: one above and one beneath the Focus Phrase (*ibid:* 297). Moreover, under an adjunction analysis for topics and the associated reiterability assumption, "nothing excludes that a comment [...] may be articulated in turn as a topic-comment structure, so that the topic phrases can undergo free recursions" (*ibid*). Frascarelli & Hinterhölzl (2007), however, argue that topics are not contra Rizzi's proposal freely recursive. Relying on data from German and Italian, it turns out that only one constituent per clause can function as an aboutness topic. They distinguish three types of topic that are realized differently phonologically and syntactically projected in a specific order in the complementizer domain, which is a

[2] The relevant glosses for the sentences in (5) are missing in the original text.

direct corollary of their disparate discourse-functional properties. This classification involves an *aboutness topic*, a *contrastive topic* and a *familiar topic*. The hierarchical order for these is given in (6). The example in (7), extracted from a discourse chunk, confirms that in Italian the number of topics in the left periphery is limited, too, and particularly that simply one aboutness topic in the sentence exists. The direct object *Inglese*, resumed by the clitic *lo*, represents a familiar topic whose referent has already been established in the prior discourse and serves as a continuing topic, say, a discourse topic in the sense of Givón (1983). The subject *io*, on the other hand, is considered to be an aboutness topic, which locates the speaker at the center of the conversation and consequently enforces a shift of attention from the previously grounded continuing topic.

(6) Shifting topic [+ aboutness] > contrastive topic > familiar topic

(7) Io, inglese non l' avevo mai fatto.
1SG English not it(CLI) have.PST.1SG never done
'I never studied English before.' (Frascarelli & Hinterhölzl 2007: 94)

The uniqueness of the aboutness topic, simply defined as an entity that the proposition is about, has a precursor in Reinhart (1982). She surmises that a sentence can only have a single topic even though it is obviously possible to talk about entities that are previously mentioned in the discourse or available situationally in the speech context. In the same vein, disciples of the discourse-based *Centering Theory* are in favor of the uniqueness hypothesis of topicality, since one of the constraints in the model is based on the fact that there is precisely one *backward-looking center*, closely identified with the topic in general and the *psychological focus* of Gundel (1999), for each utterance in a discourse segment (Walker *et al.* 1998: 3).[3]

Understood broadly from the above, there seems to be no consensus on the uniqueness of topicality in the sense of pragmatic aboutness. Some advocate the singleness of the aboutness topic, whereas some surmise that topicality can be dual, which means that two entities can be talked about at the same time with reference to the speaker's assumptions about the addressee's state of interest in getting updated on a relationship between them. What these approaches have in common is nonetheless the fact that topicality could be a gradient concept, assuming a hierarchy of topicality rather than a unique definition irrespective of the type of topics or the kind of terminology used for describing them. Lambrecht (*ibid:* 150), interestingly, adds that perhaps it is a matter of "the pragmatic salience of the various topic referents at given points in the discourse, not the difference between topics and non-topics". The implication here is that two topic expressions can be under discussion at the time of utterance

[3] Gundel (1999: 294) defines a psychological focus in terms of an entity on which both speech participants are assumed to be focused because of its salience at a given point in the discourse.

albeit one could override the other in topicality in terms of higher discourse saliency, which in reality encodes the more salient entity as the primary topic and the less salient one as the secondary topic. It is clear that saliency ad hoc, or putting it another way, being in the current awareness of the interlocutors (to a varying degree though), is the common denominator between the primary and secondary topic.

With that being said, we come to two conclusions in relation to double external possession in Persian. The first is that postulating *ra* as a marker of secondary topic (Dabir-Moghaddam 1992) would be problematic as we would have to assume that in (4), there are three secondary topics including an unmarked secondary topic, namely the direct object, and two marked secondary topics realized syntactically as external possessors in the left periphery. The markedness of secondary topics here appears to be characterized in terms of the non-canonical syntax of possessors (Dabir-Moghaddam 2005: 126). This standpoint is not compelling because it is axiomatic that the overall concept of topicality can be portrayed as non-unique in a crosslinguistic perspective, as is clear from the distinction between the primary topic vs. secondary topic or shifting topic vs. familiar topic. But, the premise concerning the gradience of the secondary topic on its own can be dispensed with, not only because a triple topic sequence has remained unprecedented (Dalrymple & Nikolaeva 2011: 53), but also, the limited cognitive capacity imposes restrictions on the mind to eschew preserving several entities under discussion at the same time. A reminder here is that the presence of the secondary topic is tightly bound to the primary topic, implying that no secondary topic is perceptible when there is no primary topic (*ibid:* 57). Therefore, if the three *ra*-marked NPs were to be construed as secondary topics, there would have to be a primary topic in the sentence, too. In this case, the imperative speech act allows us to treat the situationally present addressee as the primary topic, which indeed aggravates the situation because the total number of topics rises to four. Moreover, since Persian allows double extraction of possessor NPs from a possessive construction, this intriguing question comes to mind: What if the two external possessors do not take *ra*? Does it imply that they should be primary topics merely due to the absence of *ra*? The answer without a shadow of a doubt is negative, because it is not plausible to presume that even the primary topic can be gradient, let alone the secondary topic. This can be observed in (8) in which the external possessors illustrate inalienable possession with kinship and part-whole associations.

(8) Færzad xanum-eš kolliy-æš æz kar oftad-e.
Farzad wife-CLI.3SG kidney-CLI.3SG from work fall.PTCP-3SG
'As for Farzad's wife, her kidney has stopped functioning.'

The second conclusion, which is a reply to the question of 'what are the two external possessors then, informationally', is that the external possessors act respectively as primary and secondary topics of the sentence. Nonetheless, it needs to be borne in mind that the illocutionary force in (4) may give rise to a minor deviation from the prototypical definition of the topic as an entity the proposition is about, because the predication is in fact a non-assertive proposition that is lexico-grammatically structured in an imperative speech act form. This does not correspond precisely to the definition of the topic in an assertive speech act. However, the following definition from Gundel (1988: 210) does not rule out the possibility for an entity to be construed as a topic even in a non-assertive speech act: "An entity, E, is the topic of a sentence, S, if and only if in using S the speaker intends to increase the addressee's knowledge about, request information about, or otherwise get the addressee to act with respect to E."[4]

2.2 Left-dislocation vs. primary and secondary topicalization

In this subsection, I will first sketch out the discourse-functional motivation underlying left-dislocation in line with Lambrecht's (1994) treatment of detached constructions. Lambrecht considers left-dislocation as a grammatical means utilized in order to code an NP referent of a certain degree of pragmatic accessibility as the topic of a sentence in a "syntactically autonomous or detached position to the left of the clause which contains the propositional information about the topic referent" (*ibid:* 182). In Lambrecht's view, a topic referent and topic expression are distinguished in the fashion that the former is understood as an entity which the proposition is about, in the sense of conveying relevant information about that entity; hence increasing the addressee's knowledge in relation to it. The topic expression, on the other hand, is a linguistic expression designating the topic referent of which some newsworthy information is predicated. Given this, the left-dislocated NP in a dislocation construction is regarded as the topic referent and the pronominal clitic as the topic expression whose function is to indicate the referent's role as an argument in a proposition via co-referencing within the clause. The clause-external coding of the topic in a detached position and its concomitant intra-clausal indexation is captured by a pragmatic maxim entitled the Principle of the Separation of Reference and Role (PSRR henceforth), which states: "Do not introduce a referent and talk about it in the same clause (*ibid:* 184)." Consequently, left-dislocation is a grammatical device that serves a topic-announcing/topic-establishing function (Lambrecht *ibid:* 188 and

[4] An anonymous reviewer of the paper suggested using the canonical assertive form of (4). The point is that the assertive or imperative force of the sentence representing an external possession construction does not affect the analysis proposed here. For instance, in (8) two consecutive external possessors happen in an assertive sentence, and each serves as a primary and secondary topic respectively. The sentence in (4) is also grammatical when the predication carries an assertive force.

Gregory and Michaelis 2000) for the introduction or establishment of a 'new topic' whose referent has an accessible status in the discourse-internal or discourse-external world. In an endnote, Lambrecht (*ibid:* 353) asserts that introducing a new topic is meant to be the "new coding of an accessible referent as a topic expression".

Before going any further, we should come up with a more accurate and concise definition for a secondary topic. Lambrecht (1994), Nikolaeva (2001) and Dalrymple & Nikolaeva (2011) consider the secondary topic as an entity that stands in a relationship with the primary topic while the asserted portion of the proposition conveys not only information about the topic referents, but also about the relationship between them. The example in (9) helps us capture the distinction between the primary topic and secondary topic. In (9c), the subject and the object are respectively the primary and secondary topic of the sentence, and the communicative purpose of this utterance is to inform the addressee of the relationship between the topic referents. The information structure for (9c) is represented as in (10).

(9) a. Whatever became of John?
b. He married Rosa.
c. But he didn't really love her. (Lambrecht 1994: 148)

(10) Pragmatic presupposition: John stands in the relation X to Rosa
Pragmatic assertion: X = didn't love
Focus: didn't really love (Dalrymple & Nikolaeva 2011: 54)

Taking PSRR into consideration, I argue that left-dislocated NPs are in fact primary topics of the sentence in the light of left-dislocation as a topic-promoting device. Furthermore, the hypothesis that secondary topicality is the commonality shared by entire occurrences of *ra* has been perspicuously laid out. Let us take a look at example (11).

(11) Context
a. Næzær-et dærbare-(y)e færhad chi-(y)e?
opinion-CLI.2SG about-EZ Farhad what-be.PRS.3SG
'What is your opinion about Farhad?'
b. Æz-aš xoš-æm ne-mi-(y)a-d.
from-CLI.3SG like-CLI.1SG NEG-IPFV-come.PRS-3SG
'I don't like him.'
c. Mæn-o, be-(h)em mi-xænd-e.
1SG-RA to-CLI.1SG IPFV-laugh.PRS-3SG
'Me, he laughs at.'

From the standpoint of Dalrymple & Nikolaeva (*ibid*), it can be tentatively discussed that the detached pronoun *mæn* 'I' is a secondary topic because it holds a relationship with the

primary topic of the sentence, i.e. *Farhad*. The focus of the proposition deals with the asserted information about the relationship between the two topic referents, which is expressed by the verb. Accordingly, the information structure for the second sentence in (11a) can be potentially depicted as in (12).

(12) Pragmatic presupposition: *Farhad stands in the relation* X *to the speaker.*
Pragmatic assertion: X = *laughs at*
Focus: *laughs at*

But this is not the whole story and the representation in (12) fails to capture the exact communicative purpose served by the placement of the detached NP in an extra-clausal slot. The speaker, represented syntactically as a left-dislocated first person free pronoun, in uttering this sentence is inclined to direct the addressee's attention from Farhad as the discourse topic to himself. In other words, *Farhad* is introduced as a third party into the discourse, and the subsequent sentences are supposed to be about him. But, the second sentence in (11b) particularly brings the speaker to the very center of the addressee's attention as the entity the proposition is about. This is also in accordance with the topic-promoting function of left-dislocation in the way that the bound clitics, i.e. *et* and *æm* represent the activation status of the speaker as a potentially accessible topic referent, but employing a free first person pronoun is a clear indication of the speaker's interest in not only enforcing a topic shift on the discourse level, but also in increasing the addressee's knowledge about himself on the sentence level. Strictly speaking, the asserted part of the proposition conveys information about the speaker as the primary and only topic of the sentence, which stands in a secondary relationship to *Farhad* as the entity that the discourse is about. The status of a detached NP in a CLD construction as a secondary topic is obtained in terms its lower degree of discourse saliency vis-à-vis the higher saliency of another referent already evoked on the discourse level.

CLD, as a matter of fact, is a discourse-based operation undertaken in order to introduce a detached NP as the new topic, say, primary topic of the sentence to satisfy the consciousness presupposition and relevance/topicality presupposition put forth by Lambrecht (1994: 53–54), which are the pre-requirements for an entity to be construed as a topic. For a detached NP to be treated as the topic, the mental representation of its referent should be activated in the interlocutors' short-term memory; moreover, the addressee's knowledge store with respect to the referent currently under discussion is updated by the relevant information expressed by the clause at the utterance time.

An interesting piece of evidence in favor of the primary status of a detached NP in Persian CLD constructions is provided by the extraction of two consecutive possessors from a possessive phrase, as already exemplified by (4). From my point of view, the external possessors

are respectively the primary and secondary topic of the sentence placed in the left periphery of the core in order of their saliency in the discourse. The referents of the two external possessors are coded as new topics on the sentence level, and the inalienable possessum, despite its cognitive accessibility via the part-whole relationship with the second external possessor, constitutes part of the actual focus together with the verb. The identifiability of an inalienable entity does not rule out the possibility for it to act as a participant directly affected by the state of affairs expressed by the predicate and consequently contained within the focus of the sentence. The secondary status of the second external possessor to the primary status of the first external possessor regarding topicality, which is also visible via the order of pronominal clitics, is describable in terms of the comparatively higher degree of saliency of the former in discourse.

The discourse-functional motivation in the ascension of two possessors is compatible with the reference point construction in Langacker (1993). A reference point construction is an aspect of the human's cognitive processing which is capable of invoking the "conception of an entity for purposes of establishing mental contact with another, i.e. to single it out for individual conscious awareness" (*ibid:* 5). In (4), the referents of the two external possessors, whose activation statuses are expected to be accessible in the discourse, are conceptualized (i.e. promoted from a discourse-functional perspective) to serve as cognitive reference points for the identification of the inalienable possessum. This is in line with the propensity of possessor nominals to be coded as sentence or discourse topics by virtue of having already evoked referents in the discourse as attested by Taylor (1996: 213). Kockelman (2009: 25) in the same vein considers the inalienable possession as a grammatical category akin to deictics and prepositions that "guide the addressee's identification of a referent by encoding that referent's relation to a ground".

2.3 Left-dislocation vs. contrast

In this section, I intend to investigate the role of CLD as a grammatical means for the exhibition of contrast in addition to topicality, which was discussed in the previous subsection. This examination shows us that the status of clitic left-dislocated constituents in the left periphery of the clause is not only reducible to topicality; instead, these constituents can be contrastive depending on the particular context in which they appear. There is a trend in the ample literature on information structure toward the treatment of contrast as an independent category that should be integrated into the inventory of basic information-structural notions, and should not be regarded as a feature of topicality and focusing alone (see Vallduvi & Vilkuna 1998, Monlár 2002, 2006, McCoy 2003, Umbach 2004, Neeleman *et al.* 2009, Neeleman & Vermeulen

2012). I will argue along the lines of É. Kiss (1998) and Monlár (2006) that the contrastive function of CLD is due to the selection of an alternative from a complementary set whose members have to be explicitly mentioned in the discourse.

To begin with, I would like to cite the distinction between an information focus and identificational focus as proposed by É. Kiss (*ibid*). The information focus represents the new, non-presupposed part of a proposition that is syntactically manifested by in-situ constituents, such as in-situ interrogative elements and the narrow focus NPs replacing them. The identificational focus, on the other hand, is characterized as the exhaustive subset of a set of contextually or situationally given elements for which the predicate phrase actually holds (*ibid:* 245). É. Kiss (*ibid:* 267) also demonstrates that the range of representing an identificational focus is subject to cross-linguistic tendencies and determined with respect to the positive value of either or both of the features [±exhaustive] and [±contrastive]. For instance, in Hungarian, the immediate pre-verbal position can be assigned to an identificational focus that operates on a closed set of contrasting elements explicitly mentioned in the context, in which case the identificational focus selects both of the positive values of the above features, e.g. (13). The identifiability of the set members is so crucial to the parametric variations of the identificational focus in this language that its applicability to an open set of entities does not guarantee its contrastiveness, since the identification of the subset for which the predicate holds does not give rise to the delineation of a complementary set with clearly identifiable members. Consider (14) in which the identificational focus is [+exhaustive] and [-contrastive].

(13) Context:

a. Who introduced Peter to Mary, John or his brothers?

b. Pétert JÁNOS muttatta be Marinak.
Peter-ACC John presented PRV Mari-DAT
'It was John who introduced Peter to Mary.' (Monlár 2006: 202)

(14) Context:

a. Who wrote War and Peace?

b. Háború és békét Tolsztoj írta.
war and peace.ACC Tolstoy wrote
'It was Tolstoy who wrote War and Peace.' (Monlár 2006: 202)

The interplay of an identificational focus with contrastiveness and the position signalling it is applicable to Persian, too. The clause-initial position in the structure of the Persian sentence can be allocated to an identificational focus. Its contrastiveness is heavily dependent on explicitly mentioning the contrasting alternatives in the discourse, and consequently selecting the positive value of the features [±contrastive]. Furthermore, the representation of

the identificational focus, along with the choice of either the positive or negative value of the features [±contrastive] on the side of the speaker, is reflected by the employment of different syntactic means. Clause-initial preposing is used to demonstrate the features [+exhaustive] and [-contrastive], while clitic left-dislocation has to do with the features [+exhaustive] and [+contrastive]. For the sake of exposition, I refer to the identificational focus with the features [+exhaustive, -contrastive] as *exhaustive focus* and the identificational focus with the features [+exhaustive, +contrastive] as *contrastive focus*. Consider the examples in (15).

(15) Context:

a. Dišæb to mehmuni ba ki ræqs-id-i?
last.night in party with whom dance-PST-2SG?
'Whom did you dance with at the party last night?'

b. Ba MEHRDAD ræqs-id-am.
With Mehrdad dance-PST-1SG
'I danced with Mehrdad.'

b.' Ba MEHRDAD mæn ræqs-id-am.
with Mehrdad 1SG dance-PST-1SG
'Mehrdad I danced with.'

In (15a), the wh-element is in-situ seeking new information, and the placement of the comitative NP in the unmarked pre-verbal focus position in (15b) satisfies the unspecified value of the presupposition expressed in the question: There is an x with whom the addressee danced at the party last night. In (15b'), we are faced, on the other hand, with a case of an exhaustive focus implicitly evoking a set of alternatives one of which exhaustively identifies the focus value of the wh-element; what É. Kiss (*ibid*) calls *exhaustive identification*. Despite the in-situ position of the wh-element, the focused comitative NP has been preposed to the clause-initial position in order to display that the exhaustive focus is the only value that satisfies the presupposition in the question without explicitly mentioning other alternatives which have been capable of falling into the set of people who the speaker could have danced with at the party. Put another way, the speaker and the addressee are aware of the people who were present at the party at the time of the conversation, and the speaker, by the clause-initial positioning of the focused constituent instead of the in-situ positioning, conveys this message to the addressee that *Mehrdad* was the only person she danced with at the party, which automatically excludes the other people from the set of potential alternatives. Therefore, *Mehrdad* represents an in-situ information focus in (15b) and an exhaustive focus in (15b'). Keep in mind that *Mehrdad* in the latter is not a contrastive focus, since the positive value of the feature [±contrastive] has not been selected due the absence of an explicit set of alternatives that makes the operation of a contrasting task possible. Now take a look at the example in (16).

(16) Context:

a. Dišæb to mehmuni ki-ro ba-(h)aš ræqs-id-i, mehrdad ya
last.night in party whom-RA with-CLI.3SG dance-PST-2SG Mehrdad or
færzad?
Farzad?

'Whom did you dance with at the party last night, Mehrdad or Farzad?'

b. MEHRDAD-o mæn ba-(h)aš ræqs-id-æm.
Mehrdad-RA 1SG with-CLI.3SG dance-PST-1SG

'It was Mehrdad that I danced with.'

The left-dislocated comitative NP in (16b) is a contrastive focus because not only does it identify the proper value of the relevant set exhaustively, but also the contrasting operation gives rise to the exclusion of the other explicitly mentioned alternative, *Farzad*. Therefore, the treatment of *Mehrdad* as a contrastive focus is due to the feature [+exhaustive, +contrastive]. Based on this observation, I can define three kinds of focus in Persian in terms of the features [±exhaustive] and [±contrastive], bearing in mind that contrastiveness in this classification is regarded as the optional correlate of the exhaustive identification in such a way that that an identificational focus can represent a contrastive focus if the feature [+contrastive] is added to the feature [+exhaustive] in the presence of an explicit set of alternatives. The points are summarized in Table 1.

The example in (16) reveals a syntactic tendency for the left periphery of the clause to be occupied by contrastive foci, such that CLD in the Persian language is directly related to the depiction of contrast, which consolidates the hypothesis proposed by Monlár (2006: 213) that contrast should be superimposed as an independently motivated category of information structure on topic and focus. She, in the light of data from Finnish, introduces a type of contrast that is embodied by left-dislocated constituents in a syntactically autonomous position; the type of contrast and the syntactic position are referred to as S-kontrast and kontrast-position (*ibid:* 224). However, there is a pragmatic condition that needs to be met by the left-dislocated constituent: It must refer to "explicitly mentioned or at least contextually salient alternatives within a closed set" (Monlár *ibid:* 225).

focus type	[±exhaustive]	[±contrastive]	syntactic means of expression
information focus	–	–	in-situ positioning
exhaustive focus	+	–	clause-initial preposing
contrastive focus	+	+	clitic left-dislocation

Table 1: Focus taxonomy in Persian

3 RRG treatment of double external possession

In this section, I intend to show that the current status of the left-detached position (LDP) in the left periphery of the core cannot adequately account for double external possession in Persian. Space limitation forces me to do away with the RRG treatment of left-dislocation constructions where there is a single detached constituent, either to an external possessor or an oblique in the LDP, e.g. (1b) and (2). However, it is incumbent to look at the sequence of two external possessors, e.g. (4) and (8), in which there are two detached constituents in the left periphery of the core.

Whether RRG is able to project two left-dislocated elements syntactically or not would be a challenge. It is obvious that RRG adopts a static stance toward the notion of topicality, and the syntactic conception of the LDP fails to capture the topicality gradience of elements located in the left periphery in order of saliency. I have discussed in Section 2 that the possibility for the extraction of two possessors from a complex possessive phrase indicates that the notion of secondary topic in relation to CLD is defined in terms of the higher saliency of the first external possessor in comparison to the second external possessor. Therefore, the existence of a single LDP in the structure of the sentence would not be sufficient for the syntactic representation of two external possessors positioned hierarchically in the left periphery of the core. To represent the fact that the first external possessor reveals a higher degree of saliency than the second one, I propose that the LDP should be replaced by the conception of Leftward Detached Position (LWDP) anchoring LDP_1 and LDP_2. Hence, it enables us to syntactically locate the two topical external possessors in the left periphery in addition to the fact that the element in LDP_1 is considered to be more topical, in the sense of being more salient, than the one in LDP_2, which is a direct corollary of the positional precedence of the former over the latter.

The information structure of the sentence in RRG is related, on the one hand, to the pragmatic relations of topic and focus represented by the focus structure projection and to the activation status of discourse referents represented in the logical structure of the sentence on the other. An issue regarding the two external possessors is that both are cognitively of an identical status, each accessible as predicted by PSRR, and the semantic structure of the sentence is only able to demonstrate the cognitive category associated with the two external possessors statically. To capture the relativity of the information status of the external possessors, namely that the first one is more activated/salient than the second one, I integrate a Centering-theoretic model of discourse into RRG to reinforce the parallel cooperation between the domains of semantics and discourse.

Centers in Centering Theory (see Walker *et al.* 1998 and references therein) are discourse entities that are part of the discourse model for each utterance in a discourse segment. There are two basic types of centers: *forward-looking centers* and *backward-looking centers*. Forward-

looking centers (Cfs) are a set of discourse entities evoked by an utterance U_i in a discourse segment D, which are ranked in the Cf-list based on saliency assigned to them by the speaker. The backward-looking center (Cb) is a special member of the Cf, which represents the discourse entity that the utterance U_i is centrally concerned with. The most salient entity in the Cf-list is called the *preferred center* (Cp), which is considered the best candidate to be realized as the Cb in the utterance U_{i+1}. Centering Theory is also governed by three constrains included in (17). Taking the basics of Centering Theory into account, I can present a hypothetical mini-discourse consisting of at least three utterances in order to show how the external possessors in (4) can be organized in an actual discourse.

(17) Constraints in Centering Theory:
For each utterance U_i in a discourse segment D consisting of utterances $U_1, \ldots, U_m$:
a. There is precisely one backward-looking center Cb (U_i, D).
b. Every element of the looking-forward list, Cf (U_i, D), must be realized in U_i.
c. The center, Cb (U_i, D), is the highest-ranked element of Cf (U_{i-1}, D).

Let us assume that (4), represented as U_i, is the last utterance of a three-fold hypothetical min-discourse. We should also bear in mind that left-dislocation serves a topic-promoting function, which means the new coding of a discourse referent that is already evoked in the preceding discourse or accessible via inference. It implies that the Cb of the utterance U_{i-1} should be different from the Cb of the utterance U_i. In other words, there is a shift of attention from U_{i-1} to U_i; otherwise it would make no sense to use an NP-detached construction. It can be best hypothesized that the referents of the two external possessors do not belong to the Cf-list of U_{i-1} at all, so that the two external possessors can be promoted as new topics in U_i. Likewise, the Cb of the utterance U_{i-1} should be different from the Cp of the utterance $U_{i-n,\ n>1}$ (in our hypothetical mini-discourse, n = 2), which contains the two external possessors whose referents have to be situated within the Cf-list of U_{i-n}, n>1. A reminder is in order here that the second possessor, coded as new topic in U_i, is per se a possessum to the external possessor in U_{i-n}, such that it is ranked lower than the external possessor, i.e. the first external possessor in U_i, in the Cf-list of U_{i-n}.[5]

The hypothetical mini-discourse is represented in Figure 1. The first utterance U_{i-n} does not have a Cb shown as [?], following the Centering Theory notations. The Cbs of the next two utterances are shown as α and β, where both represent two different values, which implies that the Cb of U_{i-1} is different from that of U_i, to provide an appropriate context for the

[5] An external possession construction in Persian bears resemblance to an NP possessive construction in English in that the Cf ordering is from left to right, namely the possessor is ranked higher than the possessum (Walker & Prince 1996: 297) due to a higher degree of saliency.

topic-promotion to operate. Since $U_{i\text{-}n}$ is the first utterance of the mini-discourse, there is no transition which is represented in $U_{i\text{-}n}$ as No Cb. What we can say about the transition states of $U_{i\text{-}1}$ and U_i is that since the Cbs of the three utterances are different, there is at least a 'not continue' transition rule represented as $\neq$ CONTINUE in $U_{i\text{-}1}$ and U_i (see Walker *et al. ibid:* 6 for centering transition states). Put differently, the speaker is inclined to introduce the referents of the two external possessors of U_i in the Cf-list of $U_{i\text{-}n}$ and immediately enforces a shift in the attentional state of the addressee by introducing a discourse-new referent(s) in $U_{i\text{-}1}$ while avoiding the coding of the Cp of $U_{i\text{-}n}$ as the Cb of $U_{i\text{-}1}$. The referents of the external possessors in the end are reintroduced by the speaker to the discourse by putting them in the Cf-list of U_i.

Cb ($U_{i\text{-}n,\ n>1}$:	[?]
Cf:	[BAČČE, PIRHÆN, . . .]
Transition:	No Cb

Constraint 1: Cb ($U_{i\text{-}n,\ n>1}$) $\neq$ Cb ($U_{i\text{-}1}$) = [?]
Constraint 2: Cp ($U_{i\text{-}n,\ n>1}$) $\neq$ Cb ($U_{i\text{-}1}$)

Cb ($U_{i\text{-}1}$):	[α]
Cf:	[α . . .]
Transition:	$\neq$CONTINUE

Constraint 3: Cb ($U_{i\text{-}1}$) $\neq$ Cb (U_i)

Cb ($U_{i\text{-}1}$):	[β]
Cf:	[BAČČE, PIRHÆN, DOGME]
Transition:	$\neq$ CONTINUE

Fig. 1: Centering-based representation of double external possession

We have regarded the external possessors respectively as the primary and secondary topic on the sentence level with this point in mind that the first external possessor is more salient than the second one on the level of discourse. This fact is now visible in Figure 1. The first

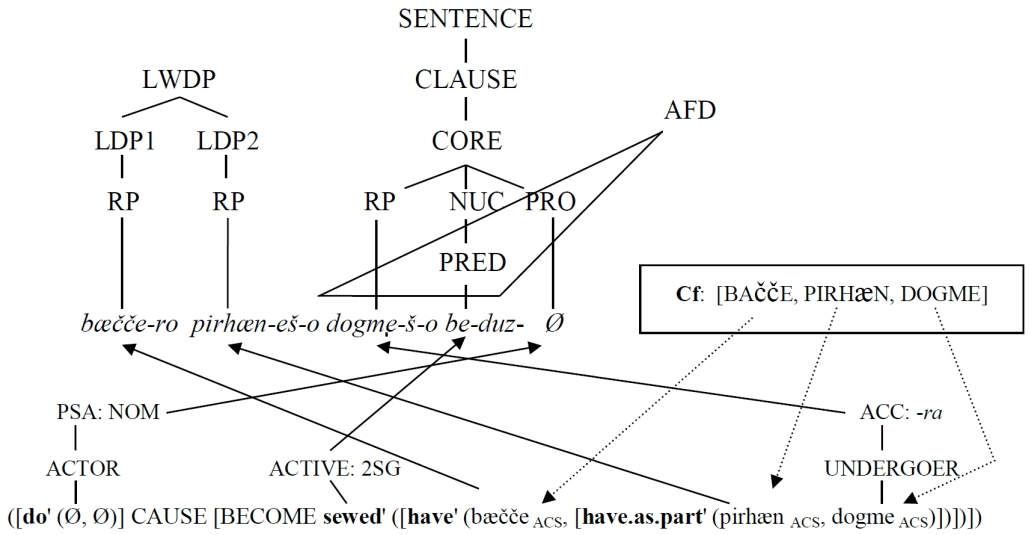


Fig. 2: External possessors in LWDP capturing the notion of discourse saliency

external possessor has been ranked higher than the second one in $U_{i\text{-}n}$, which helps us capture the higher degree of saliency of the former compared to the latter. As a result, we can use the last discourse construct in Figure 1 in order to make a link from the discourse domain to the semantic representation to eliminate the shortcoming in RRG with respect to the failure in representing the relativity of the information status of the external possessors. The linking diagram between the semantics and syntax of the sentence in (4) is given in Figure 2.

4 RRG treatment of contrastive clitic left-dislocation

As already discussed in subsection 2.3, CLD, in addition to topicality, serves a contrastive function in the sense that it operates on a set of alternatives whose referents are explicitly mentioned in the preceding discourse. The features [+exhaustive] and [+contrastive] were employed in line with É. Kiss (1998) to distinguish between the contrastive focus in the light of the above features and the exhaustive focus representing the features [+exhaustive] and [-contrastive]. Following Monlár (2006), I also demonstrated that the contrastive focus is syntactically placed in a position in the left periphery of the core termed as kontrast-position. Furthermore, the notion of kontrast is regarded as an individually pragmatically motivated category of information structure. Putting together the two proposed arguments, we can say that kontrast in Persian can be represented by a contrastive focus employing a CLD construction and monitored by a discourse-pragmatic constraint, which is the explicit mentioning of

the contrasting elements in a closed set. In this subsection, we focus on the RRG projection of a contrastive focus hosted by kontrast-position. To this end, we first need to specify where the kontrast position is situated within the layered structure of the clause in RRG. The PrCS by default is the syntactic position for focal constituents that leave a gap in the core of the clause, and the LDP hosts topical constituents such as CLD-ed topical NPs that mark their original place in the core of the clause by a resumptive pronoun (Van Valin & LaPolla 1997: 228). From an information structure perspective, the PrCS lies within the PFD while the LDP is excluded from it. The complexity regarding the RRG projection of a contrastively focused NP as a result of CLD is that it does not neatly fit into the pragmatically motivated syntactic positions in RRG. On the one hand, the contrastive focus in the kontrast position should be inside the clause to be within PFD, and on the other hand, there is a pronominal clitic in the core of the clause indicating the canonical position of the constituent with a contrastive focus interpretation. It follows that the contrastive focus cannot be in PrCS of the clause because the corresponding gap, contrary to PrCS-positioned constituents, is filled by a pronominal clitic. Similarly, it cannot be in the LDP either, because the LDP-positioned constituents are excluded from the PFD and cannot be focused. Therefore, we should postulate another pragmatically motivated position in the left periphery of the Persian core which is motivated by the discourse-pragmatic concept of kontrast that requires a set of alternatives whose referents are explicitly evoked in the discourse. Based on the different kinds of focus presented in Table 1, I will consider the exhaustive focus to be located in the PrCS and the contrastive focus in a position referred to as the kontrast position (KP), following Monlár's terminology.

Since the RRG projection of a contrastive focus in the KP is closely tied to the prior evocation of a closed set with clearly identifiable members in the discourse context, we need to Figure out how to link the set of contrasting alternatives in the preceding discourse to the information structure of a sentence containing the contrastive focus. This can be done in terms of Discourse Representation Theory (DRT, see Kamp *et al.* 2011 and references therein) brought into RRG by Van Valin (2005). Discourse-pragmatics in RRG is split into the information structure representation of a sentence via the focus structure projection that delineates the pragmatic assertion albeit lacking the representation of the pragmatic presupposition. To get a thorough picture of the pragmatic presupposition as well, it is recommended to exert Discourse Representation Structures (DRSs) of DRT that supply the information structure of a sentence with not only the pragmatic assertion but also with the pragmatic presupposition. As regards the RRG account of the contrastive focus in Persian, I will develop a DRT-based schematized analysis in the same vein according to which the explicit set of the contrasting discourse referents is spelled out by a sequence of DRSs. In (16), we have a question-answer pair in which the alternatives in a closed set, contrasted to satisfy the unspecified value/pragmatic assertion of the variable/pragmatic presupposition, are salient enough in the current attention center of the interlocutors, and the appearance of *ra* on the interrogative

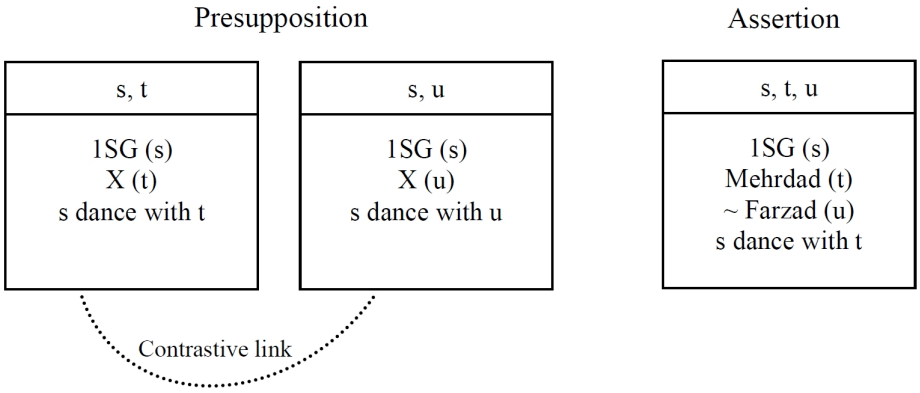


Fig. 3: DRT-based representation of the derivation of a contrastive focus

element *ki* 'whom' signals the identifiability of the referents in question. The DRSs, representing the pragmatic presupposition and assertion in (16), can be sketched out as in Figure 3 displaying how the exclusion of an alternative from a contrastive set of explicitly mentioned discourse referents brings about the derivation of a contrastive focus. The first two DRSs pertain to the pragmatic presupposition, which are connected by means of a contrastive link, following the notation used by Shimojo (2009).[6]

Now I turn to the syntactic and semantic representation of (16b). Let us begin with the logical structure of the verb *dance* including a comitative oblique core argument. Van Valin & LaPolla (1997: 379) and Farrell (2009: 196) suggest the logical structure of a comitative argument to be **do'**([x ∧ y], [**predicate'** ([x ∧ y])]). Therefore, the logical structure of (16b) can be shown as in (18).

(18) **do'**([1SG, Mehrdad], [**dance'**([1SG, Mehrdad])])

A contrastive focus is characterized by the features [+exhaustive] and [+contrastive], signifying that the discourse referent selected as the focus exhaustively identifies the only value from among an explicit set of referents in the preceding discourse satisfying the variable in the

[6] Shimojo makes a distinction between two types of referent specification in the light of different morphological markers in Japanese. The first type deals with the contrastive specification represented by a *wa*-marked argument that requires a proper contrastive link in the discourse representation structures of two discourse referents standing in an explicit contrastive relationship to one another. The second type deals with the absolutive specification represented by a zero-marked argument signaling that no search of a possible contrastive relationship is required for the exclusive identification of a discourse referent matching up with the proposition. By analogy, the contrastive specification and absolutive specification in Shimojo's approach are similar to the contrastive focus and exhaustive focus in Persian with the difference that Persian employs different syntactic means instead of morphological markers in order to differentiate between the two. Preposing and clitic left-dislocation (see (15b') and (16b)), placing their relevant type of focus in the KP and PrCS, are used to denote the contrastive and absolutive specification respectively.

pragmatic presupposition. The logical structure in (18) then is not congruent with the exhaustive identification as an abstract semantic operator and the notion of kontrast as a discourse-pragmatic notion. I assume that the logical structure underlying the sentence in (16b) should be a specificational predicate, which enables us to subsume the semantically oriented concept of exhaustive identification into the logical structure of a sentence with an exhaustive or contrastive focus, as both have the feature [+exhaustive] in common. The logical structure of a specificational predicate, after Van Valin (2005: 48), is **be'**(x, y); x and y represent the variable and the value respectively. As a consequence of this, the logical structure of (16b) can be as in (19).

(19) **be'**([**do'**([1SG ∧ x_i])], [**dance'**([1SG ∧ x_i])])], $Mehrdad_i$)

The first argument of the specificational predicate, i.e. the variable is filled by the logical structure of the verb *dance* whose internal logical structure contains the unspecified comitative NP represented by the lower case x. The second argument position is filled by the focus value, which is only compatible with the feature [+exhaustive]. The exhaustive identification is then captured by the co-indexation of the unfilled argument position in the internal structure of the verb *dance* with the second argument of the specificational predicate **be'**. The logical structure in (19) complies entirely with the sentence in (15b') and not with the one in (16b), because the feature [+contrastive], which has to do with the explicitness of the contrasting discourse referents, has been sidelined and needs to be inserted somehow into the logical structure of (19).

Another point in this regard deals with the function of the clitic in the contrastive CLD, which is different from topical CLD. In the latter, the clitic plays the role of an argument on behalf of the topical left-dislocated constituent (in the case of left-dislocated obliques only), whereas in the former, it already appears on the question (see (16a)) to enable the addressee to track the referents to which the clitic co-refers in the discourse. In other words, the pronominal clitic has a set-evoking function for the members whose referents have already been mentioned explicitly in the discourse; however, which member satisfies the variable expressed by the pragmatic presupposition in the question is determined by the addressee's communicative act provided by the reply in (16b). As a whole, the logical structure of (16b) can be represented as in (20), which not only adheres to the exhaustive identification, i.e. [+exhaustive], but also includes the explicitness condition, i.e. [+contrastive]. Moreover, the clitic has to be represented in the logical structure for satisfying the completeness constraint as it is assigned a single syntactic position in the core of the clause. The pronominal clitic has been represented in (20) by the lower case y co-indexed with the unspecified comitative x. The existence of the unlinked set members t and u runs counter to the completeness constraint, in which case since they are accessible from the discourse context, they are linked directly from

the last DRS to the logical structure to satisfy the completeness constraint, as can be seen in Figure 4.

(20) **be'**([**do'**([1SG ∧ X_i)], [**dance'**([1SG ∧ X_i [Y_i {t, u}]])])], $Mehrdad_i$)

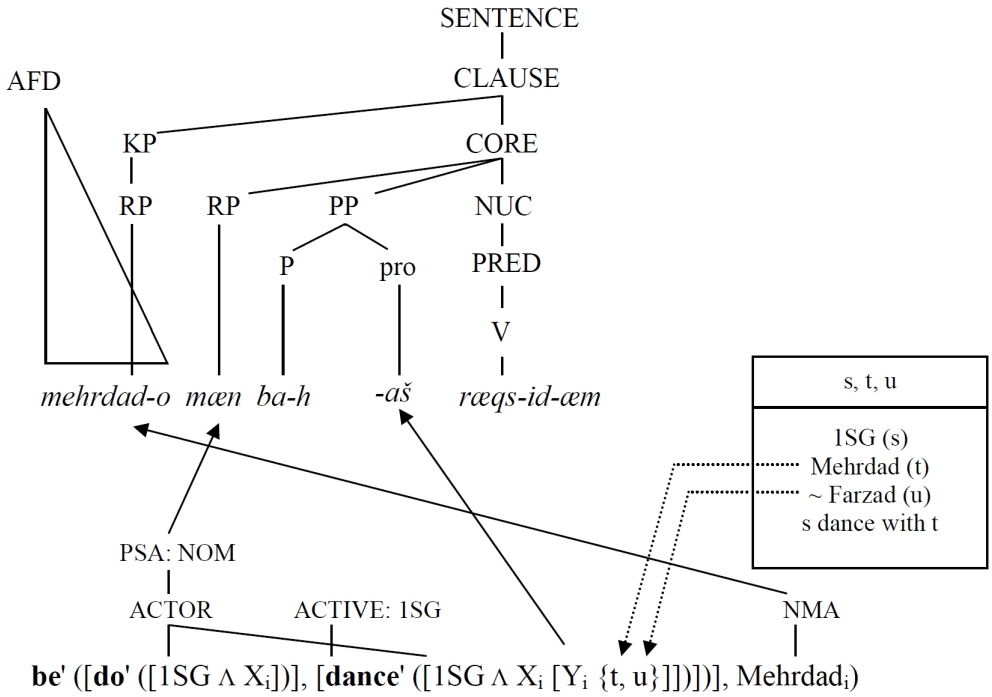


Fig. 4: Contrastive focus KP capturing the explicitness condition

5 Conclusion

The case of double external possession inevitably proves that RRG adopts a static view on the notion of topicality by postulating a single LDP in the syntactic structure of the sentence. The replacement of the LDP by a LWDP and the integration of a Centering model into RRG are assumed to be an influential resolution for illustrating the topicality gradience of left-dislocated NPs, where topicality is defined with respect to the degree of discourse saliency. Moreover, a Centering approach to the RRG projection of double external possessors helps us schematically capture the relativity of information/cognitive status of discourse referents in cases where the insertion of an absolute information status of discourse referents in the logical structure does not genuinely mirror the dynamicity of the information flow in the topical elements. Similarly, the postulation of the KP for a contrastive focus, selected as an exhaustive value from among an explicit set of contrasting alternatives, is also considered to

be a contribution to pragmatically motivated aspects of the layered structure of the clause in RRG, which can be tested on a spectrum of cross-linguistic data. Therefore, from an RRG perspective the left periphery of the Persian clause structure gets extended by assuming the four pragmatically motivated syntactic positions including the PrCS, KP, LDP_1 and LDP_2, allocated respectively to an exhaustive focus, a contrastive focus, a primary and secondary topic as in the case of two external possessors.

References

Dabir-Moghaddam, Mohammad. 1992. On the (in)dependence of syntax and pragmatics: Evidence from the postposition *-ra* in Persian. In *Cooperating with Written Texts: The Pragmatics and Comprehension of Written Texts* [Studies in Anthropological Linguistics 5], D. Stein (ed.), 549–573. Berlin: de Gruyter.

—. 2005. *Studies in Persian Linguistics: Selected Articles*. Markaz-e Nashr: Tehran.

Dalrymple, Mary & Nikolaeva, Irina. 2011. *Objects and Information Structure*. Cambridge: CUP.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74(2): 245–274.

Farrell, Patrick. 2009. The preposition with in Role and Reference Grammar. In *Studies in Role and Reference Grammar*, L. Guerrero, S. Ibáñez Cerda & V. A. Belloro (eds), 179–203. Universidad Nacional Autónoma de México.

Frascarelli, Mara & Hinterhölzl, Roland. 2007. Types of Topics in German and Italian. In *On Information Structure, Meaning and Form* [Linguistics Today 100], S. Winkler & K. Schwabe (eds), 87–116. Amsterdam: John Benjamins.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In *Topic Continuity in Discourse: A Quantitative Cross-language Study* [Typological Studies in Language 3], T. Givón (ed.), 1–43. Amsterdam: John Benjamins.

Gregory, Michelle & Michaelis, Laura. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33(11): 1665–1706.

Gundel, Jeannette. 1988. Universals of topic-comment structure. In *Studies in Syntactic Typology* [Typological Studies in Language 17], M. Hammond, E. A. Moravcsik & J. Wirth (eds), 209–239. Amsterdam: John Benjamins.

Gundel, Jeanette K. 1999. On different kinds of focus. In *Focus: Linguistic, Cognitive, and Computational Perspectives*. Peter Bosch & Rob van der Sandt (eds), 293–305. Cambridge: CUP.

Hopper, Paul & Traugott, Elizabeth C., 2003. *Grammaticalization*. CUP.

Kamp, Hans, Van Genabith, Josef & Reyle, Uwe. 2011. Discourse representation theory. In *Handbook of Philosophical Logic*, D. M. Gabbay & F. Guenthner (eds), 125–395. 2*nd* ed. Dordrecht/Heidelberg: Springer.

Kockelman, Paul. 2009. Inalienable possession as grammatical category and discourse pattern. *Studies in Language* 33(1): 25–69.

Lambrecht, Knud. 1981. *Topic, Antitopic and Verb Agreement in Non-Standard French* [Pragmatics and Beyond New Series 11: 6], Amsterdam: John Benjamins.

—. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: CUP.

—. 2001. Dislocation. In *Language Universals and Language Typology: An International Handbook*, M. Haspelmath, E. König, W. Osterreicher & W. Raible (eds), 1050–1078. Berlin: de Gruyter.

Langacker, Ronald. 1993. Reference-point constructions. *Cognitive Linguistics* 4(1): 1–38.

McCoy, Svetlana. 2003. Connecting information structure and discourse structure through 'kontrast': The case of colloquial Russian particles – TO, ŽE, and VED. *Journal of Logic, Language and Information* 12(3): 319–335.

Monlár, Valeria. 2002. Contrast-from a contrastive perspective. In *Information Structure in a Cross-linguistic Perspective*, H. Hasselgård, S. Johansson, B. Behrens & C. Fabricius-Hansen (eds), 147–162. Amsterdam: Rodopi.

—. 2006. On different kinds of contrast. In *Architecture of Focus* [Studies in Generative Grammar 82], V. Molnár & S. Winkler (eds), 197–233. Berlin: de Gruyter.

Neeleman, Ad, Titov, Elena, van de Koot, Hans & Vermeulen, Reiko. 2009. A syntactic typology of topic, focus and contrast. In *Alternatives to Cartography* [Studies in Generative Grammar 100], J. van Craenenbroeck (ed.), 15–52. Berlin: de Gruyter.

Neeleman, Ad & Vermeulen, Reiko. 2012. The syntactic expression of information structure. In *The Syntax of Topic, Focus, and Contrast: An Interface-based Approach* [Studies in Generative Grammar 113], A. Neeleman & R. Vermeulen (eds), 1–39. Berlin: de Gruyter.

Nikolaeva, Irina. 2001. Secondary topic as a relation in information structure. *Linguistics* 39(1): 1–49.

Payne, Doris L. & Barshi, Immanuel. 1999. External possession: what, where, how, and why. In *External Possession* [Typological Studies in Language 39], D.L. Payne & I. Barshi (eds), 3–33. Amsterdam/Philadelphia: John Benjamins.

Reinhart, Tanya. 1982. Pragmatics and linguistics: An analysis of sentence topic. *Philosophica* 27(1): 53–94.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar: Handbook in Generative Syntax*, L. Haegeman (ed.), 281–337. Dordrecht: Kluwer.

Shimojo, Mitsuaki. 2009. Focus structure and beyond: Discourse-pragmatics in RRG. In *Studies in Role and Reference Grammar*, L. Guerrero, S. Ibáñez Cerda & V. A. Belloro (eds), 75–95. Universidad Nacional Autónoma de México.

Taylor, John R. 1996. *Possessives in English*. Oxford: OUP.

Umbach, Carla. 2004. On the notion of contrast in information structure and discourse structure. *Journal of Semantics*, 21(2): 1–21.

Vallduví, Eric & Vilkuna, Maria. 1998. On rheme and kontrast. In *The Limits of Syntax* [Syntax and Semantics 29], P. Culicover & L. Mc Nally (eds), 79–108. New York: Academic Press.

Van Valin, Robert D., Jr. & LaPolla, Randy. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: CUP.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Walker, Marilyn & Prince, Ellen. 1996. A bilateral approach to givenness: a hearer-status algorithm and a centering algorithm. In *Reference and Referent Accessibility* [Pragmatics and Beyond New Series 38], T. Fretheim & J. K. Gundel (eds), 291–306. Amsterdam: John Benjamins.

Walker, Marilyn, Joshi, Aravind & Prince, Ellen. 1998. Centering in naturally occurring discourse: An overview. In *Centering Theory in Discourse*, M. A. Walker, A. Joshi, E. Prince (eds), 1–29. Oxford: OUP.

The Encoding of Negation in Modern Irish: Negation at the Layered Structure of the Clause and Noun Phrase

Brian Nolan
Institute of Technology Blanchardstown Dublin, Ireland

Abstract

Modern Irish utilises a variety of ways to express and encode negation on the clause and noun phrase. Additionally, negation can be realised not just over the lexical verb within the clause, but also on the auxiliary verb and copula. Each of these realisations of negation has scope considerations and morphosyntactic consequences that are interesting within an RRG account. In particular, negation within the layered structure of the clause (LSC) is sensitive to tense and a past vs. non-past distinction seems to be significant. This may be due to a realis/irrealis distinction. In this paper therefore we explore and characterise the morphosyntactic realisation of negation within the clause, over auxiliary verb constructions and copula clauses, and also as it is found within the layered structure of the noun phrase (LSNP). The general form of encoding negative polarity in a clause involves a negative particle plus a matrix verb (allowing for past or non-past tense distinction), or a negative form of the copula. In many instances, the tense and NEG operators in the LSC are directly encoded in the AUX and marked accordingly in morphosyntax. We discuss the scope of negation within the LSC and the LSNP. We outline morphological and analytical negation and the range of negative polarity particles that are deployed within the construction to encode negation along with the functional differences of each within an RRG characterisation. Within RRG, negation, tense and realis/irrealis are encoded as operators on the clause. Negation is also encoded on the LSNP as an operator.

1 Introduction

Negation is a subject of considerable interest cross linguistically (Horn 2011; Miestamo 2005 and 2010). This paper is concerned with the way in which Modern Irish utilises a variety of ways to express and encode negation on the clause, noun phrase and on the word whether this is a nominal or verb (lexical, auxiliary or copula). We therefore examine the encoding of negation in the layered structure of the clause (LSC), the layered structure of the noun

phrase (LSNP) and the layered structure of the word (LSW) in so far as it occurs in Irish. Our theoretical paradigm of choice is Role and Reference Grammar [RRG] as it is a theory of grammar that characterises the interaction of morphosyntax, semantics and pragmatics in grammatical systems.

RRG is a monostratal theory that posits a single syntactic representation for a sentence, linked directly to a semantic representation by means of a bi-directional linking algorithm. The syntactic representation of clause structure in RRG is called the layered structure of the clause (LSC). Modifying each of the layers of the clause are operators and these include grammatical categories like tense, aspect, modality, negation and illocutionary force. A clause consists of the core and the periphery, that is, the predicate, its arguments and the non-arguments. Clauses are considered to have a layered structure and each layer is semantically motivated. These operators all add additional specification to the event, state, process or action (*Aktionsarten*/event structure) of the predication indicated by the verb. The various scope levels of the operators are indicated as they are conceptualised in the RRG model in Figure 1.

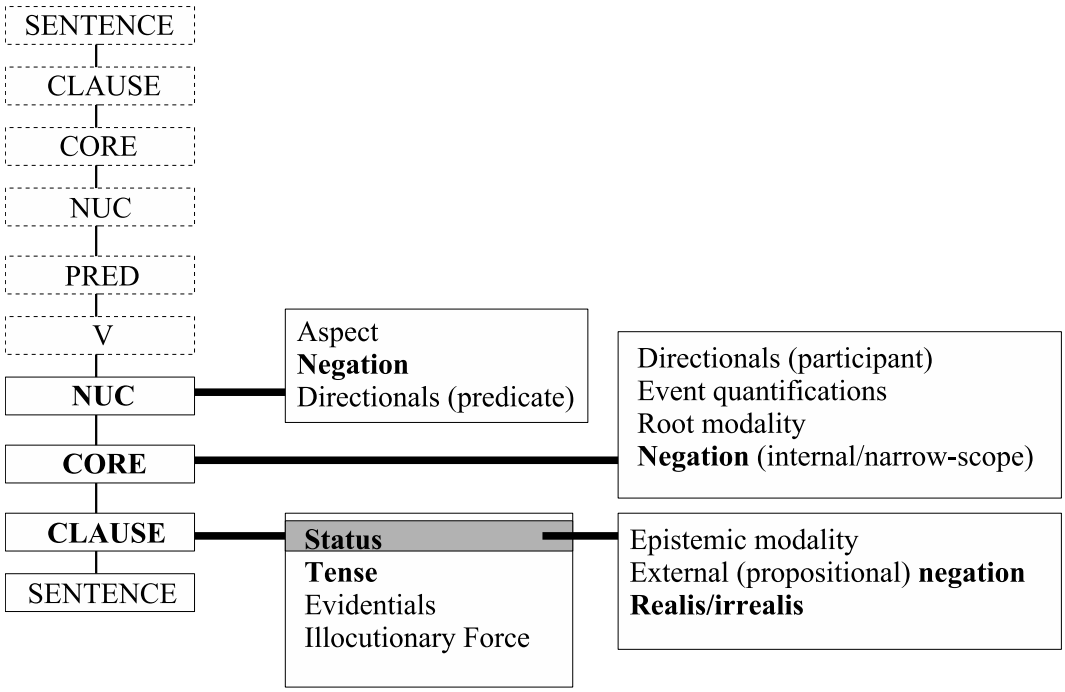


Fig. 1: The operator projection within the different levels of the layered structure of the clause (adapted from Van Valin 2005: 12)

Negation is the only operator that occurs at all three levels (Van Valin 2005: 9). Nuclear negation has only the nucleus in its scope, core negation has one or more core arguments (and possibly also the nucleus) in its scope, and clausal negation has the entire clause in its scope. Operators in the NP follow the same iconic ordering constraint as operators within the clause.

Adjectives are constituents of the nuclear_N periphery whose position is constrained by the iconicity principle and, as such, they must occur closer to the nominal nucleus than core_N and NP-level operators and modifiers.

Languages normally do not have all of these operators as grammatical categories; the ones considered to be 'universal' are illocutionary force and negation. Aspect is a common inherent verbal category. Its function is to highlight the internal temporal unfolding of the predication. Aspect indicates whether an event, state, process or action denoted by a verb is completed or unfolding. The 'perfective'/'imperfective' distinction is related to aspect. Tense indicates the time of the predication in relation to some particular moment and, in the morphology of languages cross linguistically, at least three tenses of past, present and future are usually distinguished. As operators, tense and status situate the proposition expressed by the clause within a temporal and realis/irrealis context. Typologically, we expect that the order of certain affixes in a word would reflect the order of related syntactic operations (cf. The Mirror Principle of Baker 1985: 375). That is, we would expect that the categories more relevant to the verb would occur closer to the stem than those that are less relevant. We do see some evidence of this in the encoding of negation. We can observe this with both the layered structure of the clause and the noun phrase with the respective operator projections.

In the layered structure of the noun phrase (Van Valin 2005: 24), there is a nominal nucleus and a nominal core consisting of the nucleus itself and associated arguments. The NP level corresponds to the clause/sentence levels of the layered structure of the clause. Each of these has a periphery. Nominal aspect concerns whether the referent entity is an individual, parts of an individual, a set of individuals, or a sortal kind. Core_N operators are to do with quality with respect to number, qualification and negation while the NP operators relate to definiteness and deixis, locating the NP in a discourse situation. The operators for the NP are indicated in Figure 2. The formulation of the layered structure of the noun phrase within RRG (Van Valin 2005: 24, see also Van Valin & LaPolla 1997) owes much to the work of Rijkhoff (2002).

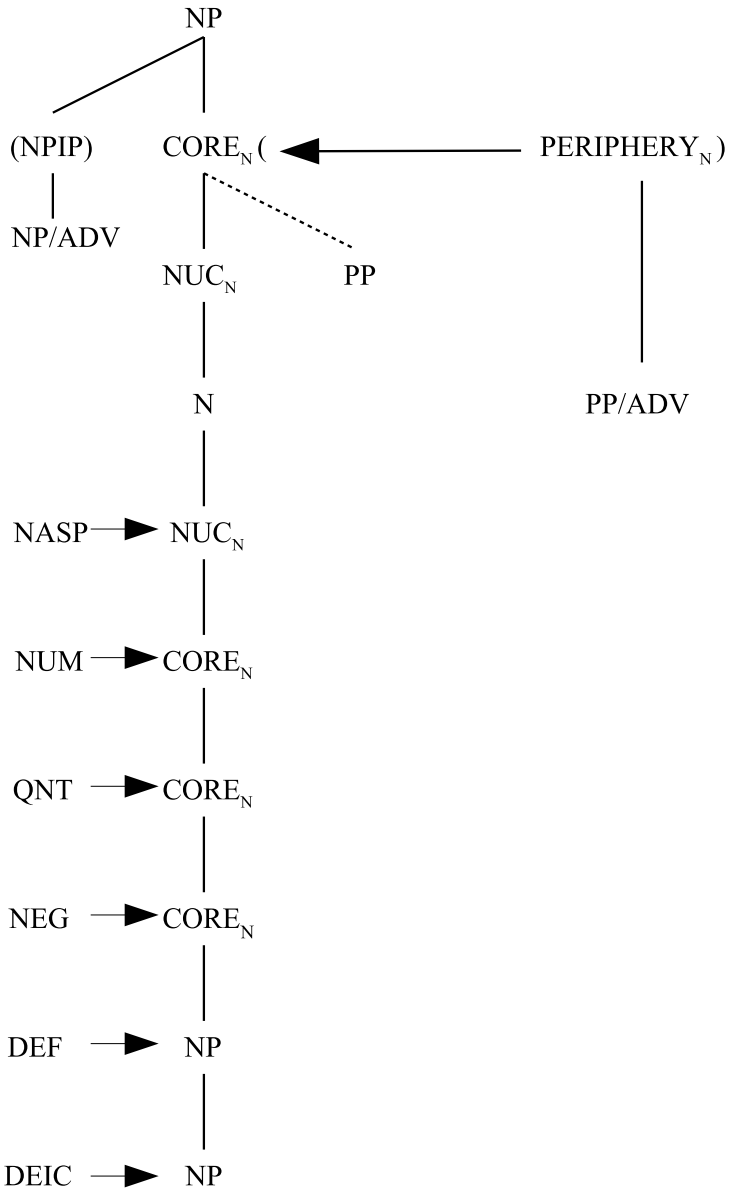


Fig. 2: The general schema for a layered structure of the noun phrase (from Van Valin 2005)

2 Expressing negative polarity on the Irish clause

A negative clause is one which makes an assertion that some event, situation or state of affairs does not hold. Irish has several interesting ways of expressing negative polarity in the clause, but does not have a single word that means 'yes' or 'no' while having strategies that arrive at the same result whether positive or negative. In particular, the language makes a distinction on the forms used according to whether a clause contains a lexical matrix verb, the auxiliary form of the verb 'to be', or is a copula clause. Within a clause, a further distinction is made based on whether the verb encodes past tense or not. We find that Modern Irish uses lexical, morphological, and analytical means to encode negation. At the level of the lexical verb, Modern Irish can employ morphological prefixes, as in examples (1)–(2), to encode negation on the verb within the layered structure of the word (LSW). This encodes negation on the verb nucleus and represents an instance of morphological negation. In contrast, lexical negation describes a situation whereby a negative concept is embedded in the lexical semantics of a verb (3).

(1) **dí**: NEG + lexical V → NEG.V [morphological negation]
díscaoil V 'unloose'
díphacáil V 'unpack'
díphlugáil V 'unplug'
díghreamaigh V 'unstick'
díliostáil V 'unsubscribe'
dícheangail V 'untie'
díchornaigh V 'unwind'
díthocraisigh V 'unwind'
díscoir V 'unyoke'

(2) **do**: NEG + LEXICAL ITEM$_{TYPE}$ → 'un-/non-'V$_{TYPE}$ [morphological negation]
doicheallaigh V 'unwilling/balk' [vs. cealaigh V 'cancel', 'rescind, 'remove']

(3) Lexical negation
a. *fásaigh* v. 'empty/lay waste/desert'
b. *folmhaigh* v. 'empty/exhaust/discharge/purge/evacuate/devastate'
c. *bánaigh* v. 'empty/clear out/devastate/lay waste'

The general form of encoding negative polarity in a clause involves analytical negation consisting of a negative particle plus a *matrix* verb (allowing for past or non-past tense distinction), a negative form of the *auxiliary*, or a negative form of the *copula*. Different forms too of a negative particle occur with an *imperative* clause. Additionally, negation is morphologically

supported in *question* and *conditional* forms of a clause. The particles that are used to encode negative polarity include (4), for the verb and clause (see also Figure 3).

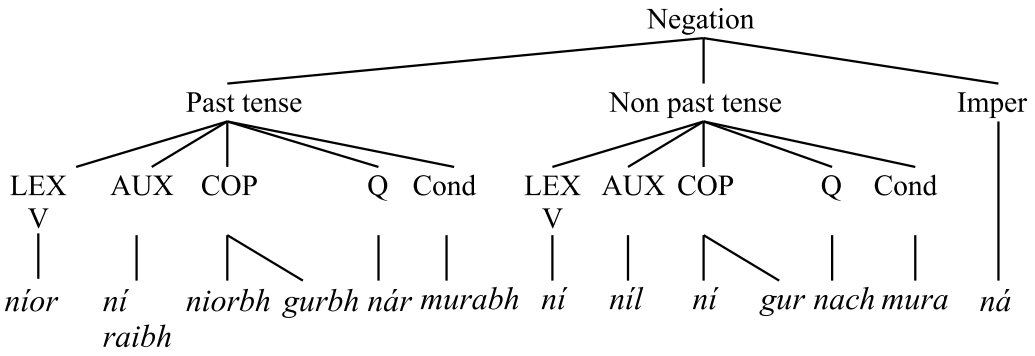


Fig. 3: The taxonomy of negation in Modern Irish

We will see that the encoding of negation on the clause is symmetric in that it always pivots on the past/non-past distinction. By symmetric we mean that the clause structure is the same for the encoding of (positive and) negative polarity with the exception of the form of the negative particle (5).

(4) The negative polarity particles on the clause in analytical negation

Clause	The clausal environment of negation
Ní 'not'	1. Form of NEG used with all *non-past* tense *matrix* verbs
Níor 'not'	2. Form of NEG used with past tense *matrix* verbs
Ní raibh 'not' AUX	3. Negative form of past tense AUX
Níl 'not' AUX	4. Negative form of *non-past* tense AUX
Nach 'not'	5. Form used in Qs with *non-past* tense matrix verbs
Nár 'not'	6. Form used in Qs with past tense matrix verbs
Nár 'not'	7. Form used in *present subjunctive* matrix verbs
Mura 'if..not'	8. Used in *conditional* clauses with *non-past* tense matrix verbs
Murar 'if..not'	9. Used in *conditional* clauses with past tense *matrix* verbs
Ná 'not'	10. Form of NEG used only in *imperative* mood

(5) Example of negative encoding on verb *éirigh* 'rise'

a. *Positive encoding*

Past Tense	*éirigh sé*	he rose.

b. *Negative encoding*

Past Tense	*níor éirigh sé*	he didn't rise.
Present Tense	*ní éiríonn sé*	he won't rise.
Future Tense	*ní éireoidh sé*	he will not rise.
Past Habitual	*ní éiríodh sé*	he didn't rise.
Conditional Mood	*ní éireodh sé*	he used to not rise/he didn't use to rise.
Imperative	*ná héirigh*	Don't not rise!/Don't rise!
Present Subjunctive	*nár éirí sé*	. . . he may not rise.

3 Negative polarity in clauses with matrix verbs

In this section, we explore the form of the negative used with matrix verbs. We find that a symmetry exists that is based on the past/non-past distinction.

3.1 *Nior*-Form of NEG used with past tense matrix verbs

In examples (6)–(8), we can see that the negative particle *níor* is used to flag negation on the clause. The matrix verb is in the past tense in these examples.

(6) *Níor chreideas é.*
NEG believe.V-PST+1SG 3SG.M.ACC
'I did not believe it/him.'

(7) *Níor tháinig Cáit fós.*
NEG arrive.V-PST Kate.N yet.ADV
'Kate did not arrive yet.'

(8) *Níor cheannaig mé an t-asal*
NEG bought.V-PST 1SG.NOM the.DET donkey.N
'I did not buy the small black donkey.'
NEG'[do'(I, [**buy'**(I, the donkey)])]

Example (9) is especially interesting in that it contains a negative encoding on the clause and, additionally, a noun with the lexicalised meaning of 'nothing'. The effect of both these together creates an ambiguity that intersects with quantification, in particular, the existential quantifier.

(9) *Níor dhúirt Patsy faic.*
NEG say.V-PST Patsy.N nothing.V
LIT: 'Patsy did not say nothing'
'Patsy did not say anything.'

3.2 *Ní*-Form of NEG used with all non-past tense matrix verbs

In examples (10)–(11), we can see that the negative particle *ní* is used to flag negation on the clause. The matrix verb is in non-past tenses in these examples.

(10) *Ach ag stáisiún na Carraige Duibhe, ní bhogfadh sí*
PART at:PREP station.N DET Blackrock.N, NEG move.V-FUT 3SG.F.NOM
as a suíochán.
from:PREP ADJ.POSS.F seat.N
'But, at Blackrock station, she would not move from her seat.'

We see in (11) the intersection of negation at clause level and nominal level with quantification. In English, we must represent <NEG (anyone)> as 'no one'. That is, as <NEG (N)>.

(11) *Ní éireodh éinne, fear nó páiste.*
NEG rise.V-PRS anyone.N man.N or child.N
LIT: 'anyone did not rise, man or child'.
'No one rose, man or child.'

The Irish nominal *einne* 'anyone' represents a generic existential quantifier and, typical of languages generally, this interacts with negation as <NEG (N)>.

3.3 Negation used in clauses with auxiliary verbs

In auxiliary verb clauses, we also find that a symmetry exists across the past/non-past distinction (Table 1). We show predicative, attributive and the *níl ach* 'only' construction, with past tense in (13) to (14) and present tense in (15) to (18). The predicative uses shown are also present tense and so, as expected, have *níl* encoding negation.

3.3.1 Attributive use of auxiliary verbs and negation

Clause negation and constituent negation interact in interesting and complex ways. In Modern Irish, the clause form based on the auxiliary verb construction is used to analytically negate a constituent. That is, the constructional form of the auxiliary clause for attribution is also used as a device for inner constituent negation. We can see this in the bolded constituents in the example translations (13)–(18).

a.	AUX Verb of 'to be' present tense + ADV|ADJ|PREP + VN clause:	*Tá mé ...*
b.	Negative AUX of 'to be' present tense + ADV|ADJ|PREP + VN clause:	*Níl mé ...*
c.	AUX Verb of 'to be' past tense + ADV|ADJ|PREP + VN clause:	*Bhí mé ...*
d.	Negative AUX of 'to be' past tense + ADV|ADJ|PREP + VN clause:	*Ní raibh mé ...*

Table 1: The syntactic pattern of the Irish AUX 'to be' over past/present tenses

(12) < TNS {PST|PRS} < NEG **BE'**(LS) > >

(13) *Ní raibh Patsy sa chistin, ach bhí an raidió ar siúl.*
NEG BE:AUX-PST Patsy.N in:PREP+DET kitchen.N PART BE:AUX-PST DET radio on:PREP playing:VN
'Patsy was **not in the kitchen** but the radio was on.'
<PST< NEG (**be-at'**(Patsy, kitchen)) > >

(14) *Ní raibh an roth ag bogadh.*
NEG BE:AUX-PST DET wheel.N at:PREP moving:VN
'The wheel was **not turning**.'
<PST< NEG (**be-at'**(the wheel, [**moving'**]) > >

(15) *Níl an madadh ag ithe.*
NEG-BE:AUX-PRS DET dog.N at:PREP eating :VN
'The dog is**n't eating**/The dog is **not eating**.'
<PRS< NEG (**be-at'**(the dog, [**eating'**]) > >

(16) *Níl arigead ag muintir Pháidín.*
NEG-BE:AUX-PRS money.N at:PREP family.N Paidín.N
'Paidín's family do**n't have money**.'
<PRS< NEG (**be-at'**(money, Paidín's family)) > >

(17) *Níl bean ag an doras.*
NEG-BE:AUX-PRS woman.N at:PREP the:DET door.N
'A woman is **not at the door**. [...it is a man, etc.]'
<PRS<NEG'[**be-at'**(door, woman)] > >

(18) *Níl lá deas ann.*
NEG-BE:AUX-PRS day.N nice:ADJ there:DEM
'The day is **not nice**.'
<PRS<NEG'[**be'**(day, [**nice'**])] > >

3.3.2 Inner negation on embedded auxiliary clause

In (19), we have an embedded auxiliary clause that encodes negation on the constituent. Sometimes negation on the constituent is used to 'fine-tune' negation scope within the clause. Here, we see use of an embedded auxiliary clause construction used to deliver negative scope over the constituents.

(19) *Deir siad nach bhfuil Bríd ná Máirtín anseo.*
Say.V-PST 3PL.NOM NEG BE:AUX-PRS Bríd.N or Máirtín.N here.ADV
LIT: 'They said that Bríd or Máirtín are not here'.
'They said that neither Bríd or Máirtín are here.'
<PST<[say'(x, [<PRS<**NEG** [**be-at'**(Bríd or Máirtín, [**here'**])]] > >

In this example, the translation captures the inner negation within the embedded AUX clause.

3.3.3 Predicative use of auxiliary verbs and negation

Negation is also found with predicative use of auxiliary constructions. We represent two different ways of articulating a predicative construction, all showing negative polarity, with the forms symmetric over the past/non-past distinction.

(20) *Ní* ***raibh*** *tuirseach orm.*
NEG BE:AUX-PST tired:ADJ on:PREP+1SG
LIT: 'Tiredness is not on me'
'I am not tired.'
<PST< NEG (**feel'**(I, [**tired'**])) > >

a. *Ní raibh mé tuirseach.*
NEG BE:AUX-PST 1SG tired:ADJ
'I am not tired.'
<PST< NEG (**feel'**(I, [**tired'**])) > >

b. *Níl tuirseach orm.*
NEG-BE:AUX-PRS tired:ADJ on:PREP+1SG
LIT: 'Tiredness is not on me'
'I am not tired.'
<PRS< NEG (**feel'**(I, [**tired'**])) > >

c. *Níl mé tuirseach.*
NEG-BE:AUX-PRS 1SG tired:ADJ
'I am not tired.'
<PRS< NEG (**feel'**(I, [**tired'**])) > >

3.3.4 The níl . . . ach 'only' construction – 'negation . . . but' → 'only'

The negative form of the auxiliary níl, when combined with ach 'but', is used to express the notion 'only' in a NÍL . . . ACH construction. The form used also pivots on the past/non-past distinction and this construction indicates the same symmetry as other clauses for negative encoding.

(21) *Ní raibh ach leabhar amháin agam.*
NEG BE:AUX-PST BUT book.N one:QTY at:PREP+1SG
'I had **only** one book.'

(22) *Níl ach leabhar amháin agam.*
NEG-BE:AUX-PRS BUT book.N one:QTY at:PREP+1SG
'I have **only** one book.'

This also reflects another instance where, at constructional level, the negative auxiliary intersects with qualification to reinforce a particular reading, here encoding the sense of 'only'.

3.4 Negation and the copula

The copula verb types of 'to be' with their schemata are indicated in a generalised form, along with the negative copula forms, in Table 2. Syntactically, while the AUX verb of 'to be' verb takes all tenses, the copula is impoverished in inflection and is never inflected for person. The copula is held to predicate for essential or inherent qualities and generally takes NP predicates. The copula has only two forms for tense: *is* 'be' is the positive polarity unmarked form used for present and future while *ba* 'be' is used for positive polarity past and conditional meanings.

The copula has negative forms: *ní* 'NOT be' in the present, *níorbh* in the past. Embedded negative copulas and negative copula questions take the respective present and past forms: *nach*, *nár(bh)*. We will see examples of these in the next section when we examine negation in relation to questions. The copula may never stand alone in a response to a question, irrespective of whether the polarity of the response is positive or negative. Typically, the copula will co-occur with another constituent, for example, an adjective or pronoun.

(23) *Ní <u>breá</u> an lá é.*
NEG.COP.PRS fine:ADJ DET day.N 3SG.M.ACC
'It is not a fine day.'
NEG'[**fine'**(day)]

(24) *Ní <u>deas</u> an lá é.*
NEG.COP.PRS nice:ADJ the:DET day.N 3SG.M.ACC
'It is not a nice day.'
NEG'[**nice'**(day)]

a.	Copula-past tense:	*is* ...
b.	Copula-past tense + VN clause:	*Ba*$_{\text{cop}}$ + ADJ|N + VN clause
c.	Negative copula present tense + VN clause:	***Ní*** ... + ADJ|N + VN clause
d.	Negative copula past tense + VN clause:	***Níorbh*** ... + ADJ|N + VN clause
e.	Copula-past tense + finite clause:	*Ba*$_{\text{cop}}$+ *féidir* + finite complement
f.	Negative copula present tense + finite complement clause:	***Ní*** ... + finite complement
g.	Negative copula past tense + finite complement clause:	***Níorbh*** ... + finite complement
h.	Copula within a complement clause, present tense:	... *gur* ...
i.	Copula within a complement clause, past tense:	... *gurbh* ...

Table 2: The syntactic pattern of the Irish Copula 'to be' clause

3.5 The negative forms used in questions

We now examine the negative forms used in the various kinds of question clauses, involving matrix, auxiliary and question clauses, respectively.

3.5.1 Question clauses

A question clause may contain an auxiliary verb, or a lexical verb. There are particular characteristics to each which are important to the clause forms of Irish.

3.5.1.1 Questions with a matrix verb

We look first at questions with a lexical verb. As we see, these are also symmetric and sensitive to the past/non-past distinction.

(25) Past

Nár *scríobh* *sí* *leitir* *agat?*
NEG.Q write.V-PST 3SG.F letter.N at:PREP+2SG
'Did she not write the letter for you?'

(26) Non-past

Nach éisteann sé?
NEG.Q listen.V-PRS 3SG.
'Does he not listen?'/'Doesn't he listen?'

3.5.1.2 Questions requiring yes/no response and verb echo forms

In the case of questions with lexical verbs, we should also note that when a question is asked that requires a yes/no response, it is necessary for grammatical acceptability that the matrix verb is repeated in an echo form of the clause along with the appropriate polarity. If the polarity is negative, then a negative particle is also required. There is no clausal form that, as a response to a yes/no question, does not require the echo form with the repeat of the matrix verb and, consequently, the matrix verb is repeated with negative polarity encoded by the presence of a negative particle. These are called echo forms and are used for encoding the appropriate positive or negative polarity. These echo forms are used across all tenses for lexical verbs, the AUX verb of 'to be' and the copula.

(27) a. *Ar cheannaigh sé é?*
Q buy.V-PST 3SG.M 3SG.M.ACC
'Did he buy it?'

b. *Cheannaigh!*
Buy.V-PST
'(He) bought (it)!' = 'Yes!'

c. *Níor cheannaigh!*
NEG buy.V-PST
'(He) did not buy (it)!' = 'No!'

(28) a. *An dtiochfá liom?*
Q come.V-FUT+2SG with:PREP+1SG
'Will you come with me?'

b. *Thiocfainn!*
Come.V-FUT+1SG
'I will come (with you)!' = 'Yes!'

c. *Ní thiocfainn!*
NEG come.V-FUT+1SG
'I will not come (with you)!' = 'No!'

3.5.1.3 Question clause with NEG in auxiliary forms

If the *question* clause employs the *auxiliary* verb then the appropriate negative (or positive) form of the AUX must occur in the response to a question. That is, the tense and NEG operators in the LSC are directly encoded in the AUX.

(29) a. *Nach bhfuil sé ann?*
NEG.Q BE.AUX-PRS 3SG.M there:DEM
'Isn't he there?'

b. *Tá!*
BE.AUX-PRS
LIT: 'Be!'
= 'Yes!' **[Positive]**

c. *Níl!*
NEG.be.AUX-PRS
LIT: 'Not be!'
= 'No!' **[Negative]**

(30) *Nach bhfuil múinteoir ansin?*
NEG.Q BE:AUX-PRS teacher.N there.ADV
'Is there **no teacher** there?'

(31) *Nach bhfuil an fhuinneog briste?*
NEG.Q BE:AUX-PRS DET window.N broken.ADJ
'Isn't the window broken?'/'Is the window **not broken**?'

(32) *Nach bhfuil tú ag ól?*
NEG.Q BE:AUX-PRS 2SG at:PREP drinking:VN
'Aren't you drinking?' / 'Are you **not drinking**?'

But, interestingly, this particle can also be found on occasion with a question form expressed in the negative with past tense auxiliary verb (33).

(33) *Nach raibh feirmeoir ar bith sa tír?*
NEG.Q BE:AUX-PST farmer.N at all in:PREP+DET country.N
'Was there **no farmer** in the country at all?'

Stating this form of question in the positive has the format shown in (34).

(34) *An raibh sé anseo?*
Q BE:AUX-PST 3SG.M here:DEIC
'Was he here?'

3.5.1.4 Question clause with the negative copula

The construction format is symmetric over positive and negative polarity but the morphological form of the negative copula is based on the past/non-past distinction.

(35) *Nach maith leat do dhinnéar?*
NEG.Q good:ADJ with:PREP+2SG 2SG-POSS dinner.N
'Do you not like your dinner?'

(36) a. *An maith leat tae?*
Q like.V-PRS with:PREP+2SG tea.N
'Do you like tea?'

b. *Is maith!*
COP like-PRS
LIT: 'Be like!'
= 'Yes!'

c. *Ní maith!*
NEG.COP like-PRS
LIT: 'Not like!'
= 'No!'

(37) a. *Nárbh amhránaite maithe iad?*
NEG.COP singers.N good:ADJ 3PL
'Weren't they good singers?'

b. *Ba ea*
COP.PST they!
'Yes, they are!'

c. *Níorbh ea*
NEG.COP.PST they!
'No, they are not!'

3.6 Negative conditional clauses with matrix verbs

The *Mura/Murar* 'IF NEG-COND' use in negative conditional clauses of Modern Irish also pivots on the past/non-past tense of the matrix verbs.

(38) ***Mura** bhfuil tú anseo, ní thiocfaidh sé abhaile.*
If neg-COND BE-AUX-PRS 2SG here:DEM NEG come:V-FUT 3SG.M home:N
'If you are not here, he won't come home.'

(39) ***Murar*** *féidir leat é a dhéanamh, fág*
If neg-COND able:ADJ with:PREP+2SG 3SG.M.ACC REL do/make:VN leave:V-IMP
é.
3SG.M
'If you cannot do it, leave it.'

3.7 *Ná* – The form of NEG used only in imperative mood

The unique negative form of the imperative construction is shown in (40)–(43).

(40) *Ná mol!*
NEG praise.V-IMP
'Don't praise!'

(41) *Ná fág!*
NEG leave.V-IMP
'Don't leave!'

(42) *Ná buail mé!*
NEG praise.V-IMP 1SG
'Don't hit me!'

(43) *Ná bris na fuinneogaí.*
NEG praise.V-IMP DET.pl windows.N.pl
'Don't break the windows!'

3.8 Negative forms in exclamatory use with adjectives.

The use of the negative in exclamatory clauses with adjectives is sensitive to the past/non-past tense of the underlying verb. Here, we show the format with non-past copula in (44) and (45).

(44) *Nach maith é!*
NEG.Q good:ADJ 3SG.M.ACC
LIT: 'Is he not good!'
'Isn't he good!' [= 'He is good!']

(45) *Ní hionann iad!*
NEG same:ADJ 3PL.ACC
'They are not the same!'

4 Expressing negative polarity within the Irish noun phrase

As with the clause and the verb forms, Modern Irish has additionally interesting ways of expressing negative polarity on the noun and the noun phrase. The means of encoding negation

on the nominal system include those shown in (46). These consist of a preposition and several morphological prefixes, as well as some lexically negative nouns. We can note that some nouns lexically encode the meaning 'nothing'. These are related to the class of generic nouns, indefinites, negative indefinites and existential quantifiers found in Irish.

(46) The negative marking on the noun/noun phrase

Noun (phrase)	The nominal environment for negation
Gan 'not=without'	1. *Prepositional* form of NEG used with a *noun phrase*.
faic	2. *Noun* with the lexical meaning of 'nothing'
dada	3. Noun with the lexical meaning of 'nothing'
náid	4. Noun with the lexical meaning of 'nothing'
neamh-	5. *Prefix* on lexical item meaning of 'non'/'un-'/
nea-	6. Prefix on lexical item meaning of 'non'/'un-'/
mi-	7. Prefix on lexical item meaning of 'non'/'un-'/
an-	8. Prefix on lexical item meaning of 'non'/'un-'/
ain-	9. Prefix on lexical item meaning of 'non'/'un-'/
do-	10. Prefix on lexical item meaning of 'non'/'un-'/

Many languages use quantifiers that are inherently negative and that have the potential to interact within the grammar with negative indefinites. English, for example, used 'none', 'nothing' along with some phrases, such as 'not many' that are negated independently. Modern Irish is no different in this regard.

PERSON	*neamhdhuin*	Nobody
TIME	*in aon chor*	Never
PLACE	*in aon áit*	Nowhere
MANNER	*ar chaoi ar bith*	
	ar aon dóigh	No how
NEGATIVE/NO!	*ní hea*	
	is cuma	No (In Irish, based on COPULA!)
THING	*neamhní*	Nothing
	faic	Nothing
	náid	Nothing
	dada	Nothing
	neodar	Nothing
	aon ní	Nothing
	faic na fríde	Nothing at all
NONE	*cuid ar bith*	No part
	aon duine	Any person
	aon rud	Any thing
	aon cheann	Any one
	nóin	None
	lán nó loic	'All or nothing'
ANY	*ar aon chaoi*	In any case
SOME	*roinnt*	Some (people)
	(duine) éigin	Some (person)
	(daoine) áirithe	Some (person)
	cuid (de)	Some (= part of)

Table 3: Inventory of negative indefinites and quantifiers of Modern Irish

4.1 The form of morphological NEG used with a noun

Generally, for morphological negation used on a noun, one may employ several prefixes. The same prefix can also be used with an adjective to encode a negative adjective. These are all instances of morphological negation. When the prefixes used in (47) are employed on a noun, it exhibits negation on the nominal core.

(47) NEG-prefixes + N → 'non-' [NEG **on** CORE_{N}]

ain-	PREFIX	'non-'
an-	PREFIX	'non-'
do-	PREFIX	'non-'
mi-	PREFIX	'non-'
nea-	PREFIX	'non-'
neamh-	PREFIX	'non-'

(48) *ain*:NEG-prefix + N LEXICAL $\text{ITEM}_{\text{TYPE}}$ → 'un-/non-' $\text{ITEM}_{\text{TYPE}}$

ainfhíreán	NM	'unjust person'
ainfhirinne	NF	'untruth'

(49) *ain*:NEG-prefix + ADJECTIVE → 'un-/non-' $\text{ITEM}_{\text{TYPE}}$

aincheart	ADJ	'unjust'
aineoil	ADJ	'unknown'
aineolach	ADJ	'unknowing'
aineolach ar rud	ADJ	'unlearned of something'
ainfhéichiúnta	ADJ	'unpunctual'
ainfhíor	ADJ	'untrue'
ainfhíréanta	ADJ	'unjust'

(50) *an*:NEG-prefix + LEXICAL $\text{ITEM}_{\text{TYPE}}$ → 'un-/non-' $\text{ITEM}_{\text{TYPE}}$

anaithnid	ADJ	'unknown'
anaoibhniúil	ADJ	'unsmiling'

(51) *do*:NEG + LEXICAL $\text{ITEM}_{\text{TYPE}}$ → 'un-/non-'N_{TYPE}

do-ríomhchlár	NM	'non-program'[non computer program]
dofheiscint	NF	'unsightliness'

(52) *do*:NEG + LEXICAL $\text{ITEM}_{\text{TYPE}}$ → 'un-/non-'ADJ_{TYPE}

do-inste	ADJ	'untold'
do-iompraithe	ADJ	'non-reversible'
do-mhaite	ADJ	'unjustifiable'
do-úsáidte	ADJ	'unusable'

dobhogtha	ADJ	'unshakeable'
dochaideartha	ADJ	'unsociable'
dochorraithe	ADJ	'unshakeable'
dochosanta	ADJ	'unjustifiable'
dodhíolta	ADJ	'unsaleable'
dofhaighte	ADJ	'unprocurable'
dofhíoraithe	ADJ	'unverifiable'
doimeartha	ADJ	'unplayable'
doléirithe	ADJ	'unplayable'
doléite	ADJ	'unreadable'
doluaite	ADJ	'unspeakable'
domhúinte	ADJ	'unteachable'
donite	ADJ	'unwashable'
doráite	ADJ	'unspeakable'
doreoite	ADJ	'non-freezing'
dosmaoinimh	ADJ	'unthinkable'
dostoptha	ADJ	'unstoppable'
dostróicthe	ADJ	'untearable'

(53) *mí*:NEG + LEXICAL ITEM$_{\text{TYPE}}$ → 'un-/non-'N$_{\text{TYPE}}$

míbhinneas	NM	'untunefulness'
míbhlastacht	NF	'unsavouriness'
mífholláine	NF	'unwholesomeness'
mífhreagracht	NF	'unresponsiveness'
mímhaise	NF	'unseemliness'
míráth	NM	'unsuccess'
míshuaimhneas	NM	'unrest'

(54) *mí*:NEG + LEXICAL ITEM$_{\text{TYPE}}$ → 'un-/non-'ADJ$_{\text{TYPE}}$

mí-ordúil	ADJ	'unmethodical'
míbhanúil	ADJ	'unladylike'
míbhinn	ADJ	'unmelodious'
míbhinn	ADJ	'unmusical'
míbhreá	ADJ	'unlovely'
míbhuíoch	ADJ	'unthankful'
mícheolmhar	ADJ	'unmusical'
míchreidiúnach	ADJ	'untrustworthy'
mídheas	ADJ	'unlovely'
mídheas	ADJ	'unpleasant'

(55) *neamh*:NEG + LEXICAL ITEM$_{TYPE}$ → 'un-/non-'N$_{TYPE}$

neamhaontach	NM	'non-conformist'
neamhbheith	NF	'non-existence'
neamhchomhlíonadh	NM	'non-compliance'
neamhchomhlíonadh	NM	'non-fulfilment'
neamhchomhraiceoir	NM	'non-combatant'
neamhchríochnú	NM	'non-completion'
neamhdhuine	NM	'nobody'
neamhdhuine	NM	'nonentity'
neamhíoc	NM	'non-payment'
neamhionsaí	NF	'non-aggression'
neamhní	NM	'nonentity'
neamhní	NM	'nothing'
neamhní	NM	'nothingness'

(56) *neamh*:NEG + LEXICAL ITEM$_{TYPE}$ → 'un-/non-'ADJ$_{TYPE}$

neamhcheadaithe	ADJ	'unpermitted'
neamhcheangailte	ADJ	'non-attached'
neamhcheirdiúil	ADJ	'unskilful'
neamhcheolmhar	ADJ	'unmusical'
neamhchoimhéadach	ADJ	'unwatchful'
neamhchóir	ADJ	'unjust'
neamhchoitianta	ADJ	'unusual'
neamhchomharsanúil	ADJ	'unneighbourly'
neamhchosúil	ADJ	'unlike'
neamheolach	ADJ	'unversed'
neamheolaíoch	ADJ	'unscientific'
neamhfhileata	ADJ	'unpoetical'
neamhfhíréanta	ADJ	'unjust'
neamhghlic	ADJ	'unsophisticated'
neamhghnóthach	ADJ	'unoccupied'
neamhghrámhar	ADJ	'unloving'
neamhléite	ADJ	'unread'
neamhullamh	ADJ	'unprepared'
neamhullamh	ADJ	'unready'

In other languages, English, for example, 'un-' is largely constrained to adjectival stems and 'non-' usually is restricted to either adjectival or nominal stems. However, Modern Irish is much more diverse in its use of prefixes to encode negation on the nominal.

4.2 Analytical negation with the preposition *gan* 'without'

The preposition *gan* 'without' can be used with a nominal to encode analytical negation over its object N and is highly productive in usage.

(57) *Bhí sé ann* ***gan*** *fhios.*
AUX-PAST 3SG.M there:DEM without:NEG knowledge.N
'He was there without knowledge.'

(58) *gan*:PREP N 'without N'$_{\text{Irish}}$ → 'un-N'

gan ábhar	PREP N	'unjustification'
gan ainm	PREP N	'unnamed'
gan ainmniú	PREP N	'unspecified'
gan áireamh	PREP N	'untold'
gan aithint	PREP N	'unperceived'
gan aithne	PREP N	'unknown'
gan athrú	PREP N	'unvaried'
gan bhearradh	PREP N	'unshaven'
gan bhrath	PREP N	'unperceived'
gan bhrí	PREP N	'unmeaning'
gan bhriseadh	PREP N	'unvanquished'
gan bhuaireamh	PREP N	'untroubled'
gan bhualadh	PREP N	'unsurpassed'
gan bpholladh	PREP N	'un-perforated'
gan buaireamh	PREP N	'unperturbed'
gan chaomhnú	PREP N	'unprotected'
gan cheadúnas	PREP N	'unlicensed'
gan cheist	PREP N	'unquestionable'
gan cheol	PREP N	'unmusical (without music)'
gan chiall	PREP N	'unmeaning'
gan chónaí	PREP N	'unoccupied'
gan chosaint	PREP N	'unprotected'
gan chruthú	PREP N	'unproved'

Additionally, the preposition *gan* 'without' can take scope over the complete noun phrase, as well as a noun. When it is used with the noun phrase, it takes within its scope the definiteness operator.

(59) *gan*:PREP NP 'without NP'

(60) ***gan*** *an t-airgead*
without:NEG DET money.N
'Without **the money**.'

(61) *Bhí sé ann* ***gan*** *fhios d'éinne beo ná*
AUX-PAST 3SG.M there:DEM without:NEG knowledge.N anyone.N alive:ADJ CONJ
marbh.
dead:ADJ
'He was there **without knowledge of anyone alive or dead**.'

(62) a. *An dá theach mhóra*
DET.SG two:NUM house.N.sg big.ADJ.pl
'The two big houses.'

b. ***Gan*** *an dá theach mhóra*
Without:NEG DET.SG two:NUM house.N.sg big:ADJ.PL
'Without the two big houses.'

c. ***Gan*** NP
without N
'**NEG'**[N]'

We have seen examples of lexically negative nouns, morphologically marked negative nouns and analytically marked negative nouns and noun phrases.

(63) NEG on the NUC_{N}: [NEG-PREFIX+N]
a. *neamhduine* 'nobody'/'nonperson'
b. *neamhní* 'nonentity'

(64) NEG on the CORE_{N}: [GAN 'WITHOUT' N]
a. *gan airgead* 'without money'
b. *gan fhios* 'without knowledge'

(65) NEG on the NP: [GAN 'WITHOUT' NP]
a. *gan an t-airgead* 'without the money'
b. *gan an dá theach mhóra* 'without the two big houses'

5 Discussion

We have earlier seen examples of negative encoding on the verb and the clause. A negative clause is one which makes an assertion that some event, situation or state of affairs does

not hold. Modern Irish makes a distinction on the forms used according to whether a clause contains a lexical matrix verb, the auxiliary form of the verb 'to be', or is a copula clause.

Within a clause, a further distinction is made based on whether the verb encodes the past tense or not and which may relate to a realis/irrealis sensitivity. *Irrealis* indicates that a certain situation or action is known to have not happened at the time of the speaker talking. With non-past tenses, *irrealis* is used to express something that is not known to be the case in reality. Then, with past tense encoding *realis*, the speaker indicates that something is actually the case (or not, if negated). That is, *realis* strategies express what the speaker considers to be a known state of affairs, broadly as in positive or negative declarative sentences.

We have seen clause negation and constituent negation interact in complex ways. Modern Irish uses the clausal form of the attributive auxiliary construction to attribute negative polarity on a particular constituent. Additionally, negation interacts in many ways with indefiniteness and quantification. We have also examined evidence that Modern Irish employs lexical, morphological and analytical means to encode negation with the appropriate level of scope on the LSC, LSNP and LSW.

(66) NEG **on** NUC_V: [NEG-PREFIX+VERB]

a. ***dí***:NEG + lexical V → NEG.V

b. ***dí***:NEG + lexical V → NEG.VA

c. ***dí***:NEG + lexical V → NEG.VN

(67) NEG **on** NUC_N: [NEG-PREFIX+NOUN]

a. ***do***:NEG + LEXICAL ITEM_TYPE → 'un-/non-'N_TYPE

b. ***do***:NEG + LEXICAL ITEM_TYPE → 'un-/non-'V_TYPE

c. ***do***:NEG + LEXICAL ITEM_TYPE → 'un-/non-'ADJ_TYPE

d. ***mí***:NEG + LEXICAL ITEM_TYPE → 'un-/non-'N_TYPE

e. ***mí***:NEG + LEXICAL ITEM_TYPE → 'un-/non-'ADJ_TYPE

f. *neamh*:NEG + LEXICAL ITEM_TYPE → 'un-/non-'N_TYPE

g. *neamh*:NEG + LEXICAL ITEM_TYPE → 'un-/non-'ADJ_TYPE

We have found in particular that in respect of the noun, negation is found on the nominal nucleus, core and noun phrase. Additionally, we have seen that Irish has a series of generic nouns that function in negative polarity contexts; for instance, *rud* functions both as a negative polarity indefinite ('anything'/'nothing') and as a generic common noun ('thing'), as does *duine* 'person, anyone'.

Therefore, a consequence of this for the RRG model is that with respect to Modern Irish, the model must allow negation to be encoded within the, NUC_N CORE_N and the NP. NEG must also be a possibility at each level in the LSNP operator projection.

(68) NP, Core_N and Nuclear_N operators for the LSNP of Modern Irish

Nuclear_N operator

- Nominal aspect
- NEGATION

Core_N operators

- Number
- Qualification (quantifiers)
- NEGATION

NP operators

- Definiteness
- NEGATION
- Deixis

This provides evidence that, for modern Irish, encoding of negation can occur at the NUC, CORE and NP levels within the layered structure of the noun phrase. That is, it reflects the negation possibilities found in the LSC where negation can also occur at the three levels. The seemingly 'simple' nature of logical negation as an operator on the LSC or LSNP that reverses positive and negative polarity hides the complex and subtle expression of negation actually found in Modern Irish (see Figure 4). We have found a rich and complex system of negative particles sensitive to the past/non-past distinction; negation over lexical verbs, auxiliary verbs and copulas, morphological and analytical negation interacting with quantifiers. We have also seen evidence of the interaction of negation with other operators, including nominal negation, in non-trivial ways. Lexical negation describes a situation in which the concept of negation is part and parcel of the lexical semantics of a particular verb.

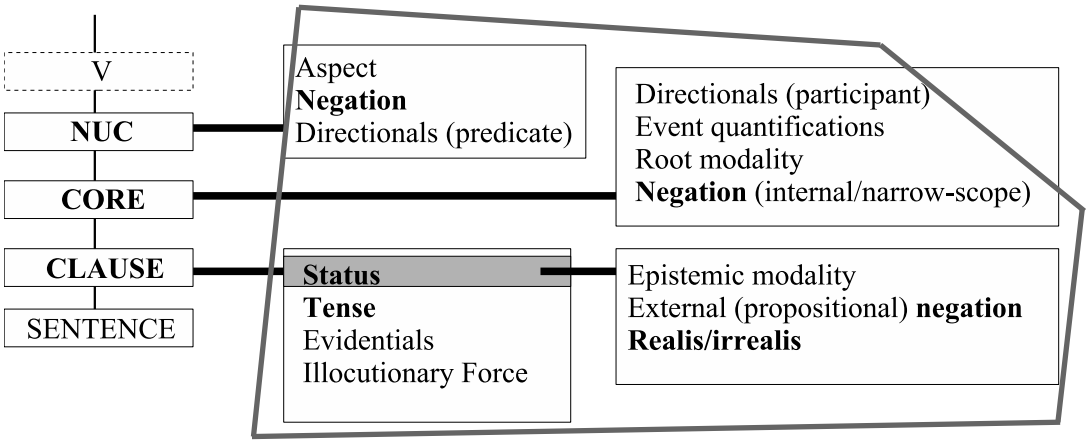


Fig. 4: Map of inter-related clausal negation space in the LSC for Modern Irish

Payne (1985) introduced the notion of standard negation as 'that type of negation which has one function, the negation of the most minimal and basic sentences'. Standard negation is defined as an operator taking sentential scope, typically in the form of a verbal predicate in a declarative clause exploiting the general strategy available in a given language. Negation, however, is not that 'standard', when it shows interaction with realis/irrealis, verbal tense and aspect, and has occurrence in question, existential and copula clauses, and the other various ways we have seen here.

References

Baker, Mark. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3): 373–415.

Horn, Laurence R. (ed.). 2011. *The expression of negation* [The expression of cognitive categories 4]. Berlin/New York: de Gruyter.

Miestamo, Matti. 2005. *Standard Negation. The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin/New York: de Gruyter.

—. 2010. Negatives without negators. In *Rethinking Universals: How Rarities Affect Linguistic Theory*, J. Wohlgemuth & M. Cyouw (eds), 169–194. Berlin/New York: de Gruyter.

Nolan, Brian. 2012. *The syntax of Irish: A functional account*. Sheffield: Equinox.

Payne, John. 1985. Negation. In *Language Typology and Syntactic Description, Volume I: Clause structure*, T. Shopen (ed.), 197–242. Cambridge: CUP.

Rijkhoff, Jan. 2002. *The Noun Phrase*. Oxford: OUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax*. Cambridge: CUP.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Some Issues Regarding (Active) Accomplishments

Robert D. Van Valin, Jr.
Heinrich-Heine-Universität, Düsseldorf,
University at Buffalo, The State University of New York

Abstract

Role and Reference Grammar is unique in recognizing bounded activities as a distinct *Aktionsart* category, labeling them 'active accomplishments', in order to distinguish them from bounded processes, which are termed simply 'accomplishments'. Active accomplishments are divided into two subtypes, motion active accomplishments, and non-motion, each with a slightly different logical structure. Van Valin (2005) argued for a change in the decomposition of accomplishments, which had serious implications for the analysis of active accomplishments. This paper argues that this change resulted in a semantic representation which made it possible to capture the incremental nature of many activity predicates and the corresponding active accomplishments. In addition, it is argued that motion active accomplishments have a different syntactic representation from non-motion active accomplishments.

1 Introduction[1]

Role and Reference Grammar (RRG) is unique in recognizing bounded activities as a distinct *Aktionsart* category, labeling them 'active accomplishments', in order to distinguish them from bounded processes, which are termed simply 'accomplishments'. In the original Vendler classification, given in (1), they contrast with achievements as change of state predicates in terms of punctuality, with achievements being punctual and accomplishments being non-punctual.

(1) Vendler's (1957 [1967]) classification

a. States, e.g. *have*, *possess*, *desire*, *want*, *like*, *love*, *hate*, ...
[inherently unbounded, non-dynamic state of affairs]

[1] I would like to thank Hana Filip, Jens Fleischhauer and Anja Latrouite for comments on an early draft of this paper. This research was supported in part by the German Science Foundation through CRC 991 'The structure of representations in language, cognition and science'.

b. Activities, e.g. *run*, *walk*, *swim*, *push*, *pull*, ...
[inherently unbounded, dynamic state of affairs]

c. Achievements, e.g. *recognize*, *realize*, *find*, *cross the border*, *reach the summit*, ...
[punctual change of state or onset of an activity]

d. Accomplishments, e.g. *paint a picture*, *write a novel*, *build a house*, *recover from illness*, *harden*, ...
[non-punctual change of state (bounded process) or onset of an activity]

This system has been reinterpreted and formalized in Dowty (1979), which has been highly influential and which motivated the systems of lexical decomposition in Foley & Van Valin (1984) and later in the work of Rappaport Hovav and Levin (e.g. Levin & Rappaport Hovav 1995, Rappaport Hovav & Levin 1998). Dowty reinterpreted achievements and accomplishments, using them in different senses from the ones proposed by Vendler. States and activities, on the other hand, were used in Vendler's original senses.

The purpose of this paper is to examine the notions of 'accomplishment' and 'active accomplishment' as they have been used in RRG and to update their semantic and syntactic representations within the theory. The discussion will proceed as follows. Section 2 summarizes Dowty's reinterpretation of Vendler and its application in early RRG, followed by a critical examination of the system in Section 3. Section 4 presents the revised system of decomposition proposed in Van Valin & LaPolla (1997), which proposes a return to Vendler's original classification. Section 5 concerns active accomplishments and their analysis. Section 6 explores the possible decomposition of the operator BECOME and its implications for lexical decomposition. Section 7 addresses the syntactic representation of active accomplishments, and Section 8 presents conclusions.

2 Dowty's (1979) formalization of the Vendler classification

In his 1979 book *Word meaning and Montague Grammar*, Dowty proposed a formalization of Vendler's classes that has been so influential that the distinctions in (1) are often referred to as the 'Dowty-Vendler' classification. Dowty proposed that states constituted the basic building blocks of the system, with the other classes being derived from states by the addition of specific operators and connectives. The system is summarized in (2).

(2) a. States: **predicate'**(x, (y)), e.g. **know'**(John, answer)
John knows the answer.

b. Activities: DO (x, [**predicate'**(x, (y))]), e.g. DO (John, [**walk'**(John)])
John walked.

c. Achievements: BECOME **predicate'**(x, (y)), e.g. BECOME **shattered'**(window)
The window shattered.

d. Accomplishments: [DO (w, [**predicate'**(w, (x))])] CAUSE [BECOME **predicate'**(y, (z))], e.g. [DO (Bill, [**affect'**(Bill, window)])] CAUSE [BECOME **shattered'** (window)]
Bill shattered the window.

'DO', taken from Ross (1972), was the defining feature of activity predicates; in Ross' original analysis, 'DO' signaled that its argument was agentive, and accordingly, in (2b) *John* in *John walked* would be an agent in thematic role terms. Achievements were constituted by a state predicate plus the operator 'BECOME', which indicated a punctual change. Finally, accomplishments in Dowty's systems were distinguished by two characteristics: causality and bounded change. The canonical accomplishment had the structure '[activity] CAUSE [achievement]', as illustrated by the example in (2d), which can be paraphrased as 'Bill's doing something to the window caused it to shatter'. [2]

Accomplishments include the transitive verbs *melt*, *freeze*, and *dry*, for example, but their intransitive forms don't fit directly into Dowty's system, since they are non-punctual changes of state, e.g. *The ice melted slowly*. Dowty proposed that such verbs form a subclass of achievements which he termed 'degree achievements'. These would have been included in accomplishments in Vendler's original classification.

Early work in RRG (Foley & Van Valin 1984) adopted the Dowty decompositional system in (2) as the basis of the semantic representations in a monostratal theory of the linking between syntax and semantics. It was used together with the Gruber-Jackendoff idea of defining thematic relations in terms of argument positions in decomposed semantic representations to create an independently motivated system of thematic relations assignment. Van Valin (1990, 1993) departed from the Foley & Van Valin system slightly. First, a distinction between punctual and durative achievements was made (but the BECOME decomposition was kept for both); this mirrored Dowty's distinction between (plain) achievements and degree achievements. Second, 'DO' occurred only in the decomposition of verbs which lexicalized agency, following Holisky (1987) (see also Van Valin & Wilkins 1996); '**do'**(x)' was introduced to represent effector arguments which may or may not be construed as agents.

[2] Dowty eventually settled on a different set of distinctions (1979: 184), which were not formalized in this way.

3 Problems with accomplishments

Since accomplishments are bounded processes (telic) and the only causative type in (2), this would seem to imply that causation and telicity are inherently linked, that is, only telic verbs can be causative. This is problematic in two ways. First, it entails that the accomplishment use of activity verbs involves causation. Activity verbs can also have accomplishment uses; in fact, Dowty maintains that all activities can have accomplishment uses. For verbs of motion, the addition of a specific goal renders them bounded, e.g. *John ran in the park for an hour* (activity) vs. *John ran to the park in an hour* (accomplishment). With verbs of creation and consumption, the addition of a specific or quantified object makes them telic and therefore accomplishments, e.g. *John drank beer for an hour* vs. *John drank a glass of beer in an hour* (consumption), *Mary wrote poetry for an hour* vs. *Mary wrote two poems in an hour* (creation). In terms of the decomposition system in Van Valin (1993), the verbs would have the following causative decompositions.

(3) a. Motion verbs
run [activity] vs. *run to the store* [accomplishment]
run'(Mary) vs. [**run'**(Mary)] CAUSE [BECOME **be-at'**(store, Mary)]

b. Consumption verbs
eat (pizza) [activity] vs. *eat a pizza* [accomplishment]
eat'(John, (pizza)) vs. [**eat'**(John, pizza)] CAUSE [BECOME **consumed'**(pizza)]

c. Creation verbs
write (poetry) [activity] vs. *write a poem* [accomplishment]
write'(John, (poetry)) vs. [**write'**(John, poetry)] CAUSE [BECOME **exist'**(poem)]

In Van Valin & LaPolla (1997) it is argued that this decomposition is fundamentally flawed. First, causative predications involve both a causer and a causee, but there is only one participant in (3a), which cannot be both causer and causee; *Mary ran to the store* does not mean the same thing as *Mary made herself run to the store*. Moreover, in languages which explicitly mark this contrast on verbs morphologically, causative morphology is never used. The following Georgian examples are of a creation verb in (4) and a consumption verb in (5); the morphology on the verb in the accomplishment versions in (4b) and (5c), the preverbs *da-* and *še-*, is unrelated to the causative morphology in (5b), in this case the suffix *-ev* and the 'version' vowel *-a-*.[3]

[3] See Samsel (1992) for an RRG analysis of Georgian causative and version forms.

(4) Georgian (Holisky 1981)

a. K'ac-i (c'eril-s) c'er-s xuti saati.
man-NOM (letter-DAT) write.PRS-3SG five hours
'The man is writing (letters) for five hours.' Atelic

b. K'ac-i c'eril-s da-c'er-s at c'ut-ši.
man-NOM letter-DAT PRV-write.PRES-3SG ten minutes-in
'The man will write the letter in ten minutes.' Telic

(5) Georgian (Aronson 1982)

a. Č'am-s
eat-3SG
'He eats/is eating (it).' Plain

b. Ø-a-č'm-ev-s
3SG-VER-eat-CAUSE-3SG
'She makes him eat (it).' Causative

c. Ø-še-č'am-s
3SG-PRV-eat-3SG
'He eats it up.' Telic

Adding causative morphology to a verb meaning 'eat' yields 'cause to eat' or 'feed', not 'eat to completion', as in (5); this is further illustrated in Lakhota (a Siouan language of North America), e.g. *wótA* 'eat [atelic]' vs. *wol-khíyA* [eat + cause] 'feed'. Adding causative morphology to verbs of motion yields verbs meaning 'send', not 'go to a goal', e.g. Lakhota *yÁ* 'be on the way going' vs. *ye-khíyA* [go+cause] 'send'. Thus, the causative analysis of the accomplishment verbs in (3) is highly problematic.

The second problem is that linking telicity and causality would appear to preclude the possibility of causative versions of atelic verbs, i.e. states and activities, yet such verbs seem clearly to exist.[4]

Consider the English sentence, *What you just said upsets me*. The verb *upset* here is stative, as it passes the English test for stativity: in the simple present tense, it can have a present tense meaning rather than a habitual reading. This example clearly has a present tense interpretation. Furthermore, it is causative, as it can be paraphrased with an explicitly causative construction: *What you just said makes me upset*. Thus, *upset* in this example has a causative state interpretation. In (6) from Mparntwe Arrernte (Wilkins 1989) there is another example of a causative state.

[4] It should be noted that Dowty gives examples of causative state and causative activity representations in his aspectual calculus (1979: 123–124), but they do not Figure prominently in his analysis.

(6) Mparntwe Arrernte [Australia]

a. Ayenge irrernte ne-ke.
1SG.ABS cold COP-PAST
'I was cold.' State

b. Kwatye-le ayenge irrernte-arle-lhile-ke.
water-ERG 1SG.ABS cold-RES-CAUSE-PAST
'The water kept me cold.' Causative state

c. Kwatye-le ayenge irrernte-arle-irre-lhile-ke.
water-ERG 1SG.ABS cold-RES-CHANGE-CAUSE-PAST
'The water cooled me down [to the point that I was cold].'
Causative change of state

(6a) is a basic state predication, 'to be cold', whereas there is an explicit causative morpheme, *-lhile-*, in (6b) which has the meaning 'to keep someone/something cold'. While it might be suggested that this involves a change of state, this can be ruled out by comparing it with (6c), a causative change of state in which both change (*-irre-*) and causation (*-lhile-*) are explicitly coded. Activity verbs can be causativized without telicity, as in the English examples in (7a, a'), and as in (7b') in which the *Arrernte* causative morpheme is added to an activity verb to code a causative activity.

(7) a. The soldiers marched in the park (for an hour). Activity

a.' The sergeant marched the soldiers in the park (for an hour).
Causative activity

b. Mparntwe Arrernte

Ayenge unthe-ke.
1SG.ABS wander.around-PAST (Wilkins 1989)
'I wandered around.' Activity

b.' Kwementyaye-le ayenge unthe-lhile-ke.
-ERG 1SG.ABS wander.around-CAUSE-PAST
'Kwementyaye made me wander around.' Causative activity

Hence, causality and telicity are independent parameters with respect to the classification of verbs.

The term 'accomplishment' has thus been used to cover four different phenomena. First, it was used by Vendler for non-punctual changes of state, e.g. *harden* as in *The ice cream hardened* (related to states, e.g. *The ice cream is hard*); some of these are termed 'degree achievements' by Dowty. Second, it has been used by Vendler and Dowty to refer to activities with an endpoint, e.g. *write a novel*, *eat a bagel* (alternates with activities, e.g. *write*, *eat*).

Third, it has been used by Dowty to refer to caused changes of state with endpoint and result, e.g. *kill*, *harden* as in *The dry ice hardened the ice cream*. Fourth, it has been used to refer to caused activities with an endpoint, e.g. *feed the cake to X*, *march X to the park*. In the next section, the revised RRG system will be introduced, in which these four uses are assigned to different *Aktionsart* classes.

4 Back to Vendler: the revised RRG system

In Van Valin & LaPolla (1997) the system of *Aktionsart*-based lexical decomposition was thoroughly revised, and an important aspect of this revision was a return to Vendler's original distinctions. States and activities are taken as the primitive building blocks of the system; they are the only classes which take arguments; it is not assumed, as in Dowty (1979), that activities are derived from states. States receive the same formal representation as before, while activities all carry **do'**, which signals that they are activities. Thus, instead of representing *run* as **run'**(x) as in (3a), it would now be represented as **do'**(x, [**run'**(x)]). With respect to achievements and accomplishments, there is a return to Vendler.

Verb Class	Logical Structure
STATE	**predicate'**(x) or (x,y)
ACTIVITY	**do'**(x, [**predicate'**(x) or (x, y)])
ACHIEVEMENT	INGR **predicate'**(x) or (x,y), or
	INGR **do'**(x, [**predicate'**(x) or (x, y)])
ACCOMPLISHMENT	BECOME **predicate'**(x) or (x,y), or
	BECOME **do'**(x, [**predicate'**(x) or (x, y)])
ACTIVE ACCOMPLISHMENT	**do'**(x, [**predicate'**(x, y)]) & BECOME **predicate'**(y)[5] [consumption/creation] or
	do'(x, [**predicate'**(x)]) & BECOME be-LOC'(y, x) [bounded motion]
CAUSATIVE	α CAUSE β, where α, β are LSs of any type

Table 1: Lexical representations for *Aktionsart* classes in Van Valin & LaPolla (1997)

The term 'achievement' is reserved for a punctual change of state or onset of activity (indicated by an 'INGRessive' operator), e.g. *shatter* in (2c) would be INGR **shattered'**(x). The term 'accomplishment' is used for a non-punctual change of state or onset of activity (indi-

[5] '&' stands for 'and then'; '∧' stands for 'and simultaneously'.

cated by 'BECOME' operator), e.g. intransitive *melt* would be BECOME **melted'**(x). Verbs in expressions like those in (3) and (4b) are termed 'active accomplishments' in order to highlight the fact that the core predicate in them is an activity predicate, not a state predicate as in pure change-of-state accomplishments. Causation is factored out as an independent parameter, such that all five classes have a causative version. The revised decompositional system is summarized in Table 1, and examples of each type are given in (8). Active accomplishments have two different decompositions, depending upon the nature of the activity predicate involved: with motion activity verbs like *run*, the activity is intransitive, and in an active accomplishment predication reaching the endpoint is represented by 'BECOME **be-LOC'**(y, x), whereas when the activity predicate has two arguments, e.g. *drink* or *write*, then the completion of the consumption or creation is indicated by 'BECOME **consumed'**(y)' or 'BECOME **exist'**(y)'.

(8) a. STATES

The window is shattered.	**shattered'** (window)
John saw the picture.	**see'** (John, picture)[6]

b. ACTIVITIES

The baby cried.	**do'**(baby, [**cry'** (baby)])
Carl ate pizza.	**do'**(Carl, [**eat'**(Carl, pizza)])

c. ACHIEVEMENTS

The window shattered.				INGR **shattered'** (window)
(Totuzen)	akanboo	ga	naki-dasi-ta.	INGR **do'** (baby, [**cry'** (baby)])
(suddenly)	baby	NOM	cry-INGR-PAST	

The baby burst out crying (suddenly). (Japanese, Toratani 2002, pers. comm.)

d. ACCOMPLISHMENTS

The snow melted.				BECOME **melted'** (snow)
(??Totuzen)	akanboo	ga	naki-hazime-ta.	BECOME **do'**(baby, [**cry'** (baby)])
(suddenly)	baby	NOM	cry-begin-PAST	

The baby started crying (??suddenly). (Japanese, Toratani 2002, pers. comm.)

[6] The order of arguments in the representations is determined by a semantic role hierarchy: the leftmost argument in the decomposition is the most agent-like, the rightmost the most patient-like. The actual linear realization of the arguments in a clause is a function of a language's semantics-to-syntax linking algorithm, which will not be discussed here. See Van Valin (2005) for detailed discussion.

e. ACTIVE ACCOMPLISHMENTS

Carl ate the pizza.	**do'**(Carl, [**eat'**(Carl, pizza)]) & BECOME **consumed'** (pizza)
Mary ran to the park.	**do'**(Mary, [**run'**(Mary)]) & BECOME **be-at'**(park, Mary)

f. CAUSATIVES

The dog scares the boy.	[**do'**(dog, ∅)] CAUSE [**feel'**(boy, [**afraid'**])]
Felix rolled the ball.	[[**do'**(Felix, ∅)] CAUSE [**do'**(ball, [**roll'** (ball)])]
Max shattered the window.	[**do'**(Max, ∅)] CAUSE [INGR **shattered'** (window)]
The hot sun melted the snow.	[**do'**(sun, ∅)] CAUSE [BECOME **melted'** (snow)]
Mary fed the pizza to the child.	[**do'**(Mary, ∅)] CAUSE [**do'**(child, [**eat'**(child, pizza)]) & BECOME **consumed'** (pizza)]
Mary sent Bill to the store.	[**do'**(Mary, ∅)] CAUSE [**do'**(Bill, [**go'**(Bill)]) & BECOME **be-at'**(store, Bill)]

The four subtypes of accomplishments discussed at the end of §3 above are now in four distinct but related categories. The non-punctual changes of state e.g. intransitive *dry*, *melt* and *freeze*, are 'accomplishments', represented as in (8d). The activities with an endpoint, e.g. *write a poem*, *drink a beer*, are 'active accomplishments', as in (8e). The caused changes of state with endpoint and result, e.g. transitive *dry*, *melt* and *freeze*, are 'causative accomplishments', as with *melt* in (8f). Finally, caused activities with an endpoint and result are 'causative active accomplishments', e.g. *feed* and *send* in (8f). This summarized in Table 2.

Predication type	*Aktionsart* class	Example
Non-punctual change of state	Accomplishment	Intransitive *dry*
Caused non-punctual change of state	Causative Accomplishment	Transitive *dry*
Bounded activity	Active Accomplishment	*run to the store*, *eat a pizza*
Caused bounded activity	Causative Active Accomplishment	*send X to the store*, *feed X a pizza*

Table 2: The four types of 'accomplishments'

Achievements are likewise split into plain and causative classes, e.g. intransitive (non-causative) *shatter* in (8c) and the transitive (causative) version in (8f).

5 Are active accomplishments a legitimate class?

It has been suggested that since many examples of active accomplishments are telic uses of activity verbs, they should not be a separate class and should be treated as derivative of activities. However, there are lexical active accomplishments, the best example of which is English *devour*, e.g. *Bill devoured the pizza in ten minutes* (**but there was some left*); even with a mass noun object the completion is still there, e.g. *Bill devoured pizza for ten minutes* (*and there was nothing left*/**but there was some left*). Furthermore, some verbs of motion involve motion to a goal, e.g. English *go*, *come*, Italian *andare* (Van Valin 1990). Finally, in some languages, the base form of verbs like *eat* and *drink* is telic and the atelic form derived;[7] this is particularly true in ergative languages, as exemplified in (9) and (10).

(9) Dyirbal [Aust.] (Dixon 1972)

a. Balam wudyu-∅ baŋgul yaɽa-ŋgu dyaŋga-ɲu.
NM.ABS fruit-ABS NM.ERG man-ERG eat-NFUT
'The man ate the fruit.' Telic

b. Bayi yaɽa-∅ dyaŋgay-mari-ɲu (bagum wudyu-gu).
NM.ABS man-ABS eat-ANTI-NFUT NM.DAT fruit-DAT
'The man is eating (fruit).' Atelic

(10) Sama [Philippines] (Walton 1986)

a. Inum na d'nda kahawa.
drink PRT woman coffee
'The woman already drank the coffee.' Telic

b. N-inum na d'nda (kahawa)
ANTI-drink PRT woman (coffee)
'The woman is now drinking (coffee).' Atelic

In Dyirbal the basic ergative-absolutive case marking pattern with a verb like *dyaŋga-* 'eat' correlates with the active accomplishment reading, as in (9a), and in order to derive the pure activity interpretation, the verb must be antipassivized, as in (9b). The same is true in Sama, a Philippine language: the basic form of the verb *inum* 'drink' has an active accomplishment interpretation, as in (10a), whereas the antipassivized form in (10b) has only the activity reading. A further example can be found in Lakhota, in which the basic, underived forms of consumption and creation verbs are telic, and the atelic forms must be derived, as illustrated in (11).

[7] It should be noted that Rothstein (2004) claims that English *read* is lexically telic and therefore would be a lexical active accomplishment verb in terms of Table 1. This analysis is not widely accepted, however.

(11) Lakhota (Latrouite & Van Valin 2014)

a. Hokšíla ki ağúyapiskuyela ki Ø-Ø-yúte.
boy the cookie the INAN-3SG.A-eat
'The boy is eating/ate the cookie(s).' Telic

b. Hokšíla ki ağúyapiskuyela eyá Ø-Ø-yúte.
boy the cookie some INAN-3SG.A-eat
'The boy is eating/ate some [specific] cookies.' Telic

c. Hokšíla ki Ø-Ø-yúte.
boy the INAN-3SG.A-eat
'The boy is eating/ate it', *'The boy is eating/ate'. Telic

d. Hokšíla ki w-Ø-óte. [*wa-* + *yútA* = *wótA*]
boy the NSO-3SG.A-eat
'The boy is eating/ate', *'The boy is eating/ate it'. Atelic

In (11a–c) the verb *yútA* 'eat [something specific]' occurs with a definite object, and indefinite specific object and no object, and as Latrouite & Van Valin (2014) show, they are all telic. To get a non-telic interpretation, the non-specific object prefix *wa-* must be added. Thus, there are non-derived active accomplishment verbs, and in some cases involving consumption and creation predicates, the base form of the verb has an active accomplishment interpretation, while the activity use is explicitly derived from it. Hence it may be concluded that active accomplishments are a legitimate class.

There has been a vigorous debate regarding whether *Aktionsart* properties are properties of verbs or of verb phrases, e.g. Verkuyl (1993) vs. Levin & Rappaport Hovav (1995), Van Valin & LaPolla (1997), and active accomplishments have been at the center of that debate. Cortés-Rodríguez (2014) argues that since there are lexical active accomplishments, e.g. English *devour* and motion verbs like *enter* and their Spanish equivalents, the category should be recognized but restricted to such verbs only, excluding phrasal expressions like *run to the park* and *eat a pizza*. This seems entirely appropriate, as it makes a clear contrast between lexical properties of verbs and the interpretation of the core. It sets up the possibility of classifying languages in terms of whether telicity is primarily a lexical property of verbs, as illustrated in Dyirbal, Sama and Lakhota, or primarily a phrasal property, as seems to be the case in English.

6 Analyzing BECOME: change + termination/result

Accomplishment change-of-state predications are represented by BECOME plus a state predicate, as in (8d), BECOME **melted'**(x) for *The ice melted*. In such a sentence, there is a process

of change, a termination of that process, and a result state. If one wanted to refer to the process alone, one can use the progressive, *The ice is/was melting*. Indeed, it is possible to explicitly refer to the process independently of the result even without the progressive, as in *The ice cream melted for half an hour, and then Mary put it back in the freezer.* Do other languages code the process separately from the termination and result state, and if so, how could such a pure process be represented? The answer to the first question is 'yes', as the following examples from Mparntwe Arrernte and Lakhota show.

(12) Mparntwe Arrernte [Australia] (Wilkins 1989)

a. Ayenge irrernte ne-ke.
1SG.ABS cold COP-PAST
'I was cold.'

b. Ayenge irrernt-irre-ke.
1SG.ABS cold-CHANGE-PAST
'I got cooler [but not to the point of being cold].'

c. Ayenge irrernte-arle-irre-ke.
1SG.ABS cold-RES-CHANGE-PAST
'I got cold.'

(13) a. Arntape urrperle ne-me.
tree.bark black COP-PAST
'Some tree bark is black.'

b. Arntape urrperl-irre-ke.
tree.bark black-CHANGE-PAST
'Some tree bark became blacker/darker.'

c. Arntape urrperle-arle-irre-ke.
tree.bark black-RES-CHANGE-PAST
'Some tree bark became black.'

In the Mparntwe Arrernte examples in (12a) and (13a) there are simple stative predications, 'be cold' and 'be black', respectively. To both of these, the suffix *-irre* is added, which signals change and creates a process without termination, as the translations of the (b) examples reflect. If one wants to indicate that the change went to completion, yielding a result state, then the suffix *-arle* must be added, as in the (c) examples.

In Lakhota these distinctions are signaled by different auxiliary verbs co-occurring with state predicates.

(14) Lakhota [Siouan; N. Amer.]

a. Ma-khúže.
1SG.U-sick
'I am sick.'

b. Khúš a<má>ye.
sick <1SG.U>AUX$_1$
'I'm getting sick(er and sicker).'

c. Khúš a<má>u.
sick <1SG.U>AUX$_2$
'I'm on the verge of being sick [e.g. on the verge of vomiting].'

Khúš a<má>hi.
sick <1SG.U>AUX$_3$
'I got sick.'

(15) a. Wakpála ki púze.
creek the dry
'The creek is dry.'

b. Wakpála ki pús áye.
creek the dry AUX$_1$
'The creek is getting dry (drier and drier).'

c. Wakpála ki pús aú.
creek the dry AUX$_2$
'The creek is (on the verge of) drying up.'

d. Wakpála ki pús ahí.
creek the dry AUX$_3$
'The creek has dried up.'

The (a) examples are simple stative predications, while the (b)–(d) examples each have a different auxiliary verb, which is characterized in (16).

(16) a. AUX$_1$: *áyA* 'to become gradually, to be changing, to grow into a new state of being, to be turning into, developing. Follows stative verbs; indicates a gradual change with cumulative progress.'(Ullrich 2008: 63) Literal meaning: 'to be on the way taking something/somebody there (away from here)...describes unfinished action . . . ' (Ullrich 2008: 63)

b. AUX$_2$: *aú* 'to gradually become, to be/keep changing, to grow into a different state of being, to be turning into (the change is in progress).' Literal meaning: 'to be coming bringing something/somebody, to carry something along while coming here.' (Ullrich 2008: 59).

c. AUX$_3$: *ahí* 'to have become, to have changed, to have grown or tuned into a different state of being (the change has been completed).' Literal meaning: 'to arrive here bringing something, to bring something here.' (Ullrich 2008: 27).

The first auxiliary, *áyA*, signals a process of change, which accumulates, but no endpoint, in the (b) examples. The second one, *aú*, also indicates a process of change, one which is on the verge of being completed, in the (c) examples, while the third one, *ahí*, marks a completed process with result state, in the (d) examples. Thus, in both Mparntwe Arrernte and Lakhota processes can be explicitly coded independent of any culmination of the process or result.

These facts strongly suggest that BECOME should be decomposed into a process plus a termination and result, and based on the Mparntwe Arrernte data in (12) and (13), Van Valin (2005) proposed to do just that, introducing 'PROC' as the logical structure element that signals a process. There is already a termination operator in the system, INGR, and consequently BECOME can be decomposed into PROC & INGR. Thus (12b) would be PROC **cold'**(1SG) 'I'm getting cooler and cooler', while (12c) would be PROC **cold'**(1SG) & INGR **cold'**(1SG) 'I got cold'.[8] Thus there is a new way to characterize the two types of non-causative accomplishments: 'process' accomplishments, as in e.g. (12c), (13c), (14d) and (15d), vs. active accomplishments.

The proposed decomposition of BECOME has implications for the analysis of active accomplishments, since they contain BECOME. If it is replaced by PROC & INGR, then the active accomplishment representations in Table 1 would be as in (17).

(17) a. Motion to a goal: **do'**(x, [**run'**(x)]) & BECOME **be-LOC'**(y, x) = **do'**(x, [**run'**(x)]) & PROC **????'**(x) & INGR **be-LOC'**(y,x)

b. Consumption: **do'**(x, [**eat'**(x,y)]) & BECOME **consumed'**(y) = **do'**(x, [**eat'**(x, y)]) & PROC **consumed'**(y) & INGR **consumed'**(y)

c. Creation: **do'**(x, [**write'**(x,y)]) & BECOME **exist'**(y) = **do'**(x, [**write'**(x,y)]) & PROC **create'**(y) & INGR **exist'**(y)

Taken literally, the revised representation in (17a) says 'x runs and then some sort of process involving x occurs and then there is a termination of the process and x comes to be located at y'. This is a rather implausible semantic representation for *x ran to y*, as this event does not consist of an activity followed by a process which then terminates. A similar problem affects the representation in (17b): it says 'x eats y and then y undergoes the process of

[8] The postulation of processes as an *Aktionsart* category of its own fills a gap in Vendler's system. Dowty (1979) is forced to analyze a predication like *the ice cream is melting* as an activity, since it is the only non-stative, unbounded class. But this is incorrect, because verbs like intransitive *melt*, *freeze* and *dry* do not take dynamic adverbs like *vigorously*, *violently*, *energetically*, or *intensely*, which true activity verbs do take.

being consumed and then the process of consumption ends'. Here again there is an activity followed by a process followed by its termination, which does not seem to express the right event structure for *x devoured y*. Similarly in (17c) the representation says there is an activity of writing followed by a process of creating which then terminates and yields y. In all three cases there does not appear to be an activity followed by a related process, i.e. two sequential events, and for this reason Van Valin (2005) proposed reformulating active accomplishments as in (18), which makes them in effect 'active achievements', which is the term used for them in Pavey (2010).

(18) Revised decompositions for active accomplishments (Van Valin 2005)

a. Motion to a goal: **do'**(x, [**run'**(x)]) & INGR **be-LOC'**(y,x)

b. Consumption: **do'**(x, [**eat'**(x,y)]) & INGR **consumed'**(y)

c. Creation: **do'**(x, [**write'**(x,y)]) & INGR **exist'**(y)

These representations express an activity which terminates and yields a result state, which seem to express what is going on better than the representations in (17).

However, Van Valin (2005) seems to have been overly hasty in rejecting the representations in (17), because the representations in (18) don't express the incremental nature of running or eating or writing (e.g. Krifka 1992; Filip 1993[1999]; Tenny 1994; Jackendoff 1996; Rothstein 2004), and this means that they cannot capture the 'measuring out' properties of what Dowty (1991) calls the 'incremental theme' and the related notion of 'incremental path' for motion predications. This can be captured in a variation on the representations in (17). One problem with them was that the process followed the activity, whereas in incremental events the process can be taken to be simultaneous with the activity. For example, the action of eating something is simultaneously a process of consumption, and this can be expressed by replacing '&' by '∧', which means 'and simultaneously' in the representation (cf. fn. 5). Moreover, the simultaneous actions constitute a single complex action, and so they should be grouped together; furthermore, the termination expressed by 'INGR predicate'' applies to both of them equally. The result for *devour* is given in (19).

(19) [**do'**(x, [**eat'**(x,y)]) ∧ PROC **consume'**(y)] & INGR **consumed'**(y)

This expresses 'x eats y, which simultaneously undergoes a process of consumption, both of which terminate and lead to the result that y is consumed'. For a creation verb, there would be analogous representation, as exemplified in (20) for *write (a poem)*.

(20) [**do'**(x, [**write'**(x,y)]) ∧ PROC **create'**(y)] & INGR **exist'**(y)

This says 'x writes y, which simultaneously undergoes a process of creation, which terminates and leads to the result that y exists'. For a motion verb like *run*, the distance covered acts as the incremental element, i.e. the incremental path, as expressed in (21).

(21) [**do'**(x, [**run'**(x)]) ∧ PROC **cover.path.distance'**(x,(y))] & INGR **be-at'**(z, x)

This means 'x runs and simultaneously effects a process of covering distance, which terminates, which leads to the result that x comes to be located at *z'*. In all three representations the incrementality of eating, writing, and running is expressed via the PROCess component, something that is missing in (18), and given that accomplishments contain PROC & INGR, the representations in (19)–(21) can be legitimately labeled 'active accomplishments', unlike those in (18).

A consequence of (18)–(20) is that incremental activities can no longer be represented simply as **do'**(x, [**run'**(x)]) or **do'**(x, [**eat'**(x, (y))]), since this does not capture their incremental nature. Rather, they must be represented as **do'**(x, [**run'**(x)]) ∧ PROC **cover.path.distance'**(x, (y)) and **do'**(x, [**eat'**(x, (y))]) ∧ PROC **consume'**(y), which contrasts with non-incremental activities such as *talk* or *shine*.[9]

The representation in (21) has an additional advantage. It has long been observed that motion predicates with a specific distance given are telic and are active accomplishments, e.g. *Sally ran two miles in 15 minutes*. Such sentences are not amenable to a traditional active accomplishment representation, i.e. **do'**(Sally, [**run'**(Sally)]) & BECOME/INGR **be-at'**(two miles, Sally). Regardless of whether this is made with BECOME or INGR, Sally does not come to be located 'at two miles', in contrast to canonical examples like *Sally ran to the park*. However, the representation in (21) can be used to express distance active accomplishments rather well, as illustrated in (22).

(22) [**do'**(Sally, [**run'**(Sally)]) ∧ PROC **cover.path.distance'**(Sally, two miles)] & INGR **be-at'**(**path.endpoint**, Sally)

The *y* variable refers to the distance covered, not the location of *x*, and the location of *x* is the endpoint of the distance covered, *z*. Moreover, it is possible to express both the distance and the goal, as in (23), which would have the representation in (24).

[9] The incrementality of a motion verb like *run* can be cancelled in a sentence like *John ran in place for an hour*, but the default interpretation of *run* involves covering some distance along a path. With a consumption verb like *eat* the incremental interpretation seems to depend on having a second argument, even a non-specific one like *pizza* in *John ate pizza*, as *John ate/was eating* seems to lack incrementality.

(23) Sally ran two miles to the park.

(24) [**do'**(Sally, [**run'**(Sally)]) ∧ PROC **cover.path.distance'**(Sally, two miles)] & INGR **be-at'**(park, Sally)

An interesting question arises: what is *two miles* in terms of the layered structure of the clause? Is it a core argument? an argument-adjunct? an adjunct? It fills an argument slot in the logical structure and is not oblique, which suggests it is not an adjunct or argument-adjunct. Is it an undergoer? It appears that it can be passivized and 'raised', as in (25), which suggests that analyzing it as an undergoer is plausible. Contrast this with the impossibility of passivizing and 'raising' superficially similar quantitative core constituents, as in (26). More work is needed on this issue.

(25) a. Two miles were run by Sally in only twenty minutes.

b. Two miles seem to have been run by Sally in only twenty minutes.

(26) a. * Five dollars were cost by the book.

b. * Five dollars seem to be cost by the book.

7 The syntactic structure of motion active accomplishments

The syntactic structure of motion active accomplishments has been assumed to be that of a core containing an RP (*John*), a verbal nucleus (*ran*) and a predicative PP (*into the room*) functioning as an argument-adjunct, as in Figure 1.

However, Van Valin (2005: 239) suggests that caused motion active accomplishments, e.g. *John pushed the chair into the room*, are nuclear junctures, with *push* and *into the room* being the linked nuclei, analogous to resultatives like *John painted the table white*. A PP as a nucleus is not unusual; cf. *John is in the room*, where *in the room* is the predicate in the nucleus, *be* being an auxiliary (Van Valin 2005: 14). This would be represented as in Figure 2.

The crucial component uniting motion and caused motion active accomplishments is the goal PP *into the room*, and therefore it is reasonable to assume that the PP should receive the same structural analysis in both forms. Is there any evidence in favor of the structure in Figure 1 or Figure 2? The scope of aspectual operators seems to provide the relevant evidence.

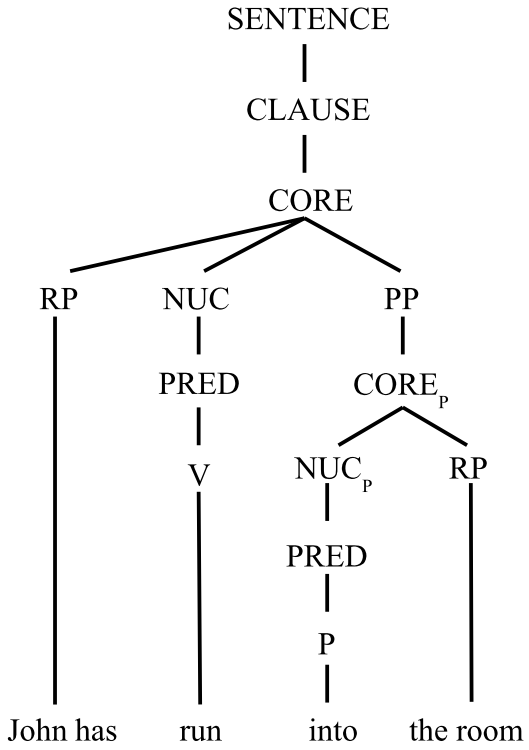


Fig. 1: Motion active accomplishment structure with argument-adjunct PP

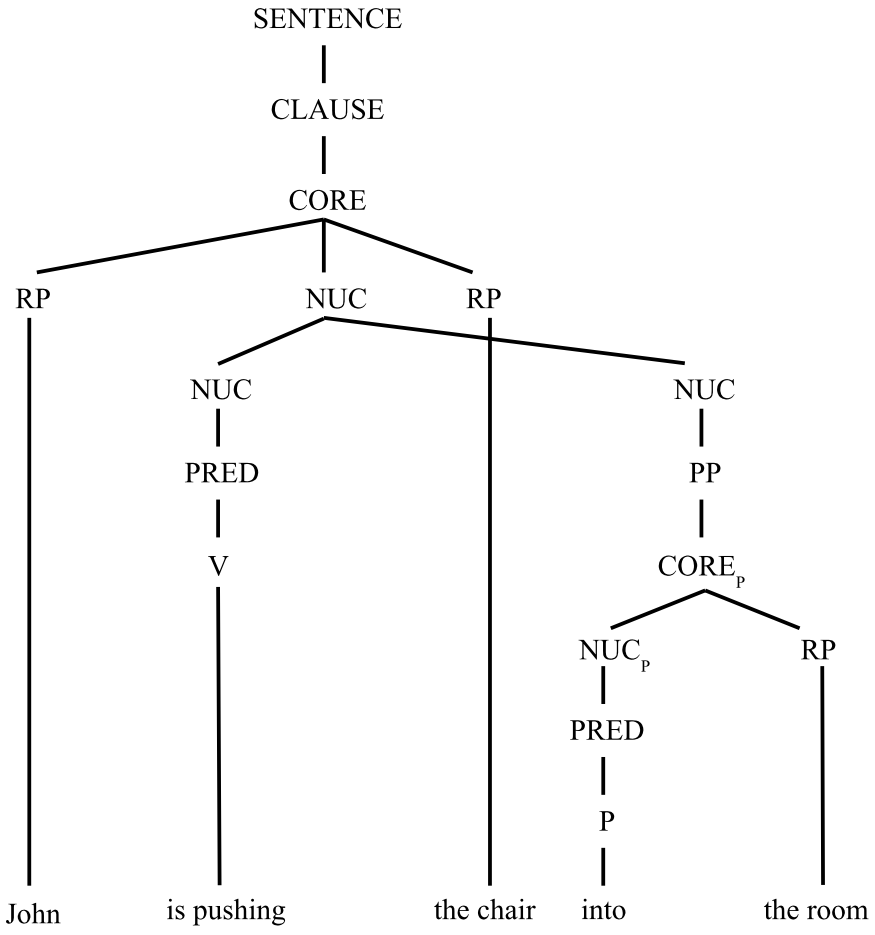


Fig. 2: Caused motion active accomplishment as nuclear juncture

(27) a. John has run into the room.

b. ⟨$_{\text{IF}}$ *DECL* ⟨$_{\text{TNS}}$ *PRES* ⟨$_{\text{ASP}}$ *PERF* [**do'**(John, [**run'**(John)])] $_{\text{ASP}}$⟩ & [INGR **be-in'** (room, John)]] $_{\text{TNS}}$⟩ $_{\text{IF}}$⟩ = mono-nuclear structure in Figure 3

c. ⟨$_{\text{IF}}$ *DECL* ⟨$_{\text{TNS}}$ *PRES* ⟨$_{\text{ASP}}$ *PERF* [**do'**(John, [**run'**(John)]) & [INGR **be-in'**(room, John)]] $_{\text{ASP}}$⟩ $_{\text{TNS}}$⟩ $_{\text{IF}}$⟩ = nuclear juncture structure in Figure 4

In the mono-nuclear argument-adjunct PP structure in Figure 3 and semantic representation in (27b), the aspect operator indicates that the running is finished but not that the transition into the room is completed; the transition is expressed by a second predication which lies outside the scope of the nuclear aspect operator, which would modify only the nucleus containing *run* in the operator projection. This is not the case in the nuclear cosubordination structure in Figure 4 and its semantic representation in (27c), in which the perfect aspect operator has scope over the running (nucleus-1) and the transition into the room (nucleus-2), both of which are completed. In the syntax, this would be represented by the aspect operator modifying the superordinate nucleus node in the operator projection, as in Figure 4. Thus, motion active accomplishments are best analyzed as nuclear junctures, specifically nuclear cosubordination.

In the past it was assumed in RRG that the logical structures for motion and non-motion active accomplishments were similar in form (cf. (8e), Table 1) and that they both were realized syntactically in simple cores with a simple nucleus. Now, however, such an analysis seems to be impossible: the difference between *Bill ate pizza* and *Bill ate a pizza* does not involve nuclear junctures; both of these sentences are mono-nuclear simple cores. The crucial difference between the motion and non-motion active accomplishment logical structures is that in the motion logical structure a new argument is introduced in the second predication, which is outside the scope of the aspect operator modifying the main verb in (27b). In a consumption or creation logical structure, on the other hand, all arguments are introduced by the main predicate and therefore would be within the scope of the aspect operator. Hence there appears to be no justification for a nuclear juncture analysis of non-motion active accomplishments. The detailed investigation of the implications of the nuclear juncture analysis of motion active accomplishments for the analysis of non-motion active accomplishments is beyond the scope of this paper and must be left for future research.

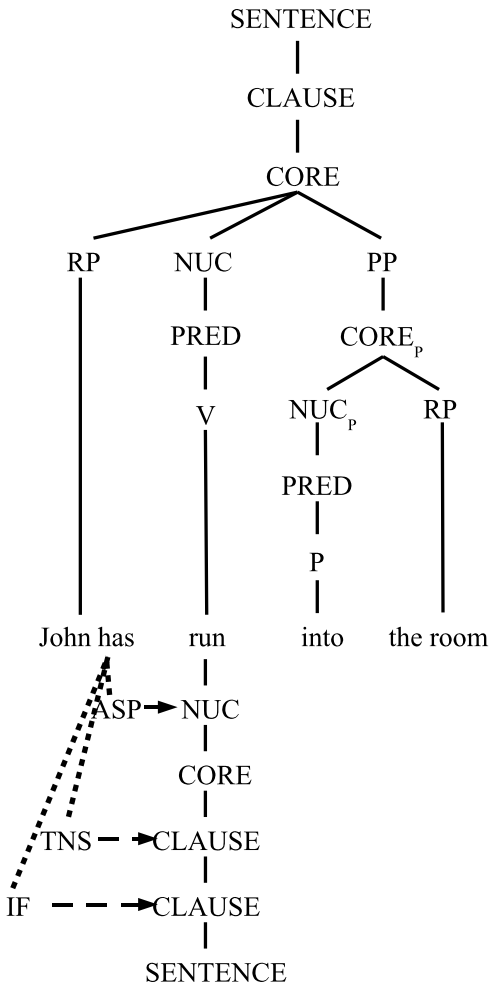


Fig. 3: Mono-nuclear structure for (27a)

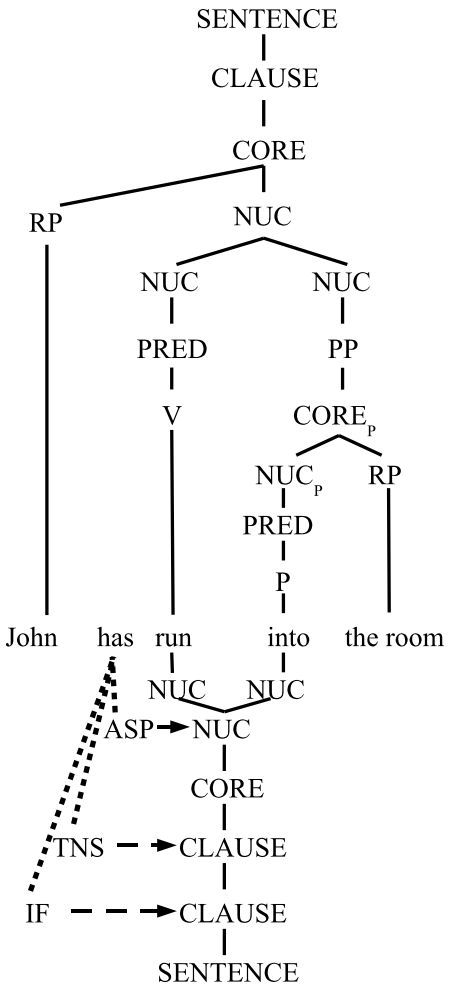


Fig. 4: Nuclear juncture structure for (27a)

8 Conclusion

This paper has looked at the treatment of accomplishment predications in a succession of theoretical frameworks, starting with Vendler's (1957[1967]) proposals. The original contrast between achievements and accomplishments has been supplanted by a more complex classification system that makes distinctions that better accommodate the facts discussed herein. This revised classification treats telic activities as a separate class, active accomplishments, whose properties were investigated. An important point is the decomposition of plain accomplishments into PROCess & INGR (achievement), as this allowed the system to capture the fact that some languages code pure processes without a termination and result state. The decomposition of BECOME into PROC & INGR led to an improved analysis of active accomplishment predications and to a solution to a long-standing problem. Furthermore, it was shown that the syntactic structure of motion active accomplishments is best captured as an instance of nuclear cosubordination, rather than as a mono-nuclear core with an argument-adjunct PP.

References

Aronson, Howard. 1982. *Georgian: A Reading Grammar*. Columbus, OH: Slavica.

Cortés-Rodríguez, Francisco J. 2014. Aspectual features in Role and Reference Grammar: A layered proposal. *Revista Española de Lingüística Aplicada* 27: 23–53.

Dixon, Robert M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: CUP.

Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

—. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.

Filip, Hana. 1993[1999]. *Aspect, Situation Types and Noun Phrase Semantics*. NewYork/London: Garland Publishing, Inc. [PhD Thesis 1993].

Foley, William & Van Valin, Robert Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: CUP.

Holisky, Dee A. 1981. Aspect theory and Georgian aspect. In *Syntax & Semantics 14: Tense and Aspect*, P. Tedeschi & A. Zaenen (eds), 127–144. New York: Academic Press.

—. 1987. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71: 103–32.

Jackendoff, Ray. 1996. The proper treatment of measuring out, telicity, and possibly even quantification in English. *Natural Language and Linguistic Theory* 14: 305–354.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In *Lexical Matters*, I. Sag & A. Szabolsci (eds), 29–54. Stanford: CSLI.

Latrouite, Anja & Van Valin, Robert D., Jr. 2014. Referentiality and telicity in Lakhota and Tagalog. In *Meaning and Grammar of Nouns and Verbs*, D. Gerland, C. Horn, A. Latrouite & A. Ortmann (eds), 401–26. Düsseldorf: Düsseldorf University Press.

Levin, Beth & Rappaport Hovav, Malka. 1995. *Unaccusativity at the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.

Pavey, Emma. 2010. *The Structure of Language: An Introduction to Grammatical Analysis*. Cambridge: CUP.

Rappaport Hovav, Malka & Levin, Beth. 1998. Building verb meanings. In *The Projection of Arguments: Lexical and Compositional factors*, M. Butt & W. Geuder (eds), 97–134. Stanford: CSLI.

Ross, John R. 1972. Act. In *Semantics of Natural Language*, D. Davidson & G. Harman (eds), 70–126. Dordrecht: Reidel.

Rothstein, Susan. 2004. *Structuring Events: A Study in the Semantics of Lexical Aspect.* Oxford: Blackwell.

Samsel, John. 1992. *A Role and Reference Grammar Analysis of Georgian Morphosyntax.* Unpublished MA Thesis, University of California, Davis. [Available on RRG website]

Tenny, Carol. 1994. *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Toratani, Kiyoko. 2002. *The morphosyntax and the logical structures of compound verbs in Japanese*. Unpublished PhD dissertation, University at Buffalo. [Available on RRG website]

Ullrich, Jan. 2008. *New Lakota Dictionary* (1st edition). Bloomington, IN: Lakota Language Consortium.

Van Valin, Robert D., Jr. 1990. Semantic parameters of split-intransitivity. *Language* 66(2): 221–260.

—. 1993. A synopsis of Role and Reference Grammar. In *Advances in Role and Reference Grammar*, R. Van Valin Jr. (ed.), 1–164. Amsterdam: John Benjamins.

—-. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Van Valin, Robert D., Jr. & LaPolla, Randy. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: CUP.

Van Valin, Robert D., Jr. & Wilkins, David P. 1996. The case for 'effector': Case roles, agents and agency revisited. In *Grammatical Constructions*, M. Shibatani & S. Thompson (eds), 289–322. Oxford: OUP.

Vendler, Zeno. 1957[1967]. Verbs and times. *The Philosophical Review* 66: 143–60; reprinted in *Linguistics in Philosophy*, Z. Vendler. 1967. Ithaca: Cornell University Press.

Verkuyl, Henk. 1993. *A Theory of Aspectuality: The Interaction between Temporal and Atemporal structure*. Cambridge: CUP.

Walton, Charles. 1986. *Sama Verbal Semantics: Classification, Derivation and Inflection*. Manila: Linguistic Society of the Philippines.

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

Aspect and *Aktionsart* in Slavic, Inflection and Derivation in RRG

Ranko Matasović
University of Zagreb

Abstract

Slavic languages are characterized by systematic pairings of verbs showing aspectual oppositions. Usually the perfective verbs are derived from imperfectives by prefixation, but the other processes, such as the derivation of secondary imperfectives and perfectives by suffixation, also exist. In this paper, it is argued that Slavic aspect needs to be represented chiefly in terms of derivation, as the system of lexical rules relating logical structures of verbs that form aspectual pairs, although it also shows some features usually associated with inflection (as a grammatical category expressed by nuclear operators in the operator projection).

1 How does Slavic aspect work?

The structure of this paper is as follows: I first briefly explain how the system of aspectual oppositions is expressed in Croatian. In Section 2, I discuss whether aspect is a derivational or inflectional category in Croatian, building upon Anelia Ignatova's (2008) RRG analysis of Bulgarian aspectual system. Section 3 is dedicated to the question how the difference between derivation and inflection should be captured in RRG. It is argued that lexical rules must be used in representation of derivational processes, but not inflectional processes. Finally, in Section 4 we discuss the types of rules that figure in the derivation of Croatian perfective verbs from imperfectives. We show that aspectual alternations involve the change of *Aktionsarten* in the lexical representation of verbs, and the current RRG system of representation of *Aktionsarten* turns out to be insufficient to capture all the *Aktionsart* types found in Croatian. This leads us to a proposal of a more economical system of features that define *Aktionsarten*, and to a theory that predicts the common cross-linguistic patterns of *Aktionsart* alternations. All examples in this paper are from Croatian, unless noted otherwise. The aspectual system of Croatian is similar to those of other Slavic languages in all the respects relevant to the argument presented

in this paper. There are two aspects: perfective and imperfective. Slavic perfective aspect expresses the event as a whole, without reference to its internal constituents:[1]

(1) Ivan je na-pisa-o knjig-u
I. AUX PFV-write-PPTC.M book-ACC.SG
'Ivan wrote a book'

Slavic imperfective aspect expresses the event with respect to its internal constituents:

(2) Ivan je pisa-o knjig-u
I. AUX write-PPTC.M book-ACC.SG
'Ivan wrote a book'

Verb stems are either perfective or imperfective, with a small number of bi-aspectual verbs, e.g. *vidjeti* 'see', *čuti* 'hear'. In Croatian, perfective verbs are not used in the present tense to express absolute present (the present tense of perfective verbs is a relative tense, and occurs mostly in subordinate clauses).

Perfective verbs are formed from imperfectives by prefixation, e.g. *pisati* (ipfv.) 'write' vs. *na-pisati* (pfv.) 'write'. There are some unprefixed perfectives; their imperfective counterparts are often formed by *Ablaut*, e.g. *skočiti* 'jump' (pfv.) vs. *skakati* 'jump' (ipfv.).

Perfectives are not formed exclusively by prefixation. The suffix *–nu-* forms perfective verbs with semelfactive meaning, e.g. *mahnuti* 'wave (once)', *viknuti* 'shout (once)'.

Secondary imperfectives can be formed from perfective verbs by suffixation (sometimes also by *Ablaut*), e.g. *prevesti* 'transport' (pfv.) vs. *prevažati* 'be transporting' (ipfv.), *donijeti* 'bring' (pfv.) vs. *donašati* 'be bringing' (ipfv.).

2 Is Slavic aspect inflectional or derivational?

The distinction between inflection and derivation is scalar, not absolute (Aikhenvald 2007), see Table 1.

[1] This standard definition of the functions and meanings of Slavic aspect is found, e.g., in Comrie 1976 and Maslov 2004.

	Inflection	Derivation	Slavic Aspect
1.	Usually obligatory	Optional	inflectional (?); but, note that there are bi-aspectual verbs
2.	Final process (if affix, on rim of grammatical relationships bet a word)	Pre-final process (if affix, between root and inflection)	derivational
3.	Forms a complete word	Derives a stem which takes inflections	derivational
4.	Defining characteristics of a word class (e.g. nouns inflect for case)	Usually specific to a word class	?
5.	Does not change word class	Either derives a stem of a different word class, or adds some semantic specification to a root without changing class	derivational
6.	May indicate grammatical relationship between words and/or participate in agreement	Never indicates grammatical relationships between words or participates in agreement	derivational
7.	Usually does not show gaps in the paradigm	Often shows gaps in the paradigm	? (no paradigm)
8.	Generally semantically regular	Often semantically irregular	derivational
9.	Tends to form smallish systems	May be large systems	derivational (the number of prefixes is large)
10.	Tends to have high frequency	Likely to have lower frequency	? (the frequency of individual morphemes is low)
11.	Tends to be monosyllabic	May be monosyllabic or longer	? (most prefixes are monosyllabic, but there are polysyllabic combinations of prefixes)

Table 1: Criteria for distinguishing inflection from derivation

Each of the distinguishing criteria adduced in Table 1 needs a few comments, when applied to Slavic aspect:

Ad 1) Aspect appears to be an obligatory category in Slavic, as all verbs are necessarily either perfective or imperfective. However, there is a small residue of the so-called "bi-aspectual" verbs that are inherently neither, e.g. *vidjeti* 'to see' in Croatian. Therefore, this criterion is not decisive.

Ad 2) Aspect is a pre-final process, as morphemes expressing it are closer to the root than morphemes expressing clearly inflectional categories such as person and number, cf. Croat. *mah-n-em* (wave-PERF.-1SG.PRS) 'I wave'.

Ad 3) Aspect morphemes generally do not form a complete word, as inflectional morphemes for person/number/tense must also be added, as in the example mentioned Ad 2).

Ad 4) Slavic aspect defines verbs as a word-class (along with other categories, such as person and number). Other word-classes do not have aspect.

Ad 5) Aspect does not change the word class, as morphemes marked for aspect are verbal roots which remain verbal. However, Aikhenvald does not consider this criterion decisive, as many derivational morphemes also do not change the word class.

Ad 6) Aspect may not involve grammatical relations or agreement, so in this sense it is clearly derivational. However, there are clearly inflectional categories that also do not involve grammatical relations (e.g. tense).

Ad 7) Except if the bi-aspectual verbs are considered as gaps (i.e. as having no inherent aspect), aspect does not show gaps in the paradigm, so in this sense it is inflectional.

Ad 8) Aspectual morphemes are often semantically irregular, in that the meaning of the verb derived by perfective prefix often cannot be predicted from the meaning of the verbal root and the meaning of the prefix. In this sense, aspect is derivational.

Ad 9) The Slavic aspectual system is small, in that there are only two aspects: perfective and imperfective. However, both within perfectives and imperfectives we can distinguish other sub-categories, and if they are taken into account, the system is quite large: for example, among perfective verbs in Croatian we find inchoatives, finitives, resultatives, etc. Thus, this criterion is not decisive.

Ad 10) Since all verbs are marked for aspect, the frequency of aspectual forms is high, and aspect is inflectional in this sense. But obviously the criterion 10) is not independent of other criteria such as 7) and 1) – all obligatory categories are bound to be frequent.

Ad 11) Apart from a few exceptions, all morphemes expressing aspect in Slavic are monosyllabic. However, this is not decisive, since derivational categories may also be expressed by short, monosyllabic morphemes (as Aikhenvald acknowledges). Only if a category is consistently expressed by relatively long morphemes is this an argument that it is derivational, but

the converse does not follow: if morphemes expressing it are short, the category may or may not be inflectional.

Ignatova (2008), working on Bulgarian, argues that Slavic aspectual prefixes are derivational, because many of them have lexical (besides purely aspectual) meaning. There are around 25 perfective-forming prefixes in Croatian, but only fifteen are frequent and productive.[2] They can express a variety of secondary verbal meanings, and many of the prefixes often have more than one meaning, so the exact meaning of the derived perfective cannot be simply predicted from the meanings of the constituent parts. We will give the examples only with the eleven most common prefixes:

1. *na-* derives sative verbs (*najesti se* 'eat one's fill' ← *jesti* 'eat'), superessives/superlatives (*nakapati* 'sprinkle onto' ← *kapati* 'sprinkle')

2. *od-* derives ablative verbs (*odagnati* 'chase away' ← *gnati* 'chase') and completives (*odsvirati* 'play (to the end)' ← *svirati* 'play (an instrument)')

3. *po-* derives distributives (*pobiti* 'kill (multiple undergoers)' ← *biti* 'strike', *poloviti* 'hunt', 'catch (multiple undergoers)' ← *loviti* 'hunt', 'chase'), inceptives (*poletjeti* 'start flying', 'take off' ← *letjeti* 'fly', *pojuriti* 'start running' ← *juriti* 'run'), diminutives (*poigrati se* 'play a little' ← *igrati se* 'play'), and it is also the default perfectivizer without any inherent meaning (*pogledati* 'take a look at' ← *gledati* 'watch').

4. *pod-* derives subteressive verbs (*podstaviti* 'place under' ← *staviti* 'place'), diminutives (*podnapiti se* 'get a little drunk' ← *napiti se* 'get drunk')

5. *pre-* derives translatives (*preploviti* 'sail across' ← *ploviti* 'sail'), repetitives (*preprodati* 're-sell' ← *prodati* 'sell') and excessives (*prepećі* 'over-bake' ← *pećі* 'bake', *preglad-njeti* 'become too hungry')

6. *pri-* derives apudessives/allatives (*privućі* 'draw close' ← *vućі* 'draw'), diminutives (*prileći* 'lay down a little' ← *leći* 'lay down'), and it can also function as a default perfectivizer (prisiliti 'force' (pfv.) ← *siliti* 'force' (ipfv.)).

7. *pro-* derives illatives (verbs denoting a piercing action), e.g. *provućі* 'draw through' ← *vućі* 'draw', *probosti* 'pierce through' ← *bosti* 'pierce', 'stab'), praeteressives (verbs denoting an action that is performed passing by an object (*proletjeti* 'fly by' ← letjeti

[2] Babić (1986: 477) claims there are 27 perfectivizing prefixes, of which 16 are productive. He included the prefixes *de-* and *re-*, as well as the prefix *dis-*, which occur only in loanwords and are of Latin and Greek origin respectively.

'fly'), inceptives (*procvasti* 'start blossoming' ← *cvasti* 'blossom'), and it is often a default perfectivizer (*proliti* 'pour' (pfv.) ← *liti* 'pour' (ipfv.), *proširiti* 'widen', 'spread' (pfv.) ← *širiti* 'spread' (pfv.)).

8. *raz-* derives verbs denoting an action aimed in several directions simultaneously (*razaslati* 'send in several directions' ← *slati* 'send', *razjuriti* 'chase in several directions' ← *juriti* 'run', 'chase'), intensives (*raspaliti* 'light up', 'burn energetically' ← *paliti* 'burn' (transitive)), privatives (*rasteretiti* 'remove the burden from' ← *teretiti* 'burden', *razuvjeriti* 'dissuade' ← *uvjeriti* 'persuade'), and it also serves as the default perfectivizer (*rashladiti* 'cool' (pfv.) ← *hladiti* 'cool' (ipfv.)).

9. *s-* derives verbs that denote centripetal, joining actions (*slijepiti* 'glue together' ← *lijepiti* 'glue', *skovati* 'hew together' ← *kovati* 'hew'), actions that denote removing from some surface (*sprati* 'wash away' ← *prati* 'wash', *sjahati* 'dismount' ← *jahati* 'ride (a horse)') and it also functions as a default perfectivizer with no lexical meaning (*slediti se* 'freeze' (pfv.) ← *lediti se* 'freeze' (ipfv.)).

10. *u-* derives illatives (*uliti* 'pour into' ← *liti* 'pour', *ugurati* 'push into' ← *gurati* 'push') and it can also have no lexical meaning (*ubrati* 'pick' (pfv.) ← *brati* 'pick' (ipfv.), *upitati* 'ask' (pfv.) ← *pitati* 'ask' (ipfv.)).

11. *uz-* is a directional denoting a vertical (top-oriented) action (*uzvesti* 'lead up' ← *vesti* 'lead') and it can also have no lexical meaning (*uzmoći* 'be able' (pfv.) ← *moći* 'be able', 'can').

12. *za-* derives inceptives/inchoatives (*zavoljeti* 'start loving' ← *voljeti* 'love'), verbs denoting immersing or covering actions (*zavući* 'draw into' ← *vući* 'draw', *zagaziti* 'wade into' ← *gaziti* 'wade'), and it can also be without lexical meaning (*zaoštriti* 'sharpen' ← *oštriti* 'sharpen').

Often different prefixes can be used to derive synonymous, or quasi-synonymous perfectives, e.g. *zaoštriti* and *naoštriti* 'sharpen', or *zapitati* and *upitati* 'ask'. There may be slight difference in meaning between such pairs of verbs, but these are mostly pragmatic, e.g. *zapitati* has a stronger force and expresses a slightly less polite way of asking than *upitati*.

As we have seen in the examples, the lexical meaning of aspectual prefixes is often directional. RRG's operator scope hierarchy predicts that directionals fused with aspectual prefixes will be nuclear directionals, i.e. those that modify the orientation of action or event without reference to participants (arguments), and this prediction is fully borne out by the facts:

slagati 'place' (ipfv.) vs. *naslagati* 'place on' (pfv.)
letjeti 'fly' (ipfv.) vs. *izletjeti* 'fly out of' (pfv.)
plivati 'swim' (ipfv.) vs. *uplivati* 'swim into' (pfv.)
trčati 'run' (ipfv.) vs. *protrčati* 'run past' (pfv.)

Some prefixes can express only directionality, without changing the aspect/*Aktionsart* of the verb:

skočiti 'jump' (pfv.) vs. *uskočiti* 'jump into' (pfv.)

Likewise, prefixes can only affect the aspect/*Aktionsart*, without having a directional value or any semantic content. The default such prefix in Croatian is *po-*:

piti 'drink' (ipfv.) vs. *popiti* 'drink' (pfv.)
jesti 'eat' (ipfv.) vs. *pojesti* 'eat' (pfv.)

Ignatova (2008) also argues that suffixes deriving secondary imperfectives are purely inflectional, rather than derivational, since they do not have any lexical meaning. However, even if we view them in isolation from the prefixes, they also show some derivational characteristics, e.g. they do not represent a final process, they do not form a complete word, but they do show gaps in the paradigm. Most importantly, they cannot be automatically formed from any perfective verb, e.g. there are no derived imperfectives **pojedavati*, **uplivavati* corresponding to perfectives *pojesti* 'eat', *uplivati* 'swim into'. Lastly, some imperfectivizing suffixes do add a component of meaning, e.g. frequentative:

pitati 'ask' (ipfv.) → *za-pitati* 'ask' (pfv.) → *za-pit-kivati* 'be asking (frequently)' (ipfv.)

Hence, we believe that Croatian aspect as a whole is best defined as a derivational, rather than inflectional category, though we do acknowledge that the difference between derivation and inflection is scalar rather than absolute (cf. also Bybee 1985: 87).

3 What is the difference between derivation and inflection in RRG?

Our initial hypothesis about the difference between inflection and derivation in RRG was that inflectional morphemes are expressed on the Operator projection, while derivational morphemes are not. But this is clearly not the case: some clearly inflectional categories are not represented on the operator projection (e.g. person markers in familiar Indo-European languages), and there is no theoretical principle in RRG stating that derivational morphemes should not be represented in the operator projection.

Arista (2008), following Everett (2002) argues that derivational morphemes are Nuclei, while inflectional morphemes are arguments. This has far-reaching consequences for their conception of the Layered Structure of the Word (parallel to the Layered Structure of the Clause), but is irrelevant for our present purposes. On the other hand, Arista (2008: 124) discusses the problem of *lexical integrity*: whether internal constituents of a word are visible to syntax. This is highly relevant, as it represents an independent criterion for distinguishing derivational from inflectional processes in word-formation. We take the view that inflectional processes are indeed invisible to syntax (i.e. syntactic rules do not refer to inflectional morphemes), while this does not necessarily apply to derivational morphemes and lexical rules. Nearly all of the word-formation processes adduced as exceptions to the Lexical Integrity Hypothesis involve derivation, rather than inflection (Lieber & Scalise 2007).

Derivational processes are handled by Lexical rules, cf., e.g., the rule relating (3a) to (3b): *pro-* + **do'**(x, [**pred$_1$'**(x, y)]) → **do'**(x,[**pred$_1$'**(x, y)]) & INGR **pred$_2$'**(y).

(3) a. Ivan je čitao knjig-u
I. AUX read-PPTC.M book-ACC.SG
'Ivan read a book'

b. Ivan je pro-čita-o knjig-u
I. AUX PFV-read-PPTC book-ACC.SG
'Ivan read a book'

Inflectional morphemes can be applied across the lexicon, and their function need not be specified in the lexicon by means of lexical rules. Therefore, the difference between derivation and inflection in RRG can be stated as follows: a morphological process is derivational if it must be represented in terms of lexical rules. If not, it is inflectional.

Morphemes expressing derivational categories are subject to the same scope restrictions in the Operator Projection as the inflectional morphemes. However, there is a separate restriction stating that derivational morphemes tend to be closer to the verbal root (Nucleus) than inflectional morphemes.

If both derivational and inflectional affixes are represented as operators, the order of derivational and inflectional affixes can interfere with the scope and order of the operators (which is predicted to be universal by RRG). Thus, in Tepehua (Totonacan) the incorporated morpheme *-alhi-*, 'always', is a core operator, but it occurs closer to the verbal root than the morpheme expressing the perfective aspect *–ta*, which is a nuclear operator (4):

(4) xix-'alhi-ta
get.dry-CONT-PFV
'S/he is always dry' (Watters 2009: 263)

In Bulgarian, we have an interesting case where (derivationally) imperfective verbs can be in the perfective aspect (aorist), and vice versa, (derivationally) perfective verbs can be in the imperfective aspect (imperfect), as in (5). The inflectional morpheme fusing person/number and aspect always follows the derivational morpheme expressing – essentially – the *Aktionsart* of the verb (see Lindstedt 1985).

(5) Štom puk-ne-še zora-ta, izkarva-x- ovc-i-te navən
As.soon.as break-PFV-IPFV.3SG. dawn-DEF.F drive-IPFV.1SG sheep-PL-DEF out
'As soon as dawn broke, I used to drive the sheep out' (Beaulieux & Mladenov 1950: 335)

Here the imperfect ending of the first verb superimposes upon it the habitual reading, although the meaning of the verb is inherently perfective (it is an achievement verb, which cannot be imperfective). A similar process exists in Georgian (Comrie 1976: 32). It is obvious that the fused person/number/tense/aspect morpheme has a wider scope than the morpheme expressing the *Aktionsart* of the verb. As predicted, in example (5), the inflectional morpheme with wider scope (*-še-*) is farther from the Nucleus than the derivational morpheme (*-ne-*).

4 What types of lexical rules express Slavic aspectual derivations?

We saw above that Slavic aspectual systems are characterized by systematic pairings of verbs showing the opposition between perfectives and imperfectives. Perfectives in Croatian generally belong to the following *Aktionsarten*:

1. Semelfactives: *bljesnuti* 'flash', *mahnuti* 'wave', *zalepršati* 'flutter'
2. Achievements: *razbiti* se 'shatter', 'break', *pročitati* 'read'
3. Accomplishments: *otopiti se* 'melt', *naučiti* 'learn'
4. Active accomplishments: *pojesti* 'eat', *popiti* 'drink', *dosegnuti* 'reach'

Imperfectives belong to the following *Aktionsarten*:

1. States: *spavati* 'sleep', *sjediti* 'sit'
2. Activities: *plesati* 'dance', *skakati* 'jump'

In some cases, the tests applied to establish the *Aktionsart* of a particular verb do not seem to work for Slavic, cf. e.g. the pair *skočiti* 'jump' (pfv.) *skakati* 'jump' (ipfv.). *Skakati* is certainly an activity: it can be used with a dynamic adverb.

(6) Snažno je skakao (vigorously/AUX/jumped) ‘he jumped vigorously’

It can be used with adverbs expressing duration:

(7) Brzo je skakao (quickly/AUX/jumped) ‘he jumped quickly’

With the *for*-PP phrase:

(8) Skakao je tri sata (jumped/AUX/three/hours) ‘he jumped for three hours’

In these examples the verb is clearly atelic and does not imply a result state. Its perfective counterpart, *skočiti*, passes nearly all of the tests that *skakati* does (*Snažno je skočio*, *Brzo je skočio*), but it is not durative (**Skočio je tri sata*). Clearly, then, it cannot be simply an activity verb. A similar argument can be produced with respect to a verb such as *zalepršati* ‘flutter’ (pfv.), which is perfectly compatible with adverbs such as *snažno* ‘strongly’, *brzo* ‘quickly’, but it is atelic (there is no result state) and it is non-durative (**Zalepršao je tri sata*).

Moreover, the current system of *Aktionsarten* does not provide for Croatian verbs such as *sagnuti se* ‘bend down’, ‘bow’, which is the perfective counterpart of *sagibati se* ‘bend down’, ‘bow’; while the latter passes all the tests for a State predicate, the former can be used with dynamic adverbs (*Snažno se sagnuo* ‘he forcefully bent down’). If this is viewed as a causative verb (‘to bend down’ = ‘to cause to be bent’, which is not the most obvious analysis in my opinion),[3] then some other verbs of change of posture can be considered better examples of this *Aktionsart*, e.g. *ustati* ‘to stand up’; these are telic (they can be construed with *in*-PP phrases), dynamic (they can be construed with *energično* ‘forcefully’, ‘energetically’), and they are punctual (they cannot be construed with for PP, e.g. *tijekom tri sata* ‘for three hours’).

It could be claimed that this is simply an inchoative/inceptive of a stative verb, and that we should deal with it by introducing an operator (BEGIN) that can be added to the logical structure of a basic verb (in this case a state predicate). However, the verb *ustati* clearly does not have an inceptive meaning ‘to begin to stand’; rather, it means ‘to perform an action that leads to one’s being in an upright position’. Inchoatives/inceptives are often perfective versions of stative imperfectives, e.g. *zaspati* ‘fall asleep’ (pfv.) from *spavati* ‘sleep’, or of imperfective activity verbs, e.g. *zapjevati* (pfv.) ‘start to sing’ vs. *pjevati* (ipfv.) ‘sing’ – but

[3] The current RRG system of *Aktionsarten* allows us to analyze any transitive verb referring to a physical action as basically causative, and we do not have clear criteria for distinguishing transitive activity verbs and causative modifications of activity verbs. Hence, *žvakati* ‘chew’ can be viewed as an activity, but also as a causative (informally ‘cause to become chewed’), and *saviti* ‘bend’ as ‘cause to become bent’. However, I would prefer to treat as causatives only those verbs that lexicalize the notion of *external* force, i.e. force not inherent in the actor. (Non-causative) Activity verbs lexicalize the notion of *internal* force.

the inchoative/ingressive prefix is generally *za-*, and these verbs do not allow secondary imperfectivization (i.e. there is no **zapjevavati* 'be starting to sing', **zaspavavati* 'be falling asleep' parallel to *ustajati* 'be standing up').

We could argue that some of the tests for *Aktionsarten* used in RRG (Van Valin 2005: 39) are not relevant in Croatian, but this would be ad hoc; rather, on the basis of Croatian data, we should consider the possibility that a few more *Aktionsarten* should be posited. Moreover, it would be nice to have a system of features that define *Aktionsarten* from which the possibility of existence of these *Aktionsarten* would follow. And indeed, such a system exists, but it is not the current RRG system (Van Valin 2005: 33):

1. State: [+static], [-dynamic], [-telic], [-punctual]
2. Activity: [-static], [+dynamic], [-telic], [-punctual]
3. Achievement: [-static], [-dynamic], [+telic], [+punctual]
4. Semelfactive: [-static], [+/-dynamic], [-telic], [+punctual]
5. Accomplishment: [-static], [-dynamic], [+telic], [-punctual]
6. Active accomplishment: [-static], [+dynamic], [+telic], [-punctual]

The causative operator (CAUSE) can be used to derive the causative versions of each basic *Aktionsart*, and the operator DO' to derive agentive versions of activity verbs (e.g. agentive *murder* vs. non-agentive *kill*, etc.).

The current RRG system allows for the possibility of other combinations of features, but some apparently never co-exist (e.g. +static and +dynamic). Does the theory *claim* that there are no verbs in any languages that have a logical structure combining, e.g. [+dynamic], [+telic], and [+punctual]? Some combinations of features are certainly going to be rare cross-linguistically, as certain types of events are rare in the human experience, e.g. there is only a handful of semelfactive activities lexicalized as the Croatian verb *zalepršati* 'flutter', and there is only a handful of semelfactive accomplishments denoting punctual changes in body posture (e.g. *sagnuti se* 'bend down'), but we what we want is a system of lexical representation in which the existence of such lexicalizations would be predicted.

Moreover, the current system of features does not help us understand why some lexical rules appear to be more common, cross-linguistically, than other theoretically possible rules. Van Valin (2005: 41) notes that the pattern deriving accomplishments from states (or vice versa) is very common cross-linguistically, but does not say why. Nearly all direct derivations of aspectual pairs in Croatian involve one of the two types of rule: either a State verb is related to an Accomplishment (as in 9a vs. 9b), or an Activity verb to an Active accomplishment (as in 9a vs. 9b):

(9) a. Led se topio
ice melted.IPFV
'Ice was melting' (State)

b. Led se otopio
ice melted.PFV
'Ice melted' (Accomplishment)

In terms of lexical rules this derivation can be expressed as **pred'**(x) → BECOME **pred'**(x).

Of course, 'to melt' is a Process, not a State in a strict sense, but Processes can be viewed as a sub-type of stative predicates – the type in which temporal sections of the state are not identical. Processes having result states are Accomplishments in the RRG system (they are derived from basic static predicates with the operator BECOME).[4]

(10) a. Ivan je jeo jabuku
Ivan ate.IPFV apple.ACC.SG
'Ivan ate an apple' (Activity)

b. Ivan je pojeo jabuku
Ivan ate.PFV apple.ACC.SG
'Ivan ate the apple' (Active accomplishment)

In terms of lexical rules this derivation can be expressed as **do'**(x) **pred'**(x, y) → **do'**(x) **pred'**(x, y) & INGR **pred'**(y).

We would like to posit a system of features in which the frequency of this pattern would be intuitively clear, and also a system in which the rareness of certain kinds of lexical rules (e.g. those relating Activities to Semelfactives) would be predictable. In order to establish a clear relation with the system of Van Valin 2005 we will use the same three semantic features: [±punctual], [±dynamic] and [±telic]. These features correspond to three important and probably innate psychological concepts: *duration* (the perception of time), *(internal) force* (as opposed to external force, or causation), and *telicity* (having an internal end, or purpose), respectively. The relevance of these concepts for cognitive processes is well-established in cognitive psychology,[5] and we would not find it surprising if categories playing a large role in other cognitive systems were also involved as semantic primitives in lexical semantics of verbs. The primitive semantic features and their combinations are represented in the Table 2.

[4] On processes as basic *Aktionsart* see Van Valin (2005: 43).

[5] On the role of force dynamics and conceptualization of time in lexical semantics see especially Talmy (2000).

	punctual (non-durative)	dynamic (= having internal force)	telic	e.g.
Semelfactives	+	–	–	*bljesnuti* 'flash'
Achievements	+	–	+	*puknuti* 'pop'
States	–	–	–	*sjediti* 'sit'
Accomplishments	–	–	+	*otopiti se* 'melt'
Semelfactive activities	+	+	–	*skočiti* 'jump', *zalepršati* 'flutter'
Activities	–	+	–	*pjevati* 'sing'
Active accomplishments	–	+	+	*pojesti* 'eat'
Semelfactive active accomplishments	+	+	+	*sagnuti se* 'bend down', *ustati* 'stand up'

Table 2: The features defining *Aktionsarten*

Moreover, we suggest that the system of *Aktionsarten* is structured, in that the features are organized hierarchically. This system can be represented as the Feature Hierarchy (Figure 1). Arrows on the Hierarchy show the adjacent *Aktionsarten* (dominated by a single node) which are likely to be related by means of lexical rules. That we do not often find lexical rules relating Semelfactives (in all three varieties) to other *Aktionsarten* should be attributed to the fact that Semelfactives are rarely lexicalized in languages, as punctual events do not often occur in everyday experience.

The system of features defining *Aktionsarten* proposed here is more intuitive and economical than the current RRG system:

- it does not posit both features Static and Dynamic; since they never co-occur in the lexical representation of verbs, having both of them is redundant.
- it does not require any *Aktionsart* to be characterized with the value [±] for any feature (the RRG system characterizes Semelfactives as [± dynamic]).
- it makes strong empirical predictions: 1) it predicts the existence of two *Aktionsarten* that the RRG system does not distinguish (Semelfactive Activities and Semelfactive Active Accomplishments); 2) it predicts that only those *Aktionsarten* that are adjacent on the Feature Hierarchy will frequently be derived from each other by means of lexical rules.

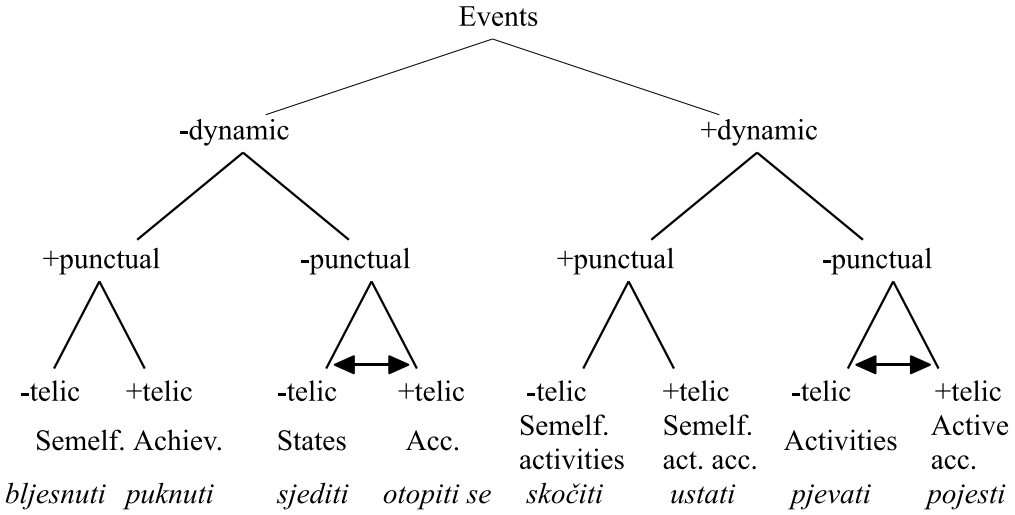


Fig. 1: The feature hierarchy

The system of lexical rules does not have to be economical. Grammars are messy. But our prediction is that it will tend to be economical, i.e. that, statistically, we will find that rules involving adding or subtracting only one feature are common cross-linguistically. This prediction has yet to be tested empirically in future investigations.

References

Aikhenvald, Alexandra. 2007. Typological distinctions in word-formation. In *Language Typology and Syntactic Description, III: Grammatical Categories and the Lexicon*, T. Shopen (ed.), 1–65. Cambridge: CUP.

Arista, Javier Martín. 2008. Unification and separation in a functional theory of morphology. In *Investigations of the Syntax-Semantics-Pragmatics Interface*, R. D. Van Valin Jr. (ed.), 119–147. Amsterdam: John Benjamins.

Babić, Stjepan. 1986. *Tvorba riječi u hrvatskom književnom jeziku* [Word Formation in Literary Croatian]. Zagreb: Hrvatska akademija znanosti i umjetnosti.

Beaulieux, Léon & Mladenov, Stefan. 1950. *Grammaire de la langue bulgare*. Paris: Institut d'Études slaves.

Bybee, Joan L. 1985. *Morphology: a study of the relation between meaning and form*. Amsterdam: John Benjamins.

Comrie, Bernard. 1976. *Aspect*. Cambridge: CUP.

Everett, Daniel. 2002. Towards a RRG theory of morphology, lecture delivered at the 2002 International Conference on Role and Reference Grammar in Logroño.

Ignatova, Anelia S. 2008. *A Functional approach to Bulgarian verbal aspect and reduplication of clitics within the framework of Role and Reference Grammar*, Ph.D. Thesis, Madrid: Universidad National de Educación a Distancia. [Available on RRG website]

Lieber, Rochelle & Scalise, Sergio. 2007. The Lexical Integrity Hypothesis in a New Theoretical Universe. In *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (*MMM5*) Fréjus 15–18 September 2005, University of Bologna*, G. Booij *et al.* (eds). https://geertbooij.files.wordpress.com/2014/02/mmm5-proceedings.pdf

Lindstedt, Jouko. 1985. *On the semantics of tense and aspect in Bulgarian*. Helsinki: Slavica Helsingensia.

Maslov, Juriy S. 2004. *Izabrannye trudy. Aspektologija, obščee jazykoznanie*. Moscow: Jazyki slavjanskoj kul'tury.

Talmy, Leonard. 2000. *Toward a cognitive semantics, Vol. 1: Concept structuring systems. Language, speech, and communication*. Cambridge, MA: MIT Press.

Van Valin Robert D., Jr. & LaPolla, Randy. 1997. *Syntax. Structure, Meaning, and Function*. Cambridge: CUP.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Watters, James K. 2009. Tepehua verb morphology, operator scope, and the encoding of arguments. In *Studies in Role and Reference Grammar*, L. Guerrero *et al.* (eds), 247–268. Ciudad de Mexico: UNAM.

Spanish *Llover* in RRG

Lisann Künkel
Albert-Ludwigs-Universität Freiburg

Abstract

The aim of this paper is to offer an RRG description of the Spanish verb *llover*, 'to rain'. We first determine the *Aktionsart* and argue that *llover* is an atypical activity verb. We then consider indications of place and time, because they occur frequently with *llover* without being attributed argument status. In the last section we apply the linking algorithm to four sentences, in which *llover* is used, to see if it makes correct predictions. We argue that, all in all, RRG is a good framework to describe *llover*, although the determination of the *Aktionsart* is not quite unequivocal and the linking algorithm has to be adapted to Spanish.

1 Introduction

The aim of this paper is to offer an RRG description of the Spanish verb *llover*, 'to rain'. *Llover* belongs to the group of so-called *verbs of precipitation*, which are a subgroup of *meteorological* or *weather verbs*. It is classified as an impersonal verb which can exceptionally be used in intransitive constructions too. The actual use is more flexible though: the constructions attested in the corpora are presented in (1).[1]

(1) *Constructions of Spanish* llover

a. las nubes llovían al revés (novel, 1990)
the clouds rain-3PL.PST at.the upside.down
'the clouds rained upside down'

b. de noche llovía hollín (novel, 1986)
at night rained-3SG.PST soot
'it rained soot at night'

[1] The data referred to in this paper are taken from CREA, the *Corpus de Referencia del Español Actual* of the *Real Academia Española* (www.corpus.rae.es). CREA contains data from various genres and its material represents both written and spoken language from 1975 to 2004.

c. el cielo llueve todas las desgracias juntas sobre nosotros (theatre, 1992)
the sky rain-3SG all the bad.luck together on us
'the sky is raining all the bad luck on us'

d. una tarde nos llovió encima (novel, 1994)
one evening us.DAT rained-3SG on
'one evening it was raining on us'

(1a) shows a sentence in which *llover*'s only argument describes the origin of the precipitation, whereas in (1b) the only argument refers to the precipitation. In (1c) both origin and precipitation are realized and in (1d) neither of them are. (1d) represents the most common use and (1c) the least common. All in all, the constructions vary with respect to the number of arguments, the syntactic function of the arguments, the meaning of the arguments and the interpretation of *llover* as a literal or metaphorical verb. If *llover* is used with one argument, that argument can refer to the precipitation, to its origin or to the place or person hit by the precipitation. An overview of the constructions' distribution is given in Table 1.

	origin only	both	precipitation only	neither	total
abs. nr.	44	17	451	2230	2742
rel. nr.	1.60%	0.62%	16.45%	81.33%	100%

Table 1: Quantitative analysis of *llover*

This paper wants to account for the *Aktionsart*, the thematic relations and macroroles, the syntactic functions and indications of place and time of *llover*. The description is based on the RRG framework. Impersonal verbs pose a challenge for linguistic theories and *llover* even more so, because it is used in different constructions. We argue that RRG nonetheless offers a good description of the facts presented by *llover*.

2 The *Aktionsart*

The aim of this section is to determine the *Aktionsart* of *llover*. To do so, we apply the tests proposed in Van Valin (2005). There are two things to keep in mind: one is the event structure, because it will be referred to in the discussion, and the other is that there may be reasons to consider the different constructions one by one.

The event structure of any precipitation event is presented in (2).

(2) *The event structure of* llover

a. begin, when first drop comes off

b. a different drops comes off, falls and lands

c. end, when last drop lands

The event structure is independent of the number or kind of argument expressed, because the arguments do not alter what is happening, but just specify who is participating in the event. Which *Aktionsart* a verb belongs to depends partly on the number and kind of arguments a verb is used with: Telicity, for example, can change according to the presence or absence of a definite argument: *Tom eats pizza* is atelic, whereas *Tom eats a pizza* is telic with *(a) pizza* being the argument. As *llover* is used in a variety of constructions, there are reasons then, to examine each type of construction by itself instead of looking at all of them at once. Such a split does not seem to be necessary, though, because the different constructions hardly ever have different test results. All occurrences of *llover* attested in CREA will therefore be examined together and exceptions highlighted. In RRG the *Aktionsart* is determined according to four semantic criteria and seven tests which partially refer to the semantic criteria. The results indicate which *Aktionsart* a verb belongs to. To determine *llover*'s *Aktionsart* we will therefore apply the tests and discuss the semantics in the following subsections.

2.1 The feature [±static]

Test 1 serves to examine if a verb can express progressive aspect. In English and Spanish the progressive aspect is expressed by the gerund and *gerundio*, respectively. The corpus contains 2742 tokens for *llover*. In 275, i.e. 10.04% of them, *llover* is used in the *gerundio*. *Llover* is thus considered to be [–static]. What stands out is that the *gerundio* seems to be incompatible with the argument referring to the origin of the precipitation: There is no token in which both the precipitation and the *gerundio* are used. As we cannot think of any semantic reason and as there seems to be no syntactic one,[2] we still conclude that, according to the first test, *llover* is [–static] and the answer thus "yes".

The semantic evaluation of the feature [±static] is based on the question *What is happening?* If the verb to be examined can serve as an answer to this question, it is considered [–static], because it then describes that something is happening and if something is happening, the verb cannot describe a state and be static. The application of this semantic test to *llover* is presented in (3). It is based on the translation of *What is happening?* to Spanish. It does not make any difference if the verb is in the present, future or past tense.

(3) A ¿Qué pasará/pasa/pasaba/está pasando? ('What will happen/happens/happened/is happening?')

B_1 las nubes llovían al revés ('the clouds rained upside down')

B_2 lloverá dinero ('it will rain money')

[2] In the grammar presented by the *Real Academia Española* (RAE), Yllera presents conditions for the use of the *gerundio* (1999: 3393). Neither of them is violated by the origin argument.

B_3 el cielo llueve todas las desgracias juntas sobre nosotros (‘the sky is raining all the bad luck on us’)

B_4 llovió débil o moderadamente en puntos de Galicia (‘it rained lightly or moderately in some places in Galicia’)

Example (3) shows that *llover* constructions can serve as an answer. *Llover* is therefore considered to be [–static] not only from a syntactic but also from a semantic point of view.

2.2 The feature [±dynamic]

Test 2 refers to the feature [±dynamic]. It is tested by coocurrence of the verb with a specified set of adverbs or prepositional phrases. If a verb is used with the expressions, it is [+dynamic], because the expressions can only describe the dynamics, if the verb has dynamics. In order to apply the test to Spanish, the expressions are translated from English to Spanish. The English and Spanish expressions are presented in (4) and (5).

(4) *English expressions of dynamic action* (Van Valin 2005: 33, 36)
actively, energetically, forcefully, strongly, vigorously, violently, in a vigorous way, with great force

(5) *Spanish expressions of dynamic action*
activamente, bruscamente, dulcemente, duramente, enérgicamente, firmemente, fuertemente, ligeramente, profundamente, rotundamente, seriamente, suavemente, vigorosamente, violentamente
a toda vela, con cuidado, con delicadeza, con dulzura, con dureza, con energía, con fuerza, con potencia, con vigor, con violencia, de manera violenta

In order to examine *llover*’s dynamics, it is necessary to know what *dynamics* means. *Dynamics* is the study of forces and their effect on the motion of an object. Dessì Schmid (2014: 20) defines a dynamic verb as one where phases of different quality can be distinguished. The sentences in (6) must, accordingly, be considered to have dynamic verbs.

(6) *Phases of different quality (with* llover*)*

a. y esa misma mañana, en que llovía a cántaros (novel, 1977)
and this same morning, in which rained.3SG buckets
‘and this very morning, when it was raining buckets’

b. eran las 3 de la mañana y llovía con intensidad (press, 2004)
were-3PL the 3 of the morning and rained-3SG with intensitiy
‘it was three o’clock in the morning and it was raining intensely’

c. cuando volvimos al teatro llovía ligeramente (novel, 1989)
when came back-1PL to the theatre rained-3SG lightly
'when we came back to the theatre, it was raining lightly'

d. sobre él seguían lloviendo abrumadoramente las pruebas de . . . (novel, 1990)
on him kept-3PL raining overwhelmingly the proofs of . . .
'the proofs of . . . kept raining on him overwhelmingly'

The expressions used in (6) are *a cántaros*, 'buckets', *con intensidad*, 'with intensity', *ligeramente*, 'lightly', and *abrumadoramente*, 'overwhelmingly'. They do not refer to any kind of force, but rather describe the rain's intensity, where *intensity* is considered to be closely related to density. Spanish expressions of intensity are presented in (7).

(7) *Spanish expressions of intensity*
a cántaros, a chaparrón, a chorros, a chuzo(s), a mares, a puñados, a raudales, con ganas, con intensidad, con menudo ahínco, con una densidad anormal, copiosamente, de forma aislada, de forma apreciable, de forma generosa, de manera torrencial, en abundancia, en cataratas,
abundantemente, excesivamente, finito/a, fino, intensamente, largamente, largo, torrencialmente, tupido

The sentences in (6) show that expressions of intensity are attested with *llover*. For the evaluation of the test, there are two options: Intensity can either be considered to be a part of dynamics or to be a distinct concept. In the first case, both kinds of expressions are relevant to the test, whereas in the second case, only expressions of dynamic action are considered. Physics and the English expressions proposed by Van Valin (2005: 33, 36) lead to a preference of the second option. There are several arguments in favour of the first option, though. One is that intensity is not captured in any other test. Another one, that the force that has an effect on the precipitation is constant: The precipitation is falling towards the earth because of gravitation. Strictly speaking, there is no difference in dynamics. The examples show, however, that differences can be perceived. The last argument is that adverbs of dynamics and of intensity have roughly the same interpretation with *llover*. This is true for expressions denoting strength as well as for those denoting weakness, as shown in (8).

(8) *Dynamics ≈ intensity (with* llover*)*

a. llueve fuerte ('it is raining strongly') ≈ llueve intensamente/a cántaros/a chorros/. . . ('it is raining intensely/from buckets/in streams')

b. llueve ligeramente ('it is raining lightly') ≈ llueve finito/fino ('it is raining finely')

Fuerte and *ligeramente* are considered expressions of dynamic action according to (5), whereas *intensamente*, *a cántaros*, *a chorros*, *finito*, and *fino* refer to the intensity. Nonetheless, there is hardly any difference in meaning, so dynamics and intensity can mean the same thing in the context of *llover* and both kinds of expressions are considered in the test.

The examination of the corpus then shows that 3.92% of *llover* tokens have an expression of dynamics or intensity. Test 2 is, consequently, answered with "yes" and *llover* is considered to be [+dynamic].

2.3 The feature [±telic]

The third feature is [±telic]. It is considered in tests 4 and 5. From a semantic perspective, it can be examined using three criteria that have been established by Dowty (1977, 1979). According to the first one, a verb is telic if it codes two different states: one at the beginning and one at the end of the event. The second criterion claims that a verb is telic if it has proper subevents, i.e. if it refers to a concrete instead of a generic event. The third criterion is Dowty's imperfective paradox which states that a verb is telic if "the entailment from the progressive tenses (also known as *continuous* tenses) to the simple tenses fails" (1977: 45), i.e. if the beginning of the event does not imply its ending.[3] All of these criteria are semantic, so they can also be applied to Spanish.

Llover does not code two different states because none of the entities mentioned undergo a change of state; neither the precipitation nor its origin. *Llover* is therefore atelic according to the first criterion. *Llover* is also classified as atelic by the second criterion: Precipitation events do not have proper subevents. This is obvious if the precipitation is not mentioned and if it is indefinite, because *llover* then behaves as other verbs like *eat* do. In the case of *llover*, there are not even any proper subevents, when the precipitation is definite as in (1c) (*todas las desgracias*) though. This is due to the event structure established in section 1. It claims that the precipitation event consists of subevents that take part almost simultaneously. The precipitation event is made up of their sum. As the single drops fall simultaneously and their falling cannot be interrupted, the precipitation event cannot be interrupted either. *Llover* therefore differs from other verbs like *eat* in that it is also classified as atelic by the second criterion. With respect to the imperfective paradox, *llover* is classified as atelic too. This is due to the fact that *estaba lloviendo* does imply *llovió* and is true for all constructions, because the implication is independent of the arguments. The reason lies, once again, in the event structure. All in all, the semantics of *llover* show that it is an atelic verb.

[3] The use of the progressive form cannot be a criterion for telicity (2005: 34), although it is considered in the imperfective paradox.

In English, the semantic properties of telic and atelic verbs are summarized by their compatibility with one of the prepositions *in* and *for*. The prepositions are combined with indications of time like *an hour* and a verb is tested for its compatibility with the two prepositional phrases. Telic verbs can be used with an *in*-PP and atelic ones with a *for*-PP. To adapt the test to Spanish, it is necessary to translate the two prepositions and to make sure that the Spanish prepositions have the same meanings. The translations can be found in (9) and (10).[4] Examples in which *llover* is used with the expressions are shown in (11).

(9) *Spanish* for ([Collins]: *for*)
future/past duration: ∅, *durante* or *por*
with English perfect tenses: *hace . . . que* + present for actions that are still going on; present + *desde hace* + time span
with past tenses: *hacía . . . que* + imperfect for actions that were going on; imperfect + *desde hacía* + time span[5]

(10) *Spanish* in ([Collins]: *in*; [Pons]: *in*)
in=during: *en*
(in=for: *hace . . . que*, *hacía . . . que*)
in=in the space of: *en*
(in=within: *dentro de*, *en*)
(in=at a later time: *dentro de*, *en*)
in=in less than: *en*

(11) Llover *with indications of time*

a. las multas le llueven diariamente desde hace años (press, 1995)
the fines him.DAT rain-3PL.PRES daily for years
'the fines have been raining daily on him for years'

b. y se solidificaba en un vapor sombrío, que llovía después
and REFL solidified in a vapor gloomy, which rained-3SG

durante días y noches sobre las ciudades (novel, 1977)
afterwards during days and nights on the cities

'and it solidified itself in a gloomy vapor, which afterwards rained on the cities for days and nights'

[4] Not all of the translations are adequate. *Hace . . . que* and *hacía . . . que* figure twice. They are only used as equivalents of *for*. *Dentro de* and sometimes also *en* are inadequate, when they correspond to *within* or to *at a later time*, because they do then refer to the beginning of the event instead of its duration. *En* is only a valid translation, if it refers to *during*, *in the space of* or *in less than*.

[5] Dessì Schmid stresses that there is no correlation between tense and telicity, as could be evoked by the distinction of pairs like *hace . . . que* and *hacía . . . que* and *desde hace . . .* and *desde hacía . . .* (2014: 119) .

c. y más cuando llueve por horas (stories, 2002)
and more when rain-3SG for hours
'and (even) more when it rains for hours'

d. hace cinco años que no llovía en la Tercera Región (press, 1997)
makes five years that not rain-3SG in the Tercera Región
'it has not rained in the Tercera Región for more than five years'

e. hacía nueve días y nueve noches que llovía sobre Peñafonte
made nine days and nine nights that rained-3SG on Peñafonte
(novel, 1993)
'it was nine days and nine nights that it had been raining in Peñafonte'

The expressions documented in (11) are *diariamente*, 'daily', *después*, 'afterwards', *hace* NP *que*, 'it has been NP since', and *hacía* NP *que*, 'it had been NP since'. The sentences show that *llover* is only used with prepositions corresponding to *for*. Two aspects are notable: The first one is that *for* is not attested in any sentence in which the origin argument is realized. This construction thus differs from the other ones, again. The construction is not attested with *in*, either, and there are so few tokens that the difference is not attributed any value for the final result. The second aspect is that the THEME's definiteness is irrelevant for the use of prepositions. This behavior distinguishes *llover* from other verbs like *eat*: *John eats pizza for/*in an hour* but *John eats a pizza *for/in an hour*. The tests with the *for*- and the *in*-PP classify *llover* as atelic.

All in all, *llover* is classified as atelic both from the semantic and from the tests' perspective with the answer to test 4 being "yes" and to test 5 "no".

2.4 The feature [±punctual]

The feature [±puncutal] is examined in test 3. Like the feature [±dynamic] it is tested through the compatibility with a special kind of expression: so-called 'pace adverbs' (Van Valin 2005: 36). These adverbs can only be used with [–punctual] verbs, because [+punctual] verbs do not have any pace to be described. They are illustrated in (12) for English and in (13) for Spanish.

(12) *English pace adverbs* (Van Valin 2005: 36)
gradually, instantly, quickly, rapidly, slowly

(13) *Spanish pace adverbs*
apresuradamente, aprisa, deprisa, despacio, despacito, gradualmente, lentamente, paulatinamente, pausadamente, rápidamente, rápido, regularmente, velozmente, a toda prisa, a vuela pluma, con paso lento, con (la mayor) prontitud, con rapidez, con regularidad, poco a poco

The use of pace adverbs with *llover* is hardly attested in the corpus. The examples in (14) show the only sentences that can be found.

(14) Llover *with pace adverbs*

a. Efrén sentía llover deprisa sobre el capote aguadero (novel, 1993)
Efrén felt rain-INF hurriedly on the CAPOTE AGUADERO
'Efrén felt a hurried rain(ing) on the CAPOTE AGUADERO'

b. ¡A cántaros llovía!, rápido y violentamente (oral)
buckets rained-3SG, quick and violently
'it was raining buckets, quickly and violently'

The adverbs are *deprisa*, 'quickly' and *rápido*, 'quickly'. The sentences show that *llover* is used with pace adverbs, but they only make up 0.0729% of all tokens. The test just asks if pace adverbs are attested, but it does not say anything about a minimum percentage. So the question is how to interpret the percentage. One interpretation is that the information regarding the pace of the precipitation event is not very likely to be given. In that case, the percentage is an indicator for answering the test with "yes", because the adverbs are used. Another interpretation is that the use is exceptional and the test is not fulfilled. Due to the fact that not even prototypical [–punctual] verbs like *eat* are used with pace adverbs very often, we claim that the first interpretation is more adequate. Test 2 is therefore answered with "yes", but does not have a very clear result.

From the semantic perspective, there is no reason to consider *llover* a [+punctual] verb: According to section 2.3, *llover* is [–telic]. From this, it follows that it must also be [–punctual], because all atelic verbs have duration and are hence [–punctual]. Another reason to consider *llover* [–punctual] is the event structure established in section 1. It consists of several steps and the precipitation event therefore cannot be punctual, but needs to take some time. This is another reason for *llover* being [–punctual]. This evaluation coincides with the intuition: The shortest precipitation event is a cloudburst, but this is already far from being a punctual event. So no precipitation event can be punctual.

All in all, the test is rather answered with "yes", meaning that *llover* is [–punctual], and the semantics clearly confirm this result.

2.5 Further tests

The last two tests refer to whether the verb has a result state and whether it is Test 6 asks if the verbal event has a resultant state that differs from the initial state. *Llover* does not cause any

change of state, but just expresses movement, so there is no new resultant state. The answer to test 6 is thus "no".

Test 7 examines if a verb is causative. A necessary condition for causativity is that the verb has two or more arguments, because the cause and its effect cannot be described otherwise. *Llover* only meets this condition in 0.66% of all cases, so the test only applies to a very small percentage of tokens. In RRG, causativity is tested by means of a paraphrase: If a paraphrase can be formed in which a causative verb is used, the verb in question must also be causative. For *llover* it is not possible to form paraphrases. This is shown by the sentences in 15, where the meaning of the first sentence is not identical to that of the second and the third.

(15) *Paraphrases as a test for causativity*

a. el cielo llueve desgracias (theatre, 1977)

b. el cielo hace que las desgracias lluevan

c. el cielo hace llover las desgracias

In Spanish, verbs can be causativized by an additional verb *hacer* that is used as an auxiliary. *Hacer* is not limited to non-causative verbs, but makes more sense if a verb is not inherently causative. *Llover* is used with *hacer* in about 1.13% of all sentences, so this is another reason to assume that it is not causative. Test 7 is therefore answered with "no".

2.6 Conclusion

Summing up, *llover* has the test results indicated in the first line of Table 2. The results do not reflect the fact that the sentences in which the origin is realized do sometimes have slightly different results, because these differences are marginal and can therefore be neglected. The results coincide with those of activities in all tests. According to the test results, *llover* is therefore considered an activity.

Llover's semantic features are represented in the first line of Table 3. They also match the features of activities, so that the semantics also imply that *llover* is an activity.

The semantics and the test results both indicate that *llover* can be classified as an activity. It is not a prototypical activity verb though. One reason is that with *llover* the feature [±dynamic] includes descriptions of intensity, mostly because both kinds of expression mean the same with *llover*. Another reason is the feature [±punctual] that does not have a clear test result, although there is no doubt semantically. A third aspect which separates *llover* from prototypical activities is that *llover* is most commonly used without an agent, but other activities typically have one. This aspect also concerns a fourth one; the number of arguments is reduced by one with *llover*. It can have zero, one, or two arguments with zero being the

Aktionsart	test 1 ger.	test 2 dyn.	test 3 pace	test 4 *for*-PP	test 5 *in*-PP	test 6 state	test 7 caus.
llover	yes	yes	yes	yes	no	no	no
state	no*	no	no	yes*	no	yes	no
achievement	no*	no	no*	no*	no*	yes	no
semelfactive	no*	no*	no*	yes*	no*	no	no
accomplishment	yes	no	yes	irrel.*	yes	yes	no
activity	yes	yes	yes	yes	no	no	no
act. acc.	yes	yes	yes	irrel.*	yes	yes	no

Table 2: Test results for *llover* and predicate classes (Van Valin 2005: 39)

Aktionsart	[±static]	[±dynamic]	[±telic]	[±punctual]
llover	[–static]	[+dynamic]	[–telic]	[–punctual]
state	[+static]	[–dynamic]	[–telic]	[–punctual]
activity	[–static]	[+dynamic]	[–telic]	[–punctual]
accomplishment	[–static]	[–dynamic]	[+telic]	[–punctual]
achievement	[–static]	[–dynamic]	[+telic]	[+punctual]
semelfactive	[–static]	[±dynamic]	[–telic]	[+punctual]
act. acc.	[–static]	[+dynamic]	[+telic]	[–punktuell]

Table 3: Semantic features (Van Valin 2005: 33)

most likely case, whereas most other activities can have one, two, or three arguments. The last reason for which *llover* is not a typical activity is its event structure that consists of many simultaneous events, while most activities only have one single event. Nonetheless, activities provide the most adequate description for *llover*, because they capture most of its properties.

The *Aktionsart* implies the logical structure (LS). It will be necessary in the following sections. As an activity, *llover* has an LS of the form presented in (16).

(16) *The* LS *of activities* (Van Valin 2005: 45)
do'(x, [**pred'**(x) or (x, y)])

Llover cannot be split into simpler verbs by semantic decomposition. **pred'** in the LS for *llover* must thus be **rain'**. *Llover* can be used with zero, one or two arguments and, accordingly, the general LS for *llover* is the one shown in (17).[6]

[6] Van Valin claims that *snow* has the LS **do'**([**snow'**]), but does not give any reasons for it apart from the fact that *snow* is a "verb with no arguments" (2005: 63). The two LS are not identical, because we consider

(17) *The* LS *of* llover *(preliminary)*
do'(x, [**rain'**(x) or (x, y)])

3 Indications of place and time

Indications of place and time both occur in almost 20% of all sentences in which *llover* is used. In total, 35.74% of all tokens have an indication of place or time, or both, and thus, there is a need for description of this phenomenon. In Spanish, the indications can be presented in several ways: Both can occur as prepositional phrases (PPs) and adverbs. Apart from that, indications of time can be realized as nominal phrases (NPs) and indications of place as dative pronouns. PPs are exemplified by the use of *sobre* ('on') and *durante* ('during') in (11b) and *por* in (11c). Another example is the *for-pp*, which was examined in test 4 (section 2.3). Adverbs are illustrated by *ayer*, 'yesterday', and *hoy*, 'today', as indications of time and *afuera*, 'outside', or *allí*, '(over) there', as indications of place. NPs only occur as indications of time like *esos días*, 'these days', or *el año pasado*, 'last year'. Pronouns like *le*, 'him', or *nos*,'us', however can only express an indication of place. An example for *le* can be found in (11a).

3.1 PP and adverbs

The formalization of PPs and adverbs is standard in RRG. The analysis depends on what the PPs and adverbs express. We first present the options for PPs and then come back to adverbs.

PPs are separated from each other by means of two criteria: One is being licensed by the verb or by the preposition and the other is having substantial proper meaning or not.[7] By these criteria, three different kinds of PP are identified (Van Valin 2005: 21–23), as summarized in Table 4.[8] Examples for the three existing options are presented in (18).

(18) *The classification of PP* (Van Valin 2005: 7, 22, 23)
a. What did Robin show [to Pat]$_{\text{non-predicative PP}}$ [in the library]$_{\text{predicative PP}}$ yesterday?
b. Kim put the book [in/on/under/…the box]$_{\text{argument-adjunct}}$

the proposal made by Belloro (2009) that Spanish subjects, which are active, i.e. not realized overtly, are represented in the LS by their corresponding feature bundles.

[7] The distinction holds not only for prepositional phrases, but for adpositional phrases in general. As *llover* only occurs with prepositional phrases, we leave out inpositions and postpositions here.

[8] The combination of being licensed by the preposition and not having substantial meaning does not exist, because in this case, the preposition does not have anything that can be licensed. As a consequence, the property "licensed by P" implies that a PP is predicational.

	licensed by V	licensed by P
no substantial meaning	non-predicational PP	—
substantial meaning	argument-adjunct	predicational PP

Table 4: The classification of PP

Llover occurs with different prepositions. Prepositions referring to locations can express a direction, as with *a*, 'to', *contra*, 'against', *en*, 'in', *hacia*, 'towards' and *sobre*, 'on', the origin of the precipitation, as with *de*, 'from', and *desde*, 'from', or a constant position, as with *en*, 'in'. The three kinds are exemplified by the sentences in (19).

(19) *PP indicating place*

a. un vapor sombrío llovía sobre las ciudades (novel, 1989) [direction]
a vapor gloomy rain-3SG.PST on the cities
'a gloomy vapor rained on the cities'

b. como si lloviera del cielo (novel, 1983) [origin]
as if rained-3SG.SUBJ from.the sky
'as if it rained from the sky'

c. llovió en Sayula (novel, 1988) [location]
rained-3SG in Sayula
'it was raining in Sayula'

The PPs in (19) are *sobre las ciudades*, *del cielo* and *en Sayula*. None of them is licensed by the verb, because *llover* does not necessarily occur with a PP. They must, therefore, be licensed by their corresponding prepositions and, consequently, denote predicational PPs. The LS by which predicative prepositions are described is exemplified and represented in (20). It shows that predicative PP are state predicates. The analogous LS for the sentences with *llover* from (19) are presented in (21). (22) shows the general LS for PP referring to the place of a precipitation event. **LOC'** represents the preposition and LOC the place of the precipitation event. [9]

[9] Note that *llover* is not forced into an accomplishment LS that would change its telicity. Rather, the LS in (22) is apt for atelic verbs only, what reinforces our claim from section 2.3 that *llover* is atelic in all constructions. Note, too, that the analysis presented here is different from the one presented for motion verbs presented in Van Valin (2005). This is due to the fact that *llover* does not express the goal of motion as does *to the park* in 'Chris is running to the park' (Van Valin 2005: 49). In Van Valin (2005: 49) the analysis of motion verbs depends on whether or not the goal-argument is expressed. The difference is exemplified by the sentences in (i).

i. *The analysis of motion verbs*

a. Chris is running in the park
be-in'(park, [**do'**(Chris, [**run'**(Chris)])])

b. Chris is running to the park
do'(Chris, [**run'**(Chris)]) & INGR **be-at'**(park, Chris)

(20) *The* LS *of predicative prepositions* (Van Valin 2005: 49)

a. Chris ran in the park

b. **be-in'**(park, [**do'**(Chris, [**run'**(Chris)])])

(21) *The* LS *of the sentences in (19)*

a. **be-on'**(ciudades, [**do'**(vapor, [**rain'**(vapor)])])

b. **be-from'**(cielo, [**do'** ([**rain'**])])

c. **be-in'**(Sayula, [**do'**(3SG, [**rain'**(3SG)])])

(22) *The general* LS *for indications of place with* llover
be-LOC'(LOC, [**do'**(x, [**rain'**(x, y)])])

If a PP expresses an indication of time, the time interval referred to can be limited to the right, to the left, to both sides, or to neither side. This is expressed by *hasta que*, 'until, by *desde*, 'for' and *a partir de*, 'from X on', and by *en*, 'in', *por*, 'for', *al + infinitive*, 'at X-ing', *durante*, 'during', and *desde hace*, 'for', respectively.[10] The use of time intervals with *llover* is exemplified in (23).

(23) PP *indicating time*

a. llueve todos los días a partir de la una (novel, 1996) [limited to the left]
rain-3SG all-PL the-PL days from the one
'it is raining every day from one o'clock on'

b. llovió mucho durante la carrera (press, 1994) [limited to both sides]
rained-3SG a lot during the race
'it was raining a lot during the race'

The PPs in (23) are *a partir de la una* and *durante la carrera*, respectively. As with the PPs indicating place, neither of the two PPs referring to the time is licensed by the verb, because *llover* does not need a PP. The indications are, consequently, licensed by the prepositions and therefore are predicative PPs. Their LS is similar to to the one proposed in (22) for PPs indicating place. The difference is that instead of **LOC'** and LOC, the LS describing indications of time has the elements **TIME'** and TIME referring to the preposition and the time interval, respectively. The LS of the sentences in (23) can be found in (24), the general LS for indications of time in (25).

---

[10] The options "both sides" and "neither side" are considered to be alike here, because they are referred to by the same prepositions.
Hasta que is not followed by a noun and therefore does not belong to a PP. It is consequently excluded from our analysis. There is no other preposition limiting a time interval to the right, so this option does not exist.

(24) *The* LS *of the sentences in (23)*

a. **be-from.x.on'**(la una, [**do'**(3SG, [**rain'**(3SG)])])

b. **be-during'**(carrera, [**do'**(3SG, [**rain'**(3SG)])])

(25) *The general* LS *for indications of time with* llover
be-TIME'(TIME, [**do'**(x, [**rain'**(x, y)])])

To sum up, the PPs occurring with *llover* have all been classified as predicational PPs, regardless of whether they express time or place. The LS are analog and show that the representation of the indication is independent of the number of arguments *llover* has. Predicative prepositions are elements of the periphery. In the syntactic template they are therefore coded as in Figure 1.

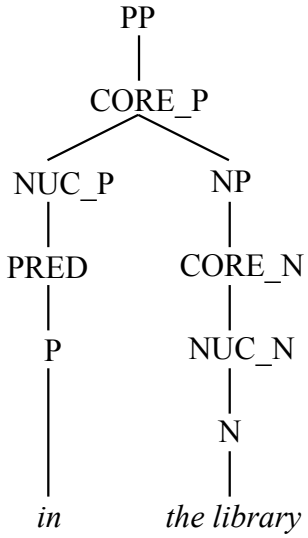


Fig. 1: Coding of predicative prepositions (Van Valin 2005: 23, 139)

Indications of time expressed by adverbs refer to the same time intervals as the corresponding PP: left bounded ones are expressed by *después*, 'afterwards', *luego*, 'then', and *pronto*, 'soon', right bounded ones by *antes*, 'before', and *ya no*, 'no longer', and those that are not limited to either side or limited to both sides by *ahora*, 'now', *aún*, 'still', *ayer*, 'yesterday', *entonces*, 'than', *hoy*, 'today', *jamás*, 'never', *mañana*, 'tomorrow', *mientras*, 'while', *nunca*, 'nunca', *siempre*, 'always', *todavía*, 'still', and *ya*, 'already'. All indications of place that are realized as adverbs express the place where it rains: *adentro*, 'inside', *afuera*, 'outside', *ahí*, 'there', *allá*, 'there', *allí*, 'there', *alrededor*, 'around', *aquí*, 'here', *detrás*, 'behind', *donde*, 'where', and *fuera*, 'outside'.

. All adverbs are treated as one-place predicates in the LS (Van Valin 2005: 49). Their argument is the part of the LS they refer to. An example with two adverbs, one for time and one for place, is given in (26).

(26) *Adverbs as indications of place or time*

a. mientras afuera llovía (novel, 1986)
while outside rain-3SG
'while it was raining outside'

b. **while'**(**outside'**(**do'**(3SG, [**rain'**(3SG)])))

In sum, the adverbs occurring with *llover* are described in RRG as any other adverb is. The LS of *llover* is therefore of the form presented in (27), when an adverb occurs.

(27) Llover *described by an adverb*
adverb'(**do'**(x, [**rain'**(x, y)]))

3.2 NPs as indications of time

NPs do not refer to indications of place. NPs used as indications of time have, to our knowledge, not yet been described in standard RRG. In the corpus, there is a great variety of NPs attested that express an indication of time.[11] Examples are presented in (28).

(28) NP*s as indications of time*

a. llovía aquella mañana (en toda la zona centro) (press, 1996)
rained-3SG that morning (in all the zone centre)
'that morning it rained (in all the central zone)'

b. en Macondo llovió cuatro años (press, 1997)
in Macondo rained-3SG four year
'it has been raining in Macondo for four years'

The indications of time in (28) are *aquella mañana* and *cuatro años*, respectively. They serve as answers to different kinds of questions: *aquella mañana* indicates the time at which it is raining, while *cuatro años* indicates the time interval during which it is raining. All NPs used with *llover* to express time can be classified as referring either to 'when' or to 'for how long'. This distinction is, therefore, used as a starting point for the formalization of NP.[12]

Indications of time making reference to 'when' that are realized as PPs have been discussed in section 3.1. The LS established in that section is **be-TIME'**(TIME, **do'**(x, [**rain'**(x, y)])). **TIME'** and TIME do not refer to the same part of the sentence: **TIME'** represents the

[11] The NPs are combinations of some kind of article or numeral with a noun. As the NPs are indications of time, the noun refers to a time of the day (*madrugada*, *mañana*, *mediodía*, ...), a day of the week, a season, or a time span like *hora*, *día*, *semana*, ... Examples are: *aquella tarde*, *cada vez*, *ese día* and *un mes seguido*, *todo el día*, and *noche y día*.

[12] There is no token in the corpus, in which an NP indicating time and serving as an answer to 'for how long' occurs with the argument that expresses the precipitation or its origin. Consequently, there is no token in which such an NP occurs with both the precipitation and its origin, either.

preposition, whereas TIME refers to the time interval mentioned. We claim that sentences like (28a) can be analyzed in an analogous way: The LS contains the TIME-element that refers to the time interval during which it rains and the **TIME'**-element is left unspecified, because there is no preposition. As with unspecified arguments, it is represented as '**∅**'. Applying this reasoning to an abbreviated version of (28a) leads to (29). The general LS for NPs referring to the time at which a precipitation event takes place, is shown in (30).

(29) llovía aquella mañana (press, 1996)
be-∅'(aquella mañana, (**do'**(3SG, [**rain'**(3SG)])))

(30) *NPs as indications of time referring to 'when?'*
be-∅'(TIME, [**do'**(x, [**rain'**(x, y)])])

Paths can be realized as bare NPs in English, Spanish and other languages. Van Valin (this volume) proposes the LS in (31) for cases in which a path is expressed.[13]

(31) *Paths realized as* NP (Van Valin this volume: 86)
Sally runs two miles | to the park.
do'(Sally, [**run'**(Sally)]) ∧ PROC **cover.path.distance'**(Sally, two miles)
| & INGR **be-at'**(park, Sally)[14]

As we did in our analysis of PPs, we claim that instead of a spatial component, there can be a temporal one in the LS: By replacing **path.distance** by **time.interval** in the LS, there is a way to represent NP referring to time and NP referring to place by analogous LS. The use of this assumption is that not only the question 'how far', but also 'for how long' can be answered. The LS corresponding to (28b) under this proposal is represented in (32). The sentence does not have an argument corresponding to 'Sally' from (31), so its position is filled by an unspecified argument '**∅**'.

[13] The basis of this idea had already been mentioned in Van Valin (2005). It was not considered an option for sentences like the one in (31), because PROC is followed by an activity and this is not semantically adequate (Van Valin 2005: 44). The issue is fixed, if the activity and the process occur at the same time: "For example, the action of eating something is simultaneously a process of consumption, and this can be expressed by replacing '&' by '∧', which means 'and simultaneously' in the representation." (Van Valin this volume: 85).

[14] The '|' is used only to indicate that 'to the park' is not necessarily part of the sentence and 'Sally runs two miles' can be analyzed on its own too.
The use of '∧' shows that there are two simultaneous events: Sally's running and her covering a distance. The LS therefore allows to represent an indication of place realized as an NP. 'Two miles' is the answer to the question 'How far is Sally running?'. The **cover.path.distance'** bit of the LS thus corresponds to the spacial interval Sally covers while she is running.

(32) llovió cuatro años (press, 1997)
do'(3SG, [**rain'**(3SG)]) ∧ PROC **cover.time.interval'**(∅, cuatro años)[15]

(33) NP *as indications of time referring to 'for how long?'*
do'(x, [**rain'**(x, y)]) ∧ PROC **cover.time.interval'**(∅, TIME)

Resuming the results obtained, the analysis of NPs as indications of time hinges on whether the indication refers to the duration or the time of the event, i.e. if it expresses for how long or when the event takes place. Both options can be represented semantically by the LS.

3.3 Pronouns as indications of place

Pronouns cannot express indications of time. They can refer to indications of place though, as is shown in (34).

(34) a. las multas le llueven diariamente desde hace años (press, 1995)
the fines him.DAT rain-3PL.PRES daily for years
'the fines have been raining daily on him for years'

b. una tarde nos llovió encima (novel, 1978)
one evening us.DAT rained-3SG on
'one evening it was raining on us'

The pronouns in (34) are *le* and *nos*. As with all pronouns expressing the place of the precipitation event, they have been assigned dative case. Most of the pronouns are in third person, although there are tokens for all persons.

Syntactically, the pronouns express dative case. Semantically, *me* is equal to the PP *sobre mí*, *te* to *sobre ti* etc. The PPs have already been analyzed in the previous section. Its LS cannot be used to describe the pronoun though, because the pronoun is syntactically different and therefore needs to be represented by a different LS. The pronouns also express a *dativus (in)commodi* and thus claim a thematic relation like the traditional BENEFICIENT. In the case of *llover*, the BENEFICIENT expresses a person or a place that is hit by the precipitation and does not have any influence on being hit or not. BENEFICIENTS are traditionally represented in causative structures.[16] We argued in section 2 that *llover* is not a causative verb. The dative

[15] The PP *en Macondo* that is part of the original sentence, has been left out in (32) in order to keep the LS simple. If it is added, the LS expands to **be-in'**(Macondo, LS from (32)), i.e. **be-in'**(Macondo, **do'**(3SG, [**rain'**(3SG)]) ∧ PROC **cover.time.interval'**(∅, cuatro años)).

[16] Belloro (2009) analyzes Spanish datives. She differentiates between clitic-only, NP-only and clitic-doubling constructions and proposes the LS in 16 for clitic-only constructions. "act" means 'active' and "ina" is short for 'inactive'. The degrees of activity refer to the cognitive state of a referent.

i. les dieron un asueto
they gave them a day off
[**do'**([3pl]$_{act}$), ∅] CAUSE [BECOME **have'**([3m.pl]$_{act}$, asueto$_{ina}$)]

occurring with *llover* therefore cannot be related to causativity[17] and an LS with the CAUSE operator is hence inadequate. Instead, all LS must be based on the LS following from the classification as an activity. It is repeated in (35) for convenience.

(35) *The* LS *of* llover *preliminary* (=(17))
do'(x, [**rain'**(x, y)])

Llover is used in a variety of constructions. Its arguments can refer to the precipitation, its origin or a beneficient and can all be left out. All combinations of the three semantic arguments are documented in the corpus except for the one in which all of them occur in the same sentence. The spectrum is represented in the first column of Table 5. "P" means precipitation, "O" origin and "B" beneficient.

The BENEFICIENT is a core argument and case marked by the dative. As dative case is assigned to non-macrorole core arguments (Van Valin 2005: 110), the BENEFICIENT is assumed to be a non-macrorole core argument. This implies two things for the representation of the BENEFICIENT in the LS: First, the BENEFICIENT will always be the less active in Dowty's (1979) terms and therefore realized as the y-argument. This explains, at least partially, the lower four rows of Table 5. Second, the BENEFICIENT is a non-macrorole argument. This implies that the number of macroroles is reduced by one if the BENEFICIENT is realized as an argument and explains the '[MR1]' and '[MR0]' additions in Table 5. Many of the aspects represented in Table 5 can thus be derived from the fact that the BENEFICIENT is a non-macrorole core argument. The missing information is, at this point, an explanation of the macroroles and case marking, i.e. columns three and four of Table 5. It will be given in the following section.

sem. arg.	LS	MR	case
—	**do'**(3SG, [**rain'**(3SG)])	actor	—
O	**do'**(x, [**rain'**(x)])	actor	NOM
P	**do'**(x, [**rain'**(x)])	actor	NOM
B	**do'**(3SG, [**rain'**(3SG, y)]) [MR0]	—	DAT
O & P	**do'**(x, [**rain'**(x, y)])	actor, undergoer	NOM, ACC
O & B	**do'**(x, [**rain'**(x, y)]) [MR1]	actor	NOM, DAT
P & B	**do'**(x, [**rain'**(x, y)]) [MR1]	actor	NOM, DAT

Table 5: The constructions of *llover*

[17] This is unlikely, because with most verbs, it is: Van Valin (2007) examines three-place predicates and claims that they all have the basic LS **do'**(x, ∅) CAUSE [BECOME **predicate'**(y, z)] (2007: 31).

4 The linking algorithm applied to *llover*

Having investigated which *Aktionsart llover* belongs to and how to analyze indications of place and time, the aim of this section is to apply the linking algorithm to *llover* in order to find out if it correctly predicts the sentences presented in the introduction (section 1). They are repeated in (36) for convenience. The algorithm in the version that is adapted to Spanish is presented in (37). It will be carried out step by step in this section.

(36) *Constructions of Spanish* llover

a. las nubes llovían al revés (novel, 1990)

b. el cielo llueve todas las desgracias juntas sobre nosotros (theatre, 1992)

c. de noche llovía hollín (novel, 1986)

d. una tarde nos llovió encima (novel, 1994)

(37) *Spanish linking algorithm: semantics → syntax* (Belloro 2009: 509)

1. Construct the semantic representation of the sentence, based on the logical structure of the predicator and select the realization of each argument based on the activation level of its referent:
   a. If active, fill in the respective argument position with relevant bundle of pronominal features.
   b. If accessible, fill in the respective argument position with the corresponding nominal, plus its pronominal features.
   c. If inactive or non-identifiable, fill in the respective argument position with the corresponding nominal exclusively.
2. Determine the actor and undergoer assignments, following the actor-undergoer hierarchy (Van Valin 2005: 61).
3. Determine the morphosyntactic coding of the arguments:
   a. Select the PSA, based on the PSA selection hierarchy (Van Valin 2005: 100).
   b. Assign the XPs the appropriate case markers and/or adpositions (Belloro 2007: 165).
4. Select the syntactic template(s) for the sentence following the appropriate template selection principles (González Vergara 2006; Belloro 2007: 185).
5. Assign the elements in each argument positions to the appropriate slots in the syntactic template:
   a. Assign pronominal features to the AGX.[18]

[18] AGX is short for 'agreement index' and denotes a node in the syntactic representation of a sentence in which

b. Assign nominals to the appropriate positions in the clause, subject to focus structure.
   i. Assign focal elements to the last position in the core (default)

c. Assign any [+WH] arguments to the precore slot.

The first step of the linking algorithm is to construct the semantic representation of the sentence.[19] To do so, the activation level of the referents must be examined. Belloro (2007, 2009) distinguishes the five activation levels presented in the hierarchy in Figure 2.

active > accessible > inactive > brand-new anchored > brand-new unanchored[20]

Fig. 2: Activation levels (Belloro 2009: 498)

The activation levels are deduced from the contexts. In (36d), the first referent is expressed by the verb ending and as such considered active. It is therefore represented by its pronominal features in the LS (Belloro 2009: 510). The other referent, *nos*, is in the focus of the addressee's consciousness and, therefore, also active. The referents found in the other sentences are all inactive, because they are identifiable, but have not been mentioned in previous sentences of the conversations. Consequently, the arguments must be realized as nominals in the LS according to step 1 of the linking algorithm. The LS of the sentences in (36) are presented in (38) and based on the LS of *llover*.[21]

(38) *Step$_1$: the semantic representation*

a. **be-at'**(revés, [**do'**(nubes$_{ina}$, [**rain'**(nubes$_{ina}$)])])

b. **be-on'**(nosotros, [**do'**(cielo$_{ina}$, [**rain'**(cielo$_{ina}$, desgracias$_{ina}$)])])

c. **be-at'**(night, [**do'**(hollín$_{ina}$, [**rain'**(hollín$_{ina}$)])])

d. **be-∅'**(tarde, [**on.top'**(**do'**(3SG$_{act}$, [**rain'**(3SG$_{act}$, 1PL$_{act}$)]))]) [MR0]

The second step determines the macroroles by means of the actor-undergoer hierarchy (AUH). The AUH is represented in Figure 3. It follows from the AUH, that the *x*-argument of *llover* is the actor and the *y*-argument the undergoer, because Spanish is an accusative language. Consequently, *nubes* is actor in (39a), *cielo* in (39b), *hollín* in (39c), and *3*SG in

the pronominal features of nouns without subject are summarized. "The AGX is a dependent of the NUCLEUS, and it receives the agreement specifications of all core argument positions present in the Logical Structure" (Belloro 2004: 43). It has been developed by Belloro to account for Spanish sentences, in which the subject is omitted, because it can be deduced from the context.

[19] *Semantic representation* is equal to *LS* here, although usually it contains operators too. The operators are omitted here in order to keep things simple and because we want to focus on semantic roles and syntactic functions.

[21] 'act' is short for 'active' and 'ina' for 'inactive'.

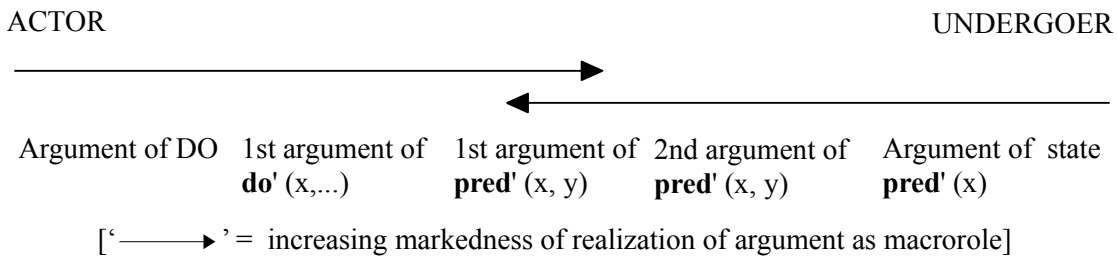


Fig. 3: The actor-undergoer hierarchy (Van Valin 2005: 126)

(39d). *Desgracias* in (39b) is attributed the undergoer MR.[22] *1*PL in (39d) is not assigned any MR, because the LS is marked by '[MR1]' and the macrorole has been attributed to the more active element, *3*SG. *1*PL must therefore be a non-macrorole argument (NMR). The output of $step_2$ is presented in (39).

(39) *$Step_2$: The assignment of* MR

a. **be-at'**(revés, [**do'**(ACT: $nubes_{ina}$, [**rain'**(ACT: $nubes_{ina}$)])])

b. **be-on'**(nosotros, [**do'**(ACT: $cielo_{ina}$, [**rain'**(ACT: $cielo_{ina}$, UND: $desgracias_{ina}$)])])

c. **be-at'**(night,[**do'**(ACT: $holl\acute{i}n_{ina}$, [**rain'**(ACT: $holl\acute{i}n_{ina}$)])])

d. **be-∅'**(tarde, [**on.top'**(**do'**(ACT: $3SG_{act}$, [**rain'**(ACT: $3SG_{act}$, NMR: $1PL_{act}$)]))]) [MR0]

The third step serves to determine the case marking of the arguments. First, the privileged syntactic argument (PSA) is selected on the basis of the hierarchy in (40) and the principles in (41).

(40) *Privileged syntactic argument selection hierarchy* (Van Valin 2005: 100):
arg. of DO > 1st arg. of **do'** > 1st arg. of **pred'**(x, y) > 2nd arg. of **pred'**(x, y) > arg. of **pred'**(x)

(41) *Accessibility to privileged syntactic argument principles* (Van Valin 2005: 100):

a. Accusative constructions: highest ranking direct core argument in terms of [(40)] (default).

b. Ergative constructions: [. . .]

c. Restriction on PSA in terms of macrorole status: [. . .]

(40) and (41) imply that the *x*-argument is selected as PSA, because it is the highest ranking argument and Spanish is an accusative language. Application of the principles to the LS in

[22] As always with acivities, the undergoer is only attributed to referential second arguments: "Multiple-argument activity verbs with a non-referential second argument, e.g. *beer* in *Kim drank beer*, never have an undergoer macrorole."(Van Valin 2005: 63). *Desgracias* is referential in (36b), because it represents the phrase *todas las desgracias juntas*.

(39) shows that *nubes* from (42a), *cielo* from (42b), (42c), and *3*SG from (42) are actors of the corresponding sentences. They are, consequently, selected as PSA and assigned nominative case (NOM). The undergoer element *desgracias* in (42b) is assigned accusative case (ACC) (Van Valin 2005: 109; Belloro 2007: 165). The NMR in (42d) is assigned dative case by default (Van Valin 2005: 109; Belloro 2007: 165). All adpositions needed are represented in the LS already, because they occur in PP. The output of step$_3$ is represented in (42).

(42) *Step$_3$: The morphosyntactic coding*

a. **be-at'**(revés, [**do'** (ACT: nubes$_{ina}$, [**rain'** (ACT: nubes$_{ina}$)])])
[PSA: NOM] active, 3PL [PSA: NOM]

b. **be-on'**(nosotros, [**do'** (ACT: cielo$_{ina}$, [**rain'** (ACT: cielo$_{ina}$,
[PSA: NOM] active, 3SG [PSA: NOM]
UND: desgracias$_{ina}$)])])
ACC

c. **be-at'**(night, [**do'** (ACT: hollín$_{ina}$, [**rain'** (ACT: hollín$_{ina}$)])])
[PSA: NOM] active, 3SG [PSA: NOM]

d. **be-∅'**(tarde, [**on.top'**(**do'** (ACT: 3SG$_{act}$, [**rain'** (ACT: 3SG$_{act}$,
[PSA: NOM] active, 3SG [PSA: NOM]
NMR: 1PL$_{act}$)]))]) [MR0]
DAT

Step four is concerned with the selection of appropriate syntactic templates. The syntactic templates proposed for English by Van Valin (2005: 15) have been adapted to Spanish by González Vergara (2006: 104–107). The Spanish templates for the nucleus are presented in Figure 4, those for the core in Figure 5 (our translations). The templates for clause and sentence are not represented due to reasons of space.[23]

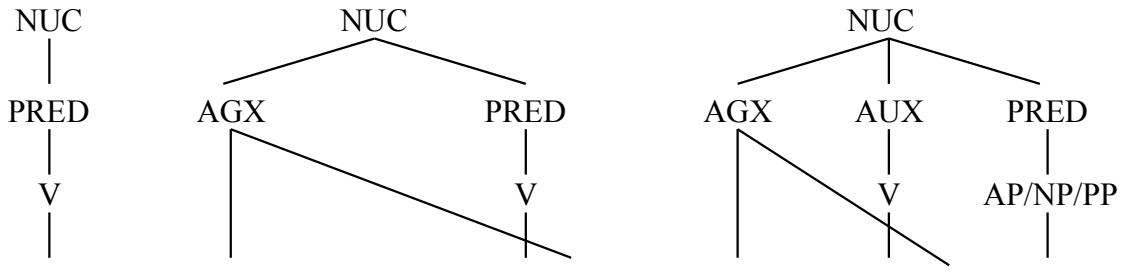


Fig. 4: Spanish nucleus templates (González Vergara 2006: 104)

[23] González Vergara distinguishes two clause templates and three sentence templates. The clauses differ because there can, but need not be a pre-Core-Slot (PrCS) and the sentences differ because both the right and the left detached position can be filled, but need not be filled (González Vergara 2006: 105–106).

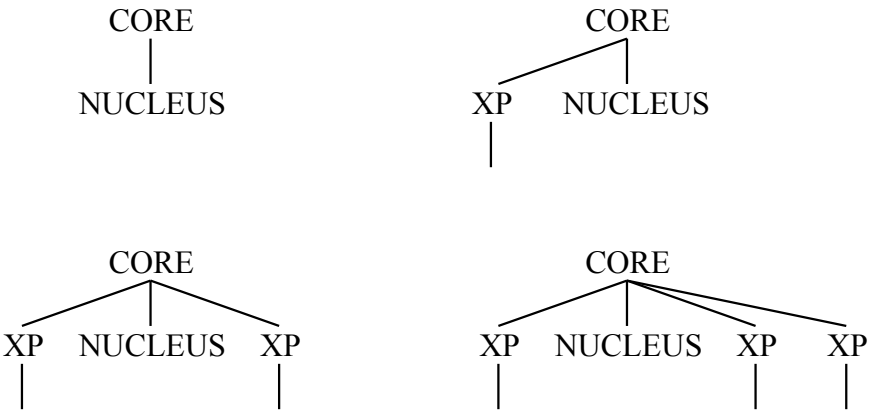


Fig. 5: Spanish core templates (González Vergara 2006: 105)

The adequate syntactic template is chosen according to the syntactic template selection principle represented in (43) and the language-specific principle in (44).

(43) *Syntactic template selection principle* (Van Valin 2005: 130)
The number of syntactic slots for arguments and argument-adjuncts within the core is equal to the number of distinct specified argument positions in the semantic representation of the core.

(44) *Spanish-specific qualifications of the principle in [a.]* (Belloro 2007: 185)
Argument positions filled exclusively by feature bundles in the semantic representation of the core do not require syntactic slots in the core template.

ex.	sem. arg.	arguments	nucleus	core
(36a)	O	las nubes	b	b
(36b)	O & P	el cielo, las desgracias	a	c
(36c)	P	hollín	a	b
(36d)	B	3SG	a	b

Table 6: Syntactic templates for *llover*

Table 6 summarizes which of the nucleus and core templates are needed in the sentences with *llover*. The nucleus templates are numbered from left to right, the core templates from top left to bottom right. Meteorological verbs are the prototypes for nucleus template a (González Vergara 2006: 104). (36a) is an exception, because its argument is in the plural and therefore the verb is different from the default form attested with *llover*, which is *3SG*. The core templates are selected according to the number of arguments realized with the verb, which are not mere feature bundles. Accordingly, there are two arguments in (36b) and one argument

in (36a), (36c) and (36d). In (36d) there is no second argument, because *3SG* is a feature bundle and as such not assigned an argument position according to (44). The output of $step_4$ is presented in the same figure as that of $step_5$, because Figure 6 consists of the parts elaborated in $step_4$. The corresponding figure is Figure 6. It does not include the operator part of the syntactic representation, because operators have been left out from the beginning on to keep things simple. The arguments listed in Table 6 are core arguments, the arguments of the PP are peripheral arguments. The representation of peripheral elements in syntactic templates was introduced in section 3 and follows standard rules.

In the last step, the elements from the argument positions are assigned to the slots in the syntactic template. To do so, AGX is assigned its features first: In the sentences with *llover* AGX occurs only once. It has the value '3PL'. After determining AGX, the other arguments are assigned their positions: The core arguments *las nubes* and *el cielo* are preverbal NPs, *todas las desgracias juntas* and *hollín* are postverbal NPs. The different positioning of the single NP of (36a) and (36c) is due to the meaning of the argument: The origin is closer to a prototypical subject than to a prototypical object and therefore realized in the preverbal position. The precipitation behaves the opposite way: It has more object properties and is therefore realized in the more object-like position, i.e. the postverbal one. The peripheral arguments *noche* and *tarde* are part of preverbal PP,[24] *nosotros* is part of a postverbal PP and *encima* is a postverbal ADVP. There is no [+WH] element, so there is no element in the precore slot.

This section was dedicated to applying the linking algorithm to *llover* in order to see if it leads to the correct formation of the sentences in (36). The results summarized in Figure 6 show that it does so.

In $step_1$, the semantic representation was established based on the results of section 2, i.e. the *Aktionsart*. The LS was taken as basis for the assignment of MR in $step_2$, which proved to give the desired output. The morphosyntactic coding was established by means of general rules and a Spanish-specific addition proposed by Belloro (2009) in $step_3$. In $step_4$ general rules did not prove to be sufficient, but the selection of syntactic templates was achieved through syntactic templates established for Spanish by González Vergara (2006). The last step showed that all sentences are predicted correctly. This shows that the modified version of the linking algorithm established by Belloro (2007) for Spanish works out sufficiently and the syntactic templates developed by González Vergara (2006) make correct predictions, even for avalent meteorological verbs.

[24] *Tarde* is analyzed as being part of a PP instead of an NP, because this corresponds to the analysis proposed in section 3.2.

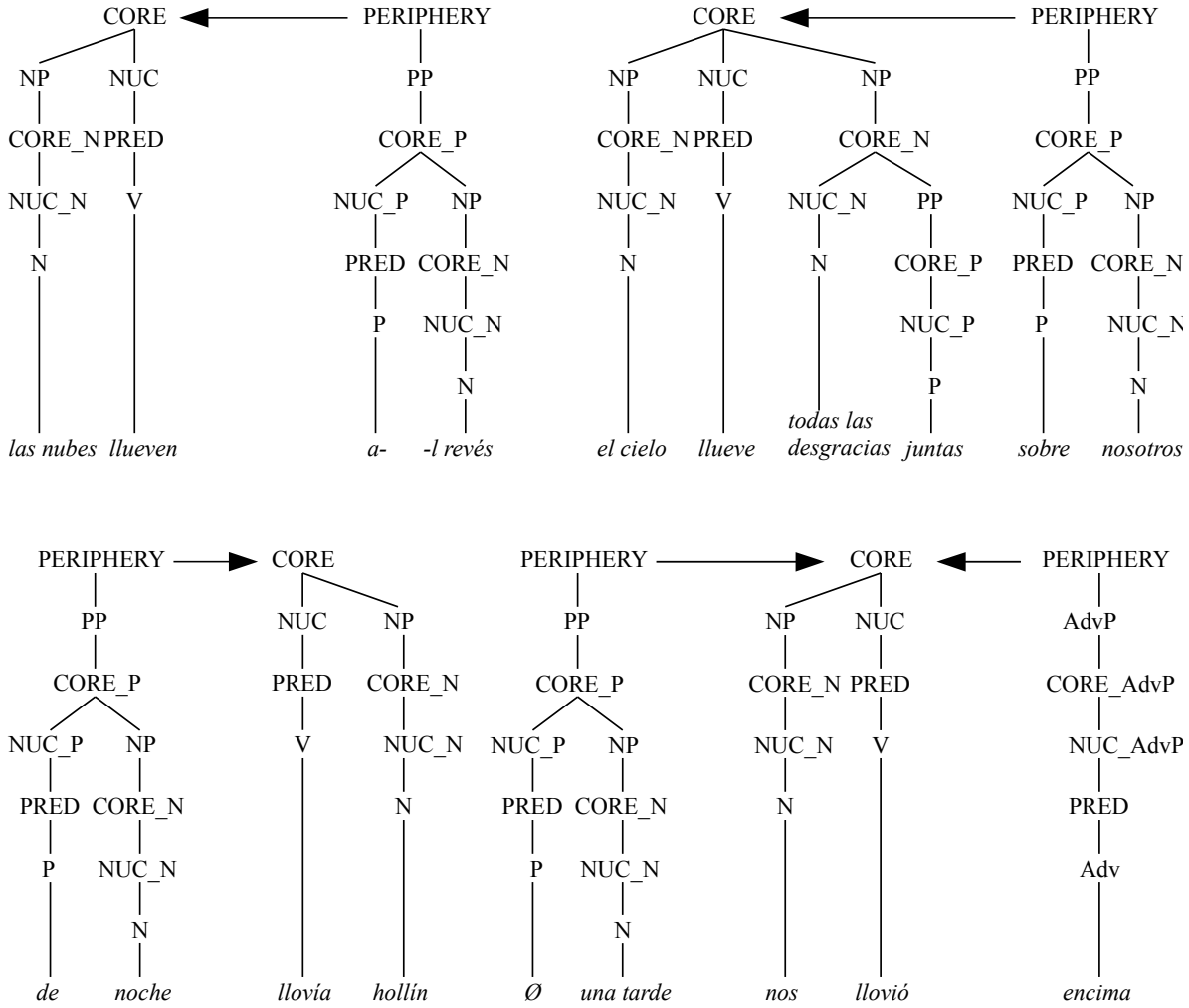


Fig. 6: Step$_5$: Syntactic templates for *llover*

5 Summary

The aim of this paper was to describe the Spanish verb *llover*, 'rain', in the framework of RRG. The first section introduced the verb and the constructions it is used in. It also presented a quantitative overview of the use of different arguments.

The objective of the second section was to determine the *Aktionsart* that *llover* belongs to. To do so, the standard tests were adapted to Spanish and applied to *llover*. The outcome is that *llover* is an atypical activity. This means that it shares most properties with activities but is not prototypical, because it has some features that differ from activities.

The third section was dedicated to the description of indications of place and time, because they are used frequently with *llover* without being arguments. They are attested in four different varieties: as PPs, adverbs, pronouns and NPs. In standard RRG, only PPs and adverbs

are considered in the context of indications of place and time, so NPs and pronouns are non-standard cases. Nonetheless, their description proved to be possible. NPs are described in the same way as PPs with the assumption of an unspecified preposition '∅'. The pronouns always occur in dative case and are therefore analyzed according to a proposal by Belloro (2009) on Spanish datives.

In the fourth section, the insights from the previous sections were used in the linking algorithm. It was applied to four Spanish sentences featuring *llover* in order to find out if it makes correct predictions. The linking algorithm used is a modified version of the standard algorithm proposed by Van Valin (2005) and was adapted to Spanish by Belloro (2009). The first two steps are carried out as in English, the morphosyntactic coding considered in the third step works with a special option developed for Spanish by Belloro (2007), the agreement index AGX. The agreement index helps to analyze sentences in which there is no overt subject. The syntactic templates for Spanish were considered in $step_4$ and summarized in $step_5$, where the arguments are assigned to the slots in the syntactic templates. The linking algorithm makes correct predictions for the Spanish sentences that were examined.

All in all, RRG can describe all phenomena occurring with *llover*: avalent verbs, different constructions of one single verb, and linking in avalent constructions. However, the determination of the *Aktionsart* is not very clear with *llover*.

References

Belloro, Valeria A. 2004. *A Role and Reference Grammar account of third-person clitic clusters.* Master's thesis, State University of New York at Buffalo.

—. 2007. *Spanish clitic doubling: A study of the syntax-pragmatics interface.* Ph.D. thesis, State University of New York at Buffalo.

—. 2009. Spanish datives: remarks on the information-structure side of the story. In *Studies in Role and Reference Grammar.* L. Guerrero, S. Ibáñez Cerda & V. A. Belloro (eds), 491–516. Mexico: Universidad Autónoma de México.

Dessì Schmid, Sara. 2014. *Aspektualität: ein onomasiologisches Modell am Beispiel der romanischen Sprachen.* [Beihefte zur Zeitschrift für romanische Philologie]. Berlin: Mouton de Gruyter.

Dowty, David R. 1979. *Word meaning and Montague Grammar.* Dordrecht: Reidel.

—. 1977. Toward a semantic analysis of verb aspect and the english 'imperfective' progressive. *Linguistics and Philosophy*, 1(1): 45–77.

González Vergara, Carlos. 2006. *Las construcciones no reflexivas con "se". Una propuesta desde la Gramática del Papel y la Referencia.* Ph.D. thesis, Universidad Complutense de Madrid.

Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface.* Cambridge: CUP.

—. 2007. The role and reference analysis of three place predicates. *Suvremena Lingvistika*, 63(1): 3–63.

Yllera, Alicia. 1999. Las perífrasis verbales de gerundio y participio. In *Gramática descriptiva de la lengua española, Vol. 2.* I. Bosque Muñoz & V. Demonte Barreto (eds), 3391–3441. Madrid: Espasa Calpe.

Online Dictionaries

[Collins] *Collins English to Spanish Dictionary* http://www.collinsdictionary.com

[Pons] *Pons Online Wörterbuch Englisch–Spanisch* http://www.pons.de

The Second Dative in Russian: A Case for the Last-Resort Case Assignment

Wataru Nakamura
Tohoku University

Abstract

The aim of this paper is to provide an RRG account of the second dative phenomena, i.e. dative case assignment to the secondary predicates in Russian, *odin* "alone" and *sam* "oneself". These so-called semi-predicatives [SemPs] must agree in case with the element they modify, but they may bear dative case, a case morpheme that is different from the one borne by the nouns they modify. I will show that the dative case assignment to the SemPs follows from an interaction among the markedness of case morphemes, the domain of case assignment, and the presence/absence of structure sharing between the controller and controllee in control constructions and will argue that dative case on the SemP shows up as the last-resort case where no other case morpheme is available.

1 Introduction

The aim of this paper is to provide an RRG account of the *second dative phenomena*, i.e. dative case assignment to a few quantifiers in Russian (e.g. *odin* "alone", *sam* "oneself") that serve as a secondary predicate to the main verb. These so-called *semi-predicatives* [SemPs] agree in case with the element they modify: when they modify the nominative-marked subject of a finite clause as illustrated in (1a), they bear nominative case. They are in contrast with adjectival secondary predicates, which bear nominative case (in agreement with the subject) or instrumental case, as illustrated in (1b) (Comrie 1974: 123–124):

(1) a. Ivan vernulsja odin/*odnomu/*odnim.
Ivan.NOM returned alone.NOM/DAT/INS
'Ivan returned alone.'

b. Ivan vernulsja ugrjumyj/ugrjumym.
Ivan.NOM returned gloomy.NOM/INS
'Ivan returned gloomy.'

What is peculiar about these SemPs is that they bear nominative, dative, or accusative case, depending on the syntactic context in which they occur. Specifically, they bear nominative case as illustrated in (2a) when they occur in a dependent core of actor control constructions; they bear dative or accusative case as illustrated in (2b) when they occur in a dependent core of undergoer control constructions; and they bear dative case as illustrated in (2c) when they occur in a subordinate core with no argumental controller (Comrie 1974, Franks 1995, Babby 1998, Landau 2008, Bailyn 2012):[1]

(2) a. Actor Control Constructions (SemP=Nominative-marked)

Vanja xočet prijti odin/*odnomu.
Vanja.NOM wants come.INF alone.NOM/DAT
'Vanja wants to come alone.'

b. Undergoer Control Constructions (SemP=Dative/Accusative-marked)

Ona poprosila ego ne ezdit' tuda *odin/odnomu(/odnogo).
she.NOM asked him.ACC NEG go.INF there alone.NOM/DAT(/ACC)
'She asked him not to go there alone.'

c. No Argumental Controller (SemP=Dative-marked)

Ivan dumaet čto pojti domoj *odin/odnomu važno.
Ivan.NOM thinks that go.INF home alone.NOM/DAT important
'Ivan thinks that it is important to go home alone.'

Table 1 is a summary of the correlation between the case marking of SemPs and their syntactic environments in (2a)–(2c).

Case Marking of SemPs	Environment	Example
Nominative-marked SemPs	A dependent core of actor control constructions	(2a)
Dative(/Accusative)-marked SemPs	A dependent core of undergoer control constructions	(2b)
Dative-marked SemPs	A subordinate core with no argumental controller of the unexpressed subject	(2c)

Table 1: The Case Marking of SemPs and Their Syntactic Environments

(3a, b) are generalized from (2a)–(2c), under the assumption that these SemPs agree in case with the unexpressed subjects they modify (Comrie 1974, Babby 1998):

[1] Comrie (1974) and Babby (1998) note that the accusative marking of SemPs as illustrated by (2b) is colloquial, while Landau (2008: 888–889) casts doubt on their claim.

(3) a. Actor control constructions such as (2a) require their unexpressed controllees to bear nominative case.

b. All the other controllees bear dative case except for those of undergoer control constructions such as (2b) that may bear accusative as well as dative case.

The case agreement requirement leads us to reason that the unexpressed controllees in (2a) and (2b), respectively, bear nominative and dative/accusative case, but it remains unclear how to explain the dative marking of the SemPs in (2b, c) within the framework of RRG as outlined in Van Valin & LaPolla (1997) and Van Valin (2005).

The rest of this paper is organized as follows. Section 2 gives a brief summary of the RRG account of control and matrix-coding (i.e. 'raising') constructions and their case assignment. Section 3 proposes a modification of the RRG account of control constructions (so that they may or may not involve structure sharing between the controller and controllee) and argues that the nominative-marked SemP in (2a) and the accusative-marked SemP in (2b) show up as a consequence of the local case agreement, while the dative-marked SemPs in (2b, c) result from the last-resort case assignment that would take place when the unexpressed controllee receives no case value for the SemP to agree with. Section 4 is a conclusion.

2 Role and Reference Grammar

RRG characterizes complex sentences in terms of a combination of levels of juncture (i.e. *nuclear*, *core*, and *clause*) and nexus relations (i.e. *coordination*, *cosubordination*, and *subordination*) and defines control and matrix-coding constructions as embodying a non-subordinate core juncture (i.e. *core coordination* or *core cosubordination*) in which the matrix and dependent core share one argument. (4a)–(4d) are English examples:

(4) Control Constructions

a. Actor control:
John wanted to watch the movie.

b. Undergoer control:
John wanted the students to watch the movie.

Matrix-Coding ('Raising') Constructions

c. Matrix-Coding as PSA:
John seemed to watch the movie.

d. Matrix-Coding as non-PSA:
John expected the students to watch the movie.

The major difference between them is that the shared argument is a semantic argument of the matrix verb in control constructions, while it isn't in matrix-coding constructions.

Both control and matrix-coding constructions involve argument sharing, but it has been customary in the literature to assume that they differ with respect to degree of argument sharing, as shown in Table 2. Specifically, control constructions involve a referential dependency between the controller and unexpressed controllee (the controller is a referential antecedent of the controllee serving as a null pronoun), while matrix-coding constructions embody their structure sharing. What is crucial about structure sharing is that the matrix-coded argument and its counterpart in the dependent core share not only the referential values (i.e. person, gender, and number), but also the case value (Table 2).

		Control ('PRO') Referential Dependency	Matrix-Coding Structure Sharing
Referential Values	Person Gender Number	○	○
Case Value		×	○

Table 2: Referential Dependency vs. Structure Sharing

Table 2 states that the controller and controllee in control constructions share their referential values alone, while the matrix-coded argument and its counterpart in the dependent core in matrix-coding constructions share both their referential values and case value.

(5) is a set of case assignment rules for accusative case-marking systems that crucially refers to the privileged syntactic argument [PSA] selection hierarchy in (6):

(5) Case Assignment Rules (Accusative) (Van Valin & LaPolla 1997: 359)

a. Assign nominative case to the highest-ranking macrorole argument.

b. Assign accusative case to the other macrorole argument.

c. Assign dative case to non-macrorole arguments (default)

(6) Privileged Syntactic Argument [PSA] Selection Hierarchy
Actor > Undergoer

The case assignment rules in (5) work as follows: (5a) assigns nominative case to both S and A arguments, which correspond to "the highest-ranking macrorole argument" in (5a), while (5b) assigns accusative case to O arguments. (5b) defines accusative as a dependent case assigned to a macrorole argument only when it occurs with an A (i.e. highest-ranking macrorole) argument in the same core. This means that when (5b) applies to control constructions with a subordinate core that contains a transitive verb, for example, it applies to both an O argument in a matrix core (if any) and the one in a subordinate core. Finally, (5c) states that dative is the default case for non-macrorole core arguments (Van Valin 1991; cf. Silverstein 1976, 1993).

The next question to ask is how to extend (5a)–(5c) to control and matrix-coding constructions. The fact that they have two cores (i.e. the matrix and dependent core) brings up a question of whether (5a)–(5c) apply to each individual core separately or they apply to all of the cores jointly within the clause. What is at issue here is whether the domain of case assignment [DCA] is the core or the clause. Van Valin & LaPolla (1997) propose that Icelandic and English, for example, choose the core and clause as their DCA, respectively. Nakamura (1999) also proposes that the DCA for Japanese is the core. The evidence that the DCA for Icelandic is the core comes from (7a, b), in which each of the matrix and dependent core has one nominative argument in it (Andrews 1982: 451, Van Valin & LaPolla 1997: 578):

(7) a. Hún bað hann að vera góðan/góður.
she.NOM request.PAST him.ACC to be good.ACC/NOM
'She requested him to be good.'

b. Jón-ø tel-ur mér haf-a alltaf þótt Ólaf-ur
John-NOM believe-PRS me.DAT have-INF always think.PPTC Olaf-NOM
leiðinleg-ur.
boring-NOM
'John believes me to have always considered Olaf boring.'

If the DCA for Icelandic were the clause, we would make an incorrect prediction that only the highest-ranking macrorole arguments in the matrix cores of (7a, b) (namely, *Hún* and *Jón*) would bear nominative case. This would make it impossible to explain why the predicate adjective in the dependent core in (7a), which is in case-agreement with the unexpressed controllee, may be nominative-marked or why (7b) has a nominative-marked argument in the dependent core.

(8a)–(8c) are a summary of this section:

(8) a. Distinction between Control and Matrix-Coding Constructions (to be revised in Section 3):
Control constructions involve *referential dependency* between the controller and controllee, while matrix-coding constructions involve *structure sharing*.

b. Case Assignment Rules (Accusative):
Assign nominative case to the highest-ranking macrorole argument.
Assign accusative case to the other macrorole argument.
Assign dative case to non-macrorole core arguments (default).

c. Domain of Case Assignment [DCA]:
Core e.g. Icelandic, Japanese, Hungarian
Clause e.g. English, Enga, Newari

3 Proposal

3.1 The Last-Resort Case Assignment

Let us begin with the question of whether the DCA for Russian is the core or the clause. The clue comes from (2b), an undergoer control construction. Given that a SemP is in case-agreement with a subject it modifies, we may see that the unrealized controllee in (2b) may bear dative case, a different case than the one assigned to the controller. This observation, in turn, suggests that (5a) cannot apply to the matrix core and dependent core independently.[2] The above consideration suggests that the DCA for Russian is the clause.

Before addressing the question of the case marking of the SemPs in (2), let me make three more assumptions. First, I assume that verbal arguments must receive some case marking in inflected languages like Russian. Second, I posit the markedness hierarchy of case morphemes in (9), proposed by Silverstein (1976, 1993), together with the *NP lexical content hierarchy*:

(9) Case Hierarchy [CH] (Silverstein 1976, 1993; cf. Van Valin 1991)
Nom > Dat > {Acc, Erg} > Gen

What is notable about the case hierarchy [CH] is that dative is taken to be the *second* least marked case morpheme.[3]

Finally, I follow Hudson (1998, 2003) (cf. Przepiórkowski & Rosen 2005) in assuming that the distinction between control and matrix-coding constructions is orthogonal to the one between referential dependency and structure sharing and that the only difference between control and matrix-coding constructions comes down to whether or not the shared argument

[2] If (5a) were able to apply to the matrix and dependent core independently, we would make an incorrect prediction that the controllee may bear nominative (instead of dative) or accusative case (or, more generally, the same case as the controller), as illustrated in Icelandic (Andrews 1982).

[3] Silverstein (1993) attributes the unmarkedness of dative as opposed to accusative/ergative to the fact that the former tends to show up (together with nominative) where the latter cannot show up (e.g. nominalized clauses) and argues that the nominative and dative serve as two focal points (as found in the system of color terms) in the variation of case-marking systems. See Silverstein (1993: 485–497) for further discussion.

is a verbal argument of the matrix verb. This is a departure from the traditional assumption in (8a) that only matrix-coding constructions embody structure sharing. The consequence is that control constructions may involve either a referential dependency between the controller and controllee or their structure sharing. This move leads to an extended four-way typology of non-subordinate core junctures in Table 3.

	Referential Dependency	Structure Sharing
Control Constructions		
Matrix-Coding Constructions		

Table 3: An Extended Typology of Non-Subordinate Core Junctures

Given these assumptions, we are ready to account for the case marking of the unexpressed controllees in (2a)–(2c). Let us begin with (2c), repeated below for convenience:

(2) c. Ivan dumaet čto pojti domoj *odin/odnomu važno.
Ivan.NOM thinks that go.INF home alone.NOM/DAT important
'Ivan thinks that it is important to go home alone.'

(2c) involves no argumental controller for the unexpressed subject in the subordinate core. Under the proposal that (5a) can't apply to the subordinate core in Russian, the unexpressed subject cannot receive nominative case. In fact, none of (5a)–(5c) can apply to the controllee.

Against this backdrop, I propose to assign the least marked case morpheme available to the unexpressed infinitival subject. This means that the unexpressed controllee (the highest-ranking macrorole argument) in the dependent core in (2c) bears the least marked case morpheme other than nominative (i.e. dative) as the last-resort measure to avoid leaving the controllee noun caseless:

(10) The Last-Resort Case Assignment[4]
When (5a) is unable to apply to an argument NP for some reason, it receives the least marked case morpheme other than nominative (i.e. dative).

This last-resort case assignment requires the SemP in (2c) to bear dative case as a consequence of the local case agreement with the unexpressed controllee.

3.2 Control Constructions May Involve Structure Sharing

The case marking of the SemPs in (2a, b) requires more than the last-resort case assignment summarized in (10). The SemP in (2a) has to bear nominative case, while the one in (2b) may

[4] (10) suggests that this dative marking has no semantic import, contrary to Fortuin (2003).

bear accusative as well as dative case. First, the last-resort account extends to (2b) with no modification when the SemP bears dative case:

(2) b. Ona poprosila ego ne ezdit' tuda *odin/odnomu/odnogo.
she.NOM asked him.ACC NEG go.INF there alone.NOM/DAT/ACC
'She asked him not to go there alone.'

Suppose that (2b) is a control construction with a referential dependency between the controller and controllee. The unexpressed subject of the dependent core fails to receive nominative case from (5a) and has no choice but to bear dative case by appeal to the last-resort procedure.

However, this leaves unaccounted for those cases in which the controller and unexpressed controllee bear the same case morpheme, nominative or accusative, as illustrated in (2a, b). It is important to recall at this juncture that control constructions may involve structure sharing. Suppose that (2a) embodies a structure sharing of the controller and controllee:[5]

(2) a. Vanja xočet prijti odin/*odnomu.
Vanja.NOM wants come.INF alone.NOM/DAT
'Vanja wants to come alone.'

They share both the referential and case values. Given that the controller bears nominative case, it follows that the controllee in (2a) also bears nominative case. The SemP in the dependent core receives nominative case as a consequence of the local agreement with the controllee.

Furthermore, I propose that (2b) may involve *either* a referential dependency between the controller and controllee *or* their structure sharing. To put it differently, I propose to analyze undergoer control constructions such as (2b) as syntactically ambiguous as to whether or not they involve structure sharing, as shown in Table 4.

	Relation between the Controller and Controllee	Shared Values
(2b)	Referential Dependency	Referential Values
	Structure Sharing	Referential and Case Values

Table 4: Syntactic Ambiguity of Undergoer Control Constructions in Russian

When (2b) involves structure sharing of the controller and controllee, they share the case value as well as the referential values. This explains why the unexpressed subject of the

[5] This itself is an assumption with no independent motivation, but see below for why only actor control constructions always involve structure sharing of the controller and controllee.

dependent core of (2b) may bear accusative case, the same case as the one assigned to the controller in the matrix core by (5b). The SemP in (2b) may receive accusative case as a result of the local agreement with the accusative-marked controllee. In contrast, when (2b) involves no structure sharing, the controllee in the dependent core receives dative case by appeal to the last-resort case assignment in (10) and the SemP inherits dative case from the dative-marked controllee.

The above account of (2a, b) leads to an important prediction that when no structure sharing is allowed, the unexpressed infinitival subjects always bear dative case. This prediction is borne out by examples (11a)–(11c) (Landau 2008: 884, 886):

(11) a. Ivan ne znaet kak tuda dobrat'sja *odin/odnomu.
Ivan.NOM NEG know how there reach.INF alone.NOM/DAT
'Ivan doesn't know how to get there by himself.'

b. Ivan poprosil razrešenija prijti na večerinku *odin/odnomu.
Ivan.NOM asked permission come.INF to party alone.NOM/DAT
'Ivan asked permission to come to the party alone.'

c. Ljuda priexala čtoby pokupat maslo *sama/samoj.
Ljuda.NOM came in.order buy.INF butter herself.NOM/DAT
'Ljuda came in order to buy butter herself.'

d. Ljuda priexala pokupat maslo sama/*samoj.
Ljuda.NOM came buy.INF butter herself.NOM/DAT
'Ljuda came to buy butter herself.'

The unexpressed controllees in (11a)–(11c), respectively, occur in a *wh*-infinitive, an infinitive complement of a noun, and a purpose clause introduced by a complementizer *čtoby*. We may see that they receive dative case by appeal to the last-resort case assignment and that the dative marking of the controllers is transmitted to the SemPs that modify them.

(11c,d) constitute a minimal pair with respect to the presence/absence of the complementizer. Since (11d) is an actor control construction, we may derive the nominative marking of the SemP in (11d) from the structure sharing of the controller and controllee. In contrast, it is impossible to posit any structure sharing of the subject argument of the matrix core and the unexpressed controllee in the purpose clause. This is why the SemP in (11c) ends up bearing dative case by appeal to the last-resort case assignment.

Two questions remain to be answered here: why undergoer control constructions such as (2b) optionally involve structure sharing and why actor control constructions such as (2a) always do so. Let us begin with the first question. It has been observed that an accusative-marked SemP as in (2b) is unstable, but that it sometimes replaces the dative-marked counterpart in colloquial speech (Comrie 1974, Babby 1998). These two observations suggest that

the accusative-marked SemP in (2b) is a relatively recent innovation. It is important to note in this connection that in the Polish counterpart to (2b), only the dative-marked SemP would be allowed, while in the Czech, Slovak, and Slovene counterpart, only the accusative-marked SemP would be possible.

(12a)–(12d), respectively, illustrate undergoer control constructions in Polish, Czech, Slovak, and Slovene that have a SemP in their infinitival cores (Comrie 1974: 138, 141; Landau 2008: 917):

(12) a. Polish

Zmusiłem Barbare sprzedać wóz samej/*sama.
forced.1SG Barbara.ACC sell.INF car alone.DAT/ACC
'I forced Barbara to sell the car herself.'

b. Czech

Donutil jsem ho přijít samotného.
forced.1SG AUX.1SG him.ACC come.INF alone.ACC
'I forced him to come alone.'

c. Slovak

Necháva ju starat'sao domácnost' samu.
he.leaves her.ACC look.after.INF housework herself.ACC
'He leaves her to look after the housework herself.'

d. Slovene

Zdravnik jo je poslal delat bolno.
doctor her.ACC AUX.3SG sent work sick.ACC
'The doctor sent her to work sick.'

Comparing (2b) with the Polish, Czech, Slovak, and Slovene counterpart in (12) reveals that undergoer control constructions in Russian are now in the process of change: these control constructions have started to optionally involve structure sharing of the controller and controllee. This means that Russian is midway between Polish, on the one hand, and Czech, Slovak, and Slovene, on the other hand, as shown in (13):

(13) Undergoer Control Constructions in Polish, Russian, Czech, Slovak, and Slovene

Dative	⟶	Dative/Accusative	⟶	Accusative-marked
SemP		SemP		SemP/Adj
Polish		Russian		Czech, Slovak, Slovene

This seems to be a unidirectional process whereby the controllee in the dependent core becomes more of a part of the matrix core and may be interpreted as an instance of *grammaticalization* in which the dependent core (or, more precisely, the unexpressed controllee) is gradually losing its morphosyntactic autonomy.

Finally, the last question that remains unanswered is why actor control constructions always involve structure sharing between the controller and controllee. Comrie (1974) proposes that actor control constructions involve a stronger cohesion between the main verb and embedded infinitival verb than undergoer control constructions and argues that the two verbs form "a particularly cohesive unit" (cf. Babby 1998).[6]

However, this explanation doesn't seem to hold against (14a,b) (Arylova 2006: 52; see Arylova (2006: 44, 79–82) for related discussion):

(14) a. Ja rešil ne xodit' tuda odin.
1SG.NOM decided NEG go there alone.NOM
'I decided not to go there alone.'

b. On staraetsja ne pojavljat'sja na publike odin.
he.NOM tries NEG appear on public alone.NOM
'He tries not to appear alone in public.'

These examples indicate that some actor control constructions allow negation (i.e. one of the core-level operators) to occur in their subordinate cores and that just like undergoer (or, more generally, non-actor) control constructions illustrated by (2b), they involve a core juncture.

Given that the two types of control constructions are not different syntactically, I would suggest that the reason actor control constructions always involve structure sharing between the controller and controllee has to do with the relational status of the controller and controllee. Specifically, non-actor control constructions exhibit a relational mismatch: the controller in the matrix core is a non-PSA argument, while the unrealized controllee is a PSA. They are in contrast to actor control constructions, which involve no such mismatch: both the controller and controllee serve as a PSA in actor control constructions. The contrast between actor control and non-actor control constructions is summarized in Table 5.

I suggest that sharing the PSA status by the controller and controllee in actor control constructions prompts the controllee to share not only the referential values, but also the case value with the controller and that the relational mismatch between the controller and controllee in non-actor control constructions makes it unlikely for the controllee (PSA) to share the case value with the controller (non-PSA) in addition to its referential values. The

[6] The only (albeit indirect) piece of evidence that actor control constructions always involve structure sharing comes from English (not Russian), in which the so-called *wanna*-contraction (e.g. *I wanna go see the movie*) applies to actor control constructions alone.

		Controller	Controllee
Actor control constructions	(2a)	PSA (Actor)	PSA
Undergoer control constructions	(2b)	Non-PSA (Undergoer)	PSA

Table 5: Relational Status of the Controller and Controllee in (2a, b)

logic behind the above argument is that the more functionally similar the controller and controllee are, the more dependent the controllee is on the controller with respect to the referential and case values.

4 Conclusion

The main target of this paper is the second-dative constructions in (2b, c). I have proposed to derive the dative marking of the SemPs in (2b, c) by appeal to the last-resort case assignment in (10), which applies when (5a) fails to apply to the highest-ranking macrorole argument in the dependent core of control constructions:

(5) Case Assignment Rules (Accusative) (Van Valin & LaPolla 1997: 359)

a. Assign nominative case to the highest-ranking macrorole argument.

b. Assign accusative case to the other macrorole argument.

c. Assign dative case to non-macrorole arguments (default).

I have also proposed to attribute the contrast between second-dative constructions as in (2b) and second-nominative constructions as in (2a) to the availability of structure sharing of the controller and unexpressed controllee: when they are subject to structure sharing, the controller and controllee receive the same case (nominative or accusative), while when they aren't, the controller ends up bearing dative case by appeal to the last-resort case assignment.

Finally, I leave it to another occasion to explore whether or not it is possible to extend the present account of the second-dative constructions in (2b, c) to the so-called *dative-infinitive* constructions in (15a)–(15d) (Moore & Perlmutter 2000: 387–389):

(15) a. Mne ne sdat' èkzamen.
me.DAT NEG pass.INF exam.ACC
'It's not (in the cards) for me to pass the exam.'

b. Mne ne rabotat' odnomu.
me.DAT NEG work.INF alone.DAT
'It's not (in the cards) for me to work alone.'

c. Cvetam zdes' ne rasti.
flowers.DAT here NEG grow.INF
It's not (in the cards) for flowers to grow here.'

d. Toj rukopisi ne byt' opublikovannoj zarubežnym
that manuscript.F.SG.DAT NEG be.INF published.F.SG.INS foreign
izdatel'stvom.
publishing house.INS
'It's not (in the cards) for that manuscript to be published by a foreign publisher.'

References

Andrews, Avery D. 1982. The representation of case in Modern Icelandic. In *The Mental Representation of Grammatical Relations*, J. Bresnan (ed.), 427–503. Cambridge, MA: MIT Press.

Arylova, Aysa. 2006. *Infinitival Complementation in Russian*. MA Thesis, University of Tromsø.

Babby, Leonard. 1998. *Subject control as direct predication: evidence from Russian.* In *Formal Approaches to Slavic Linguistics: The Connecticut Meeting*, Ž. Bošković, S. Franks & W. Snyder (eds), 17–37. Ann Arbor, MI: Michigan Slavic Publishers.

Bailyn, John Frederick. 2012. *The Syntax of Russian.* Cambridge: CUP.

Comrie, Bernard. 1974. The second dative: a transformational approach. In *Slavic Transformational Syntax*, R. D. Brecht & C. V. Chvany (eds), 123–150. Ann Arbor, MI: Michigan Slavic Materials.

Fortuin, Egbert. 2003. The conceptual basis of syntactic rules: a semantic motivation for the second dative in Russian. *Lingua* 113(1): 49–92.

Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford: OUP.

Hudson, Richard. 1998. Functional control with and without structure sharing. In *Case, Typology, and Grammar*, A. Siewierska & J. J. Song (eds), 151–169. Amsterdam/Philadelphia: John Benjamins.

—. 2003. Case agreement, PRO, and structure sharing. *Research in Language* 1: 1–33.

Landau, Idan. 2008. Two routes of control: evidence from case transmission in Russian. *Natural Language and Linguistic Theory* 26(4): 877–924.

Moore, John & Perlmutter, David. M. 2000. What does it take to be a dative subject? *Natural Language and Linguistic Theory* 18(2): 373–416.

Nakamura, Wataru. 1999. An Optimality-Theoretic account of the Japanese case system. *Studies in Language* 23(3): 607–660.

Przepiórkowski, Adam & Rosen, Alexandr. 2005. Czech and Polish raising/control with or without structure sharing. *Research in Language* 3: 33–66.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical Categories in Australian Languages*, R. M. W. Dixon (ed.), 112–171. New Jersey: Humanities Press.

—. 1993. Of nominatives and datives. In *Advances in Role and Reference Grammar*, R. D. Van Valin Jr. (ed.), 465–498. Amsterdam/Philadelphia: John Benjamins.

Van Valin, Robert D., Jr. 1991. Another look at Icelandic case marking and grammatical relations. *Natural Language and Linguistic Theory* 9(1): 145–194.

—. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: CUP.

A Role and Reference Grammar Description of the Finnish Partitive

Koen Van Hooste[1]
Heinrich Heine Universität, Düsseldorf

Abstract

This paper attempts to capture the argument-level occurrences of the partitive case in Finnish within the Role and Reference Grammar framework (henceforth: RRG). Partitive case in Finnish and Estonian has been studied quite extensively in the literature (Kiparsky 1998, de Hoop 1992, Nielsen 2007, Miljan & Cann 2013 etc.) and is treated as an instance of differential case marking (DCM). I will primarily draw from Kiparsky's 1998 analysis and Nielsen's 2007 paper and from the RRG-approach to DCM. This paper is conceived of as an exploratory study with a small dataset, not as an exhaustive answer to the complete complexity of the partitive. In summary, I will present an RRG-based formalization as a method for analyzing DCM, specifically the partitive.

1 Introduction

Finnish is a member of the Finno-Ugric language family (along with Estonian and Hungarian amongst others) and exhibits a comparatively large case system that lacks the traditionally present dative while featuring other cases such as the partitive (Karlsson 2004: 32–33). Finnish has a highly agglutinative structure which clearly differentiates between number and case affixes. In its 15-case array, Finnish has a subset of six cases which fulfill (primarily) locative functions. There is no single accusative marker. Rather, accusative is to be understood as a collective designation for several markers that stand in opposition to the partitive (Nielsen 2007: 39 & Karlsson 2004: 122). In other words, there is allomorphy: each marker has specific conditions on its application. For example, the *-t* morpheme is an accusative marker in

[1] 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. I would also like to express my gratitude to Robert Van Valin, Jr., Jasmin Pfeifer, Lea Kawaletz and Jens Fleischhauer for their insightful comments. The general discussion of Finnish is based on Karlsson (2004) unless indicated otherwise.

personal pronouns and noun plurals. In the noun plurals, the accusative is syncretic with the nominative (i.e. the accusative plural marker can be defined as -∅). For nominal singulars, *-n* is the marker for the accusative (syncretic with the genitive). The partitive is marked with *-a/-ä*, *-ta/-tä* or *-tta/-ttä* depending on phonological criteria. As Finnish exhibits vowel harmony, the first morpheme in each set is the back vowel-version; the other is the front vowel-version. In this paper, I will present an overview of the Finnish partitive and its behavior and discuss DCM within the RRG-framework. I will conclude by presenting an RRG-analysis of the Finnish partitive and by proposing modified case assignment rules for Finnish.

I have applied a uniform style of glossing to all examples throughout this paper. This means that the glossing used here is sometimes distinct from the original glossing used by the respective authors.

2 Differential Object Marking

Differential Object Marking (henceforth: DOM) is a phenomenon where in a certain language one set of direct objects is marked in one way and another set of objects is marked in a different way, depending on the features of the object, the verb's referent or the object's referent (Malchukov & de Swart 2009: 345). Aissen (2003: 435) points out that DOM is often driven by animacy and definiteness. Malchukov and de Swart (2009: 341) state that in *fluid case alternations*, the case alternation occurs with one and the same governing verb. This stands in opposition to *split alternations*, where specific case morphology is paired with specific verbs. DOM is a cross-linguistically widespread phenomenon and literature on the topic is vast (Bossong 1985, 1991 & 1997, de Hoop & Malchukov 2008, Malchukov & de Swart 2009). De Hoop & Malchukov (2008) and Aissen (2003), for example, propose an Optimality Theory account. Differential *case* marking (henceforth: DCM) is a term covering multiple types of alternation whereas DOM is only a specific subtype of DCM. DCM and non-canonical argument marking provide interesting challenges to linking-based accounts of grammar, such as RRG. German, for example, exhibits non-canonical argument marking. Verbs like *helfen* (*to help*) take dative-marked direct objects rather than accusative-marked direct objects. Van Valin & LaPolla (1997: 355–356) point out that these German case alternations or Icelandic quirky case (Van Valin 1991, Van Valin & LaPolla 1997: 357ff.) can be handled in terms of exceptional macrorole transitivity. With the proper annotation of M-transitivity, the theory correctly predicts the case in these respective languages. Parallel to capturing certain case alternations in terms of exceptional M-transitivity, case assignment rules can be a possible strategy for dealing with DCM as RRG allows for a considerable degree of language-specific elements or principles in such rules. From Malchukov and de Swart's perspective, the accusative-partitive

alternation in Finnish represents a hybrid situation: the objects of some verbs can be marked with either the accusative or the partitive, depending on the properties of the object (amongst others). However, the direct objects of other verbs can *only* be marked with the partitive (see section 3.3); marking them with the accusative would result in ungrammaticality. I believe that the flexibility in case assignment rules is indeed the proper way to describe the behavior of the Finnish partitive as it exhibits an intricate interplay of multiple factors, rather than just one (e.g. exceptional M-transitivity).

3 Functions of the partitive

Kiparsky (1998) and de Hoop (1992) characterize the partitive as a structural-semantic *hybrid case*. It exhibits features of structural cases (Blake's terminology, see Blake 2001) – i.e. cases that mark grammatical relations on the verb's arguments – and of semantic cases. The latter cases mark adjuncts for peripheral information. An example would be the locative case that many languages feature in their inventory. Nielsen (2007: 37) characterizes the partitive and the accusative in Finnish as the cases that are used to mark an object. To account for the choice of case, I will employ data from Kiparsky (1998) and Nielsen (2007) to draw up a decision tree (see section 3.3). Based on this tree, I will then attempt to design case assignment rules for the Finnish partitive. Determining how the partitive works and how those workings can be accounted for in RRG can shed light on how case assignment rules function in Finnish in general. Kiparsky (1998: 267 & 286) states that the partitive case in Finnish has two functions: An *NP-related function* and an *aspect-related function* (Nielsen calls these the *quantity distinction* and the *aspectual distinction*, respectively). Nielsen (2007: 41) also introduces a third function, however. He calls this the *actuality distinction*. All three will be discussed in the following sections.

3.1 NP-related function of the partitive

The NP-related function of the partitive can be explained by looking at an intrinsically telic verb like *get*: The objects of intrinsically telic verbs are marked with the partitive when they are quantitatively indeterminate (Kiparsky 1998: 268).[2] In all other circumstances, the object appears in the accusative.

[2] The concept of boundedness is complex and subject to criticism. In this paper, I follow Nielsen (2007) in equating aspectual boundedness with telicity.

(1) Saa-n karhu-n
get-1SG bear-ACC
'I'll get the bear'

(2) * Saa-n kahta karhu-a
get-1SG two(PART) bear-PART
'I'll get two bears'

(3) Saa-n karhu-j-a
get-1SG bear-PL-PART
'I'll get bears'

In example (1), the quantity of *bear* is determined (one) and thus the object appears in the accusative. In example (3), the quantity is not determined (one only knows this quantity to be plural in nature) and thus the object stands in the partitive. Using the partitive in example (2) is ungrammatical, as the quantity is determined (two bears). Nielsen (2007) uses a Functional-Grammar-analysis (henceforth: FG) to capture the occurrences of the partitive. As such, he uses the FG-notion of operators (Dik 1997). The marking used to express quantity distinctions is, in FG, triggered by so-called ω-operators, and more specifically, the *quantifier* type (Nielsen 2007: 53). In FG, quantification involves the relation between three so-called ensembles (Dik 1997: 170 & Nielsen 2007: 53). The reference ensemble (R) is the ensemble that the term in question refers to. This R is a subset of the domain ensemble (D). U is the universal ensemble. Nielsen (2007: 53) states that with an expression such as *five books*, the five books are a subset of U (all the books). In this case, D is said to be identical to U. Ergo, the relation that *five books* expresses is R$\subset$D=U. If one were to say *three of the five books*, the relation can be characterized as R$\subset$D$\subset$U (Nielsen 2007: 53). He theorizes that it is grammatically relevant in Finnish to decide whether there is identity between R and D or whether this relation is not specified. If an object is quantitatively bounded (in the terminology used here: quantitatively determined), then the expressed relation is R=D. If an object is quantitatively unbounded (in the terminology used here: quantitatively undetermined), then the relation can be characterized as R$\subseteq$D. The former relation triggers accusative case, the latter triggers partitive case. I follow Nielsen's approach for my definition of quantitative determinacy or [$\pm$QD]:

(4) [+QD]: R=D
[-QD]: R$\subseteq$D

My characterization of [$\pm$QD] can be considered to be part of the properties of the RP (Reference Phrase). However, it is possible to account for certain [-QD] contexts another way.

Multiple argument activity verbs with a non-referential second argument do not have an undergoer macrorole (Van Valin 2005: 63–64). This is because the undergoer-RP refers to a non-instigating, affected participant. *Beer* in *John drank beer* is non-referential. It does not refer to a specific entity and therefore cannot receive the undergoer macrorole. If it cannot become undergoer, it must become the non-macrorole argument (henceforth: NMR). NMRs in case languages are marked with the default case (Van Valin 2005: 110). Should one posit the partitive as default case form (in absence of the dative), sentences like (5) can be explained:

(5) Pekka juo olut-ta
Pekka drink.PRS.3SG beer-PART
'Pekka drinks beer' (Karlsson 2004: 105)

In (5), *beer* is non-referential: it becomes the NMR and therefore receives partitive case by default. This is a general background rule of RRG (Van Valin 2007: 63–64) and therefore such verbs do not need to be marked with an [MR1] label. However, concluding that a function of the partitive is to mark NMRs does not explain the occurrence of the partitive as the marker for the second argument in achievements (or any other *Aktionsart* for that matter):

(6) Pekka löy-si vet-tä
Pekka find-PST.3SG water-PART
'Pekka found water' (Nielsen 2007: 41)

Here, the verb is fully transitive, taking two macroroles; yet the undergoer is marked with the partitive as well. Hence, one cannot simply account for the occurrences of the partitive by positing it as the default case for NMRs. Further evidence against such an approach is that three-place predicates (e.g. *give*) mark the NMR with the allative case (primarily a local case):

(7) Minä anna-n lahja-n vaimo-lle-ni
1SG.NOM give-PRS.1SG gift-ACC wife-ALL-POSS.1SG
'I give a gift to my wife' (adapted from Karlsson 2004: 138 & 154)

This issue concerns the complex concept of the default case in Finnish. From the examples in (6) and (7), it is clear however that it would be premature to generalize the partitive as default case for NMRs.

3.2 Aspectual function of the partitive

Kiparsky (1998: 267ff.) explains the aspectual function of the partitive as being characterized by either *resultativity* or *boundedness*. Nielsen (2007: 45) points out that his concept of boundedness is comparable to *telic* vs. *atelic* in other descriptions. Ackerman & Moore

(1999) also state that the telicity of the verb influences case assignment. Following Nielsen and Ackerman & Moore, I will assume that telicity is central to the aspectual workings of the partitive. The basic idea then is that accusative appears with telic predicates and partitive with atelic predicates. Consider (Ackerman & Moore 1999: 1):

(8) Matti ost-i maito-a (tunni-n)
Matti buy-PST.3SG milk-PART (hour-ACC)
'Matti bought milk (for an hour)'

(9) Matti ost-i maido-n (tunni-ssa)
Matti buy-PST.3SG milk-ACC (hour-INE)
'Matti bought the milk (in an hour)'

In the first sentence, *buy* is an activity, similar to *Dana ate fish* (Van Valin 2005: 33). It passes the *X for an hour*-test, as the authors indicate with *tunnin*. As an activity, it is atelic. In the second sentence, the verb is an active accomplishment, similar to *Dana ate the fish*. Here the verb passes the *X in an hour*-test. As an active accomplishment, the object of this verb receives the accusative. Nielsen (2007: 45) states that if the situation is bounded (i.e. telic), the accusative appears. If it is unbounded (i.e. atelic), the partitive appears (Nielsen 2007: 45). I will now test this hypothesis with several examples: Consider the sentence in (10) (Kittilä & Malchukov 2009: 549):

(10) Henkilö tappo-i karhu-n
person.NOM kill-PST.3SG bear-ACC
'The person killed a bear'

The verbal predicate *kill* is a causative accomplishment and as such is telic. As Kittilä and Malchukov indicate, using the partitive here would not be grammatical. Consider the following example (Kittilä & Malchukov 2009: 549):

(11) Henkilö jo-i maido-n/maito-a
person.NOM drink-PST.3SG milk-ACC/milk-PART
'The person drank the milk/some milk'

In the second reading, the verb is an activity (and thus atelic). In the first, it is an active accomplishment (and thus telic). One could claim that the accusative is replaced with the partitive when the telic interpretation of the predicate shifts towards an atelic one. This would fall in line with the claim that the [±telic] feature in the matrix of the verbal predicate determines the marking of the second argument in an M-transitive predicate. This approach also accounts for the example in (5): The verb is an activity (and thus atelic) and therefore receives partitive

marking. Kiparsky (1998: 281–282) provides a (non-exhaustive) list of verbs that can have a variable telicity reading. Verbs with both readings as a possibility appear to be common in Finnish (Kiparsky 1998: *ibid.*). He offers an example with the verb *shoot*. This verb is said to belong to a class of verbs that assign the two different cases to their object, depending on the aspectual interpretation. Consider (Kiparsky 1998: 267):

(12) Ammu-i-n karhu-n
shoot-PST-1SG bear-ACC
'I shot the bear (dead)'

(13) Ammu-i-n karhu-a
shoot-PST-1SG bear-PART
'I shot at the bear'

The latter example features an activity predicate (i.e. atelic) and is translated as 'to shoot at'. The whole process of killing the bear by shooting is not complete. The former counts as an active accomplishment and is thus telic (it is entailed in Finnish that the bear has been killed). As such the object receives accusative case. There are, however, examples of telic verbs that can receive the partitive on their object (Nielsen 2007: 45) which cannot be explained by referring to the quantity function:

(14) Perhe raken-si möki-n
family build-PST.3SG cottage-ACC
'The family built a cottage'

(15) Perhe raken-si mökki-ä
family build-PST.3SG cottage-PART
'The family was building a cottage'

In RRG terms, *build* is an active accomplishment and is thus telic. As such it receives the accusative case in (14). However, in (15), the object is marked with the partitive. RRG is capable of capturing this in its conceptual system. Consider the logical structure of the sentence in (14):

(16) **do'** (perhe, [**build'** (perhe, mökki)]) & INGR **exist'** (mökki)

Active accomplishments are telic, as is clear from their INGR-component. However, with a progressive reading, the INGR-component is cancelled. Accomplishments (BECOME) are compatible with the progressive: cancelling INGR leaves PROC (Van Valin 2005: 44). I therefore theorize that the desired progressive reading cancels INGR, changing the verb class from

active accomplishment to activity. Such a conversion necessarily changes the telicity in the matrix from [+telic] (active accomplishment) to [-telic] (activity). The telicity account therefore remains intact: in absence of any quantity related case assignment, atelic predicates assign partitive, telic predicates assign accusative.

3.3 Priority of the two functions

So far, this description does not adequately explain the interaction between these two distinctions (or functions). Apart from describing them in their own right, it is equally important to describe how they interact. Which function has priority over the other? For these purposes I define function A having priority over B when A applies in more contexts: If distinction B is irrelevant when A has a value x, then A has priority over B. All the examples discussed thus far point in the direction of the clean division proposed above: if the verb has an atelic reading, then the object will appear in the partitive. If the verb has a telic reading, the object will appear in the accusative. There are, however, verbs that cannot 'shift' and are intrinsically atelic or telic. *Etsiä* ('seek') belongs to the class of intrinsically atelic verbs. Such verbs thus only assign partitive case (Kiparsky 1998: 268):

(17) Etsi-n karhu-a
seek-1SG bear-PART
'I am looking for the/a bear'

(18) Etsi-n karhu-j-a
seek-1SG bear-PL-PART
'I am looking for (the) bears'

One would expect that the intrinsically telic verbs always require accusative. However, the discussion of the NP-function made clear that this is not the case. Consider the following verbs that are intrinsically telic:

(19) Saa-n karhu-t
get-PRS.1SG bear-ACC.PL
'I'll get the bears' (Kiparsky 1998: 268)

(20) Saa-n karhu-j-a
get-PRS.1SG bear-PL-PART
'I'll get bears' (*ibid.*)

The verb in (19) is telic and receives the accusative. It is, however, possible to mark the object in question with a partitive, as example (20) illustrates. This marking does not change the aspectual interpretation: the verb is still telic. The difference lies in the interpretation of the

object-RP. In (20), the quantitatively indeterminate reading of *bears* changes the case assignment to partitive rather than maintain the accusative, as aspect would predict. In other words, the semantic-pragmatic reading of the object overrides aspectually governed case assignment. The reverse is not possible. If a verb is atelic, the object properties become irrelevant:

(21) * Ets-i-n karhu-t
seek-PST-1SG bear-ACC.PL
'I looked for the bears' (Kiparsky 1998: 268)

(22) Ets-i-n karhu-j-a
seek-PST-1SG bear-PL-PART
'I looked for the bears/ bears' (*ibid.*)

In (22), the object can either have a quantitatively determined or an indetermined reading. This distinction cannot be expressed morphologically. As the RP is already marked with the partitive, [-QD] cannot be shown (as one would replace the partitive with a partitive). It is also impossible to replace the partitive with the accusative as example (21) illustrates. Ergo, by replacing an accusative with a partitive, one can show that the RP is [-QD]. One cannot, however, make clear that the RP is [+QD]. In other words, case substitution is asymmetrical: The accusative can only be replaced with a partitive, but not vice versa. This asymmetry gives rise to ambiguities if the verb is atelic. With a verb like *ampua* ('shoot') a three-way ambiguity can exist, since such a verb can be atelic or telic. Kiparsky (1998: 268) points out that *ammuin karhuja* can mean that the speaker (a) shot at bears, (b) shot bears (dead) or (c) shot at the bears. Here, the verb can be [+telic] with the RP read as [-QD] or the verb can be [-telic] with a [-QD] RP or the verb can be [-telic] with the RP being [+QD]. In *I shot at bears*, the [-telic] value blocks any potential morphological realization of the [+QD] property of the RP. In summary: if the verb is atelic (irrespective of RP properties), it obligatorily has a partitive object. This means that the quantity function is only relevant in a subset of cases. The decision tree in Figure 1 clearly shows that aspect has priority over quantity.
Nielsen (2007: 56ff.) arrives at a similar conclusion within his Functional-Grammar-approach ((23)–(25); (25) shows the FG-analysis of (15)):[3]

(23) Pekka löy-si heti vede-n
Pekka find-PST.3SG right away water-ACC
'Pekka found the water right away'
Operators: [[Pf [+DomID [water]]]] → resulting case: ACC

[+telic] → [+QD] → resulting case: ACC

[3] Nielsen (2007: 53ff.) uses perfective and imperfective as values for the aspectual distinction. Both in his analysis and in mine, boldface marks the partitive trigger. My analysis is inserted beneath his.

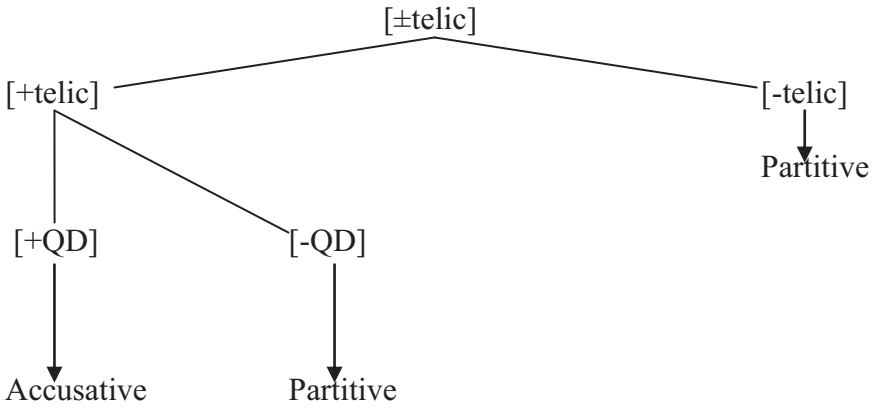


Fig. 1: Decision tree for the Finnish accusative-partitive alternation (preliminary).

(24) Pekka löy-si heti vett-ä
Pekka find-PST.3SG right away water-PART
'Pekka found water right away'
Operators: [[Pf [-**DomID** [water]]]] → resulting case: PART

[+telic] → [**-QD**] → resulting case: PART

(25) Operators: [[**Impf**[+DomID [cottage]]]] → resulting case: PART

[**-telic**] → [+QD] → resulting case: PART

One can draw the preliminary conclusion that [-telic] always triggers partitive marking and accusative marking is only obtained if a [+telic] predicate also happens to have a [+QD] argument-RP. This is echoed by Nielsen (2007: 40–41) when he states that accusative is the marked choice for object case and that "the partitive always wins". In all other cases, partitive is assigned, with certain ambiguities as a result.

It would be erroneous to equate quantity with *definiteness*. This would indeed be highly problematic (Nielsen 2007: 41). In a sentence like 'I shot a bear dead' the object is marked with the accusative (Kiparsky 1998: 268), even though the RP has an indefinite reading. Despite being indefinite, it is still quantitatively determined (*a bear* is indefinite, but the quantity of bear is one). Therefore, it would count as [+QD] and thus receive the accusative.[4] And this is precisely what happens. In this instance, there is ambiguity with respect to definiteness: *Ammuin karhun* can be read as 'I shot the bear dead' or 'I shot a bear dead'.

Kiparsky (1998: 280–283) divides Finnish verbs into three classes: (1) Verbs that are telic, (2) verbs that are atelic and (3) a class of verbs that can be telic or atelic. The first class of verbs (like *saada* – 'get') assigns accusative to the objects, unless the object is quantitatively inde-

[4] The fact that *bear* is a count noun in the singular automatically makes it [+QD].

terminate in which case the partitive is assigned. In other words, telic verbs assign accusative, unless the object is quantitatively indeterminate. Here, the NP-function seems to overrule the default assignment principle. The second class of verbs (like *etsiä* – 'seek') assigns partitive to the objects, regardless of whether the object is quantitatively determinate or not. The third class of verbs (switch verbs like *ampua* – 'shoot') can be [+] or [–] telic. If they are telic, then accusative is assigned, unless the object is quantitatively indeterminate (then the partitive is assigned). Ergo, when they are telic, these verbs function the same way the verbs in the first class do. When these verbs are atelic, then the object is always in the partitive, irrespective of its quantitative determinacy. This can be summarized as follows:

a) [-telic] → assign partitive
   [+telic] → assign accusative by default; rule b) (case substitution) may apply.

b) Replace case marking on the argument with partitive case marking if the argument is read as quantitatively indeterminate.

3.4 Actuality distinction of the partitive

Nielsen (2007: 37, 41 & 47ff.) states that a third function of the partitive can be identified. He calls this the *actuality distinction*. He believes that this distinction (following Dik 1997: 242) is based on a continuum ranging from a purely asserted proposition to a purely negated one (Nielsen 2007: 41). Such a continuum is said to contain several points (Nielsen 2007: 47):

(26) certain – probable – possible – improbable – impossible

Polarity (negative or positive) is to be understood as the extreme points of this scale. The selection of the Finnish object case (apart from the other two functions) depends on epistemic objective modality which is assigned to the propositional content (Nielsen 2007: 47–48). He divides this scale into *actual modality* and *non-actual modality* and theorizes that actual clauses trigger accusative while non-actual clauses trigger partitive. Such an assumption would not only explain negated sentences, but also questions expecting a negative answer (in both cases, the object stands in the partitive) as both questions and negation are on the non-actual side of the scale.
Consider:

(27) En osta auto-a
NeV.PRS.1SG[5] NeLV car-PART
'I do not buy a car' (Karlsson 2004: 102)

[5] This notation is my own: NeV = negation verb, NeLV = Negation-Lexical Verb

(28) Tuskin Kaisa tappa-a hiir-tä
unlikely Kaisa kill-PRS.3SG mouse-PART
'It is unlikely that Kaisa kills a/the mouse' (Nielsen 2007: 48)

(29) Ol-et-ko teh-nyt yhtään hyvä-ä teko-a?
be-2SG-Q do-PTCP single good-PART deed-PART
'Have you done a single good deed?' (Nielsen 2007: 48)

Finnish uses a verb for negation, as exemplified in (27) (Karlsson 2004: 85–86). This negation verb is inflected for person. It is followed by the stem of the lexical verb that is to be negated. The marking of tense and aspect in negative constructions is very complex and beyond the scope of this paper. In the case of (29), Nielsen (2007: 48–49) points out that the expected negative answer to this question makes the interpretation of the proposition non-actual. This (expected) non-actuality then triggers partitive marking on the object. In (28) the proposition is less non-actual than in (29) or (27), yet it is clearly on the non-actual side of the scale in (26). However, a clear-cut distinction between [+actual] triggering accusative and [-actual] triggering partitive is too simple and fails to explain many instances of the partitive. Take sentences (12), (13) and (19), (20) into consideration. These sentences are all actual, yet the second one in each pair is marked with the partitive. Nielsen further expands his FG analysis by adding actuality as an operator with scope over both the aspect distinction and the quantity distinction (Nielsen 2007: 57):

(30) Kaisa ei tapa hiir-tä
Kaisa NEV.PRS.3SG NELV mouse-PART
'Kaisa does not kill a/the mouse'
Operators: [**NonAct**[Pf[+DomID [mouse]]]] → resulting case: PART

With non-actual sentences, there does not seem to be a choice with respect to the object case. If a sentence is non-actual, then the object is partitive. This is similar to the telicity distinction: If the verb is atelic, the object always appears in the partitive. It would thus appear that Nielsen's characterization that the "partitive always wins" is true to some degree. Nielsen's scope analysis, with the actuality distinction included, necessitates the rewriting of the decision tree in Figure 1. In terms of the revised decision tree, there are more branches that lead to partitive than to accusative.

The revised decision tree in Figure 2 now covers more contexts in which the partitive appears:

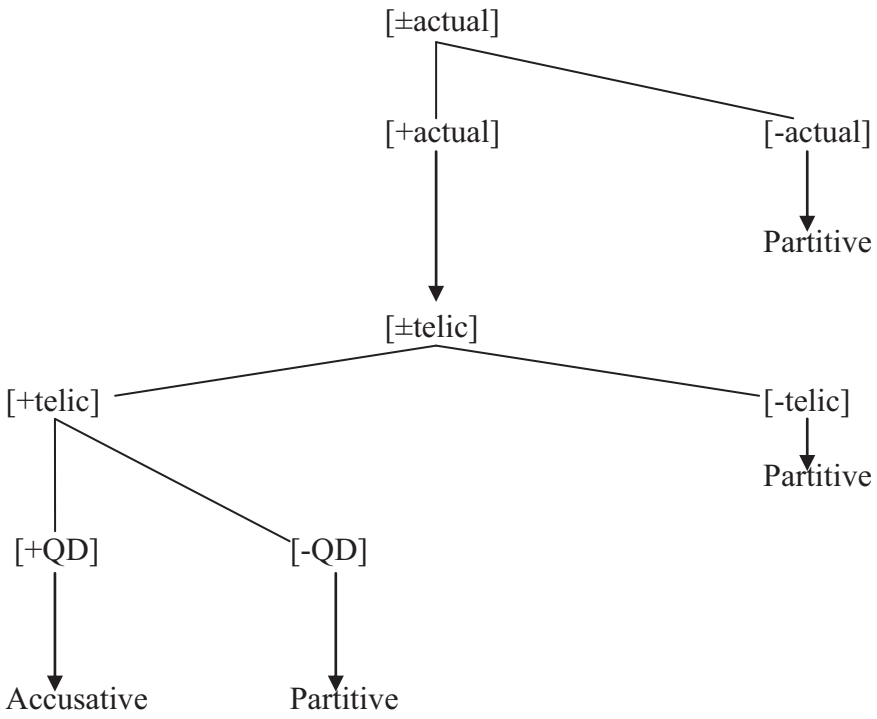


Fig. 2: Revised decision tree for Finnish case assignment.

I theorize that the aspectual distinction and the quantity distinction are best captured with case assignment rules. I believe that this so-called third function can best be handled in RRG by placing the triggers discussed by Nielsen in the domain of the operators in the operator projection. I propose to treat this distinction as a part of the clausal operator *status*. The status operator in RRG covers external negation and epistemic modality (Van Valin 2005: 9ff.). I therefore contend that, in these cases, the partitive is simply the Finnish-specific expression of the clausal operator *status* modifying the layered structure of the clause. The continuum in (26) is relevant as the driver of *status*. As a general rule in the grammatical system of Finnish, case assignment by rules can be overridden by the expression of the status operator. The layered structure of the clause for the sentence *I do not buy a car* is represented in Figure 3. Nielsen (2007: 49) also lists a fourth function of the partitive (example (31)), pointing out that it is a very marginal one: the downtoning of *illocutionary force* (henceforth: IF).

(31) Ote-taan paita-a pois
take-PASS shirt-PART away
'Let's take our shirt off, shall we?' (*ibid.*)

I believe it is possible to treat such an occurrence of the partitive similarly to the third function. This so-called downtoning can be considered as a modifier of IF. Finnish might have a broad

conceptual range of distinctions in terms of illocutionary force and modifications thereof. As we have established, the partitive is not only a case form governed by properties of the predicate and of the RP itself, but also the morphological expression of the status operator in the layered structure of the clause. I believe it is therefore not unlikely that IF-modification can (under certain circumstances) be expressed by means of the partitive. As the example in (31) was the only one of its kind listed by Nielsen, this remains an open question and warrants further investigation.

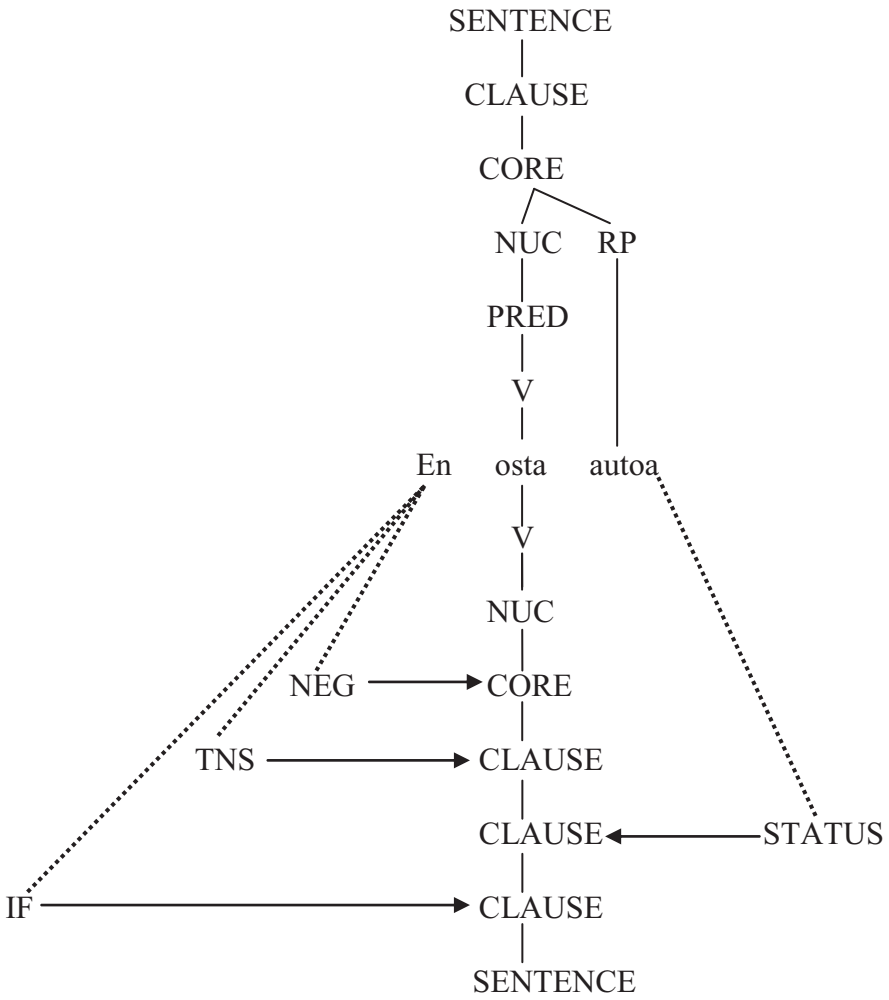


Fig. 3: Layered Structure of the Clause for *En osta autoa*.

In section 3.1, the possibility of the partitive as default case for NMRs was raised. As (7) illustrated, allative case can mark the NMR as well. However, examples (5) and (11, second reading) suggest that partitive marking does not depend on macrorole-hood. Bearing in mind that the functions driving the partitive case are subject to priority relations amongst themselves, one could theorize that [-actual], [-telic] and [-QD] simply have priority over other

principles (such as default case for NMRs). The question of the default case in Finnish falls outside the scope of this paper, however, and is a promising topic for future research.

4 The proposed solution of VP-level expression of boundedness

Kiparsky (1998: 266) theorizes that the aspectual and the NP-related function of the partitive can be unified by stating that the partitive licenses unboundedness at the VP-level. RRG does not have the concept of a VP. Furthermore, if one assumes that the VP is the domain of the partitive, one requires at least two explanations (or a very complicated one) of partitive behavior: Finnish notably also has partitive subjects. These cannot be explained by placing the solution in the domain of the VP, as VPs do not contain a subject position. Kiparsky (1998: 285–286) proposes two rules that capture all instances of the partitive and thus unify the two functions:

a) A VP argument is unbounded if it either has an unbounded head or an unbounded argument.

b) The object of an unbounded VP is obligatorily partitive.

Kiparsky (1998: 269) defines *(un)boundedness* as follows: "A predicate is intrinsically unbounded if it can be modified by degree adverbs such as (some) more, a lot, very much, a bit, somewhat less, considerably, slightly, referring to the extent of a single eventuality." He stresses the importance of *gradability*, stating that "bounded predicates, telic or atelic, admit of no degree". Therefore, according to Kiparsky, bounded predicates cannot be modified by degree adverbs and are not gradable whereas unbounded predicates *are* gradable. This clearly indicates that Kiparsky's boundedness is a concept that either overarches telicity or one that encompasses more than just telicity. Kiparsky proposes a test for determining which predicates are bounded and which ones are not. In order to determine the class of predicate, one simply has to add the degree adverbs from the definition above and check whether the resulting sentence is grammatical or not. This test also works for determining whether or not an NP predicate is bounded (Kiparsky 1998: 270):

(32) a lot of coffee (unbounded NP-predicate)

(33) * a lot of bear (bounded NP-predicate)

In my view, the definition of boundedness provided by Kiparsky is a tool for testing boundedness rather than a proper definition. Kiparsky (1998: 271–272) also points out that the lack of such a definition is why many (if not most) Finnish linguists consider the partitive

as having two fundamentally distinct functions rather than one unified function. I do not follow Kiparsky's idea of boundedness as a grammatical concept. Boundedness could perhaps be conceived of as a vague background concept with cognitive significance in the Finnish language system with separate interpretations (and repercussions) in the verbal and nominal domains. However, I do not believe it directly translates to a grammatical concept. I therefore propose to maintain the solution that has been put forth in previous sections.

5 The issue of partitive subjects

Even though the (vast) majority of partitive marked arguments are objects, both Kiparsky (1998: 297) and Nielsen (2007: 40) acknowledge the existence of partitive marked subjects (see differential subject marking, Malchukov & de Swart 2009: 351–353). Kiparsky points out that they only appear with intransitive verbs. It would indeed appear that there can only be one partitively marked argument in the core at any given time. A sentence with two such arguments is ungrammatical (Kiparsky 1998: 297–298):

(34) * karhu-j-a sö-i hunaja-a
bear-PL-PART eat-PST.3SG honey-PART.SG
'There were bears eating honey'

(35) Piha-lla leikk-i laps-i-a
courtyard-ADE play-PRS.3SG child-PL-PART
'There are children playing in the courtyard'

Nielsen (2007: 40) points out that partitive subjects occur in existential clauses when the argument in question is quantitatively unbounded (or, in the terminology used here: [-QD]). In (35) the number of children is not specified, therefore, the RP is quantitatively indeterminate. However, there is a further constraint: partitive subjects only appear with certain intransitive verbs, a class which Kiparsky (2008: 297) calls *presentational verbs* (*existential verbs* in Nielsen's account). Consider for instance (Kiparsky 1998: 298): In Finnish *There are children playing in the playground* or *A child has been born* are grammatical but **There are children smiling in the playground* is not.

In RRG terms, one could theorize that the occurrence of the partitive case is limited to the core. Nor can there be more than one partitively-marked argument in each core. Partitive subjects can only appear with intransitive verbs if the reading of the RP is one of quantitative indeterminacy (i.e. only the NP-related function is possible in the case of partitive subjects). I believe it is possible to treat this phenomenon as a grammatical construction concerning existential constructions, where partitive marking is part of the morphosyntactic information

of the constructional schema. That could also account for the altered word order. Kiparsky (1998: 300) states that the basic position of the partitive subject is VP-internal. Judging from his examples, partitive subjects appear postverbally. Nielsen's example (2007: 40) also features a postverbal subject:

(36) Tässä on leipä-ä
DEM.PROX.INE be.PRS.3SG bread-PART
'Here is some bread'

It is however possible to maintain the subject preverbally, although it does not seem to be the neutral position. If we treat partitive subjects as an occurrence of quantitative indeterminacy linked to existential utterances captured in constructional schemas, then a reference to the partitive suffices. Evidence in favor of a constructional schema comes from the fact that such verbs very often have the option to have a nominative subject as well. Consider (Kiparsky 1998: 297):

(37) karhu-t kuol-i-vat
bear-NOM.PL die-PST-3PL
'The bears died'

(38) karhu-j-a kuol-i
bear-PL-PART die-PST.3SG
'Bears died'

In (37), the verb agrees with the Privileged Syntactic Argument (or: PSA) in number (both plural). In (38), the verb is in the singular, but the PSA is in the plural. In such a construction, the finite verb is always third person singular (Karlsson 2004: 100). Syntactic templates for this construction are given in Figure 4 and the constructional schema is given in Figure 5.[6]

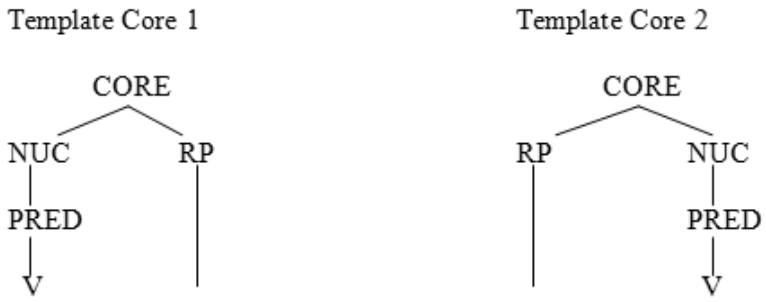


Fig. 4: Finnish core templates.

[6] "4.15a" in the schema refers to the principles governing PSA-assignment in accusative constructions (Van Valin 2005: 100): the highest ranking direct core argument in terms of the PSA-selection hierarchy becomes the PSA.

Construction: Finnish presentational construction (indeterminate)	
Syntax:	template: core 1 (default), core 2 PSA: 4.15a
Morphology:	PSA: partitive case Verb agreement: 3SG
Semantics:	PSA is [-QD] part of existential/presentational state of affairs
Pragmatics:	unspecified

Fig. 5: Constructional schema for 'partitive subjects'.

6 Problems with the telicity account

There are some issues with the telicity account presented here. Due to the exploratory nature of this paper, they will remain unresolved. I nevertheless briefly wish to highlight them and explore a possible explanation for them. Kiparsky (1998: 283) points out that there are some atelic verbs that take accusative objects and a class of telic verbs that take partitive objects (regardless of the nature of the interpretation of the RP). As far as the first class of verbs is concerned, it is entirely possible that they are, in fact, not members of the same *Aktionsart*-class as their English equivalents. It is common for verbs in different languages describing the same state of affairs to belong to different *Aktionsart*-classes (Van Valin & LaPolla 1997: 103ff.). For example, *omistaa (own)* is said to assign case as if it were an achievement or an accomplishment rather than a state. If, with the proper *Aktionsart*-testing, such a verb turns out to pattern like an achievement or an accomplishment, then case assignment would follow naturally: Both accomplishments and achievements are telic and therefore their objects are marked with the accusative. The object could then appear in the partitive as a result of the NP-related function. Consider (Kiparsky 1998: 283):

(39) Näe-n häne-t
see-1SG 3SG-ACC
'I see him/her'

See usually tests as a state (Van Valin 2005: 46). However, it is entirely possible that in Finnish the verb see rather tests as an achievement. If this were true, then *come to see* would be a closer approximation of *nähdä* than *see*. Likewise, *come to own* would be closer than *own* in the case of *omistaa*. Kiparsky (1998: 280ff.) presents a list of verbs as members of this verb class which include: *realize (oivaltaa)*, *acknowledge (myöntää)* and *notice* (*huomata,*

havaita, keksiä). This hypothesis is very preliminary and extensive testing of this verb class is required. Due to the limited data set available and the exploratory nature of this study, this hypothesis remains to be proven and is an interesting topic for future research.

The objects of the second class of verbs receive partitive, whereas they should receive accusative, like *punish* (Kiparsky 1998: 284). According to Kiparsky (1998), this verb passes the standard tests for accomplishments (for example, the in-PP test):

(40) rankais-i-t he-i-tä
punish-PST-2SG 3-PL-PART
'You punished them'

However, Kiparsky only has this one verb as the example for the class as a whole. One might hypothesize that this verb used to be (in earlier stages of the language) a pure activity meaning *to inflict harm motivated by past offenses*. With such a reading, the object would appear in the partitive (as the verb would be atelic). Later, the verb might have shifted to accomplishment or active accomplishment, without the case marking making the matching shift. However, this theory needs diachronic data in order to be substantiated.

7 Rewriting the rules

Based on Kiparsky's analysis (1998), the case assignment rules for the Finnish partitive could be potentially rewritten as follows:

a) Assign nominative case to the highest-ranking macrorole argument (in terms of the actor-undergoer hierarchy).

b) Assign partitive case to the lowest-ranking argument if the verb is [-telic]. If the verb is [+telic], assign accusative by default unless the conditions specified in c) apply. If so, apply rule c).

c) Replace case marking on the lowest-ranking argument with partitive case iff it is read as quantitatively indeterminate.

Such a set of rules, however, is faced with two problems: 1) it tacitly assumes the partitive to be the marked choice for the object case (which is contradicted by the revised decision tree in Figure 2) and 2) it does not capture the actuality distinction discussed in section 3.4. It is possible – by way of refinement – to integrate the actuality distinction and simplify the rules considerably, by assuming that the accusative is the marked choice, rather than the partitive. In my analysis, it is crucial to place the actuality distinction in the domain of the operators *status*

and (possibly) *illocutionary force*. Case 'assignment' driven by operators is to be treated as an override of the normal case assignment principles. I justify the override metaphor by referring to the fact that actuality has clausal scope contrary to the *aspect* (a predicate property) and *quantity* (an RP-property) distinctions. Consider the refined case assignment principles for the Finnish partitive:

a) Assign nominative case to the highest-ranking macrorole argument (in terms of the actor-undergoer hierarchy).

b) Assign accusative case to the lowest-ranking argument iff (1) the predicate is [+telic] and (2) the argument-RP is [+QD]. Otherwise, assign partitive case.

These revised rules do not conflict with the constructional schema in Figure 5.

8 Conclusion

I have argued that there are three distinct functions that trigger partitive marking on the lowest-ranking argument. I have also determined that the complex interplay of these functions can be explained by priority relations: *actuality* > *aspect* > *quantity*. I have attempted to write case assignment rules for a Finnish transitive predication. In order to do this, I have translated the priority hierarchy into RRG-compatible notions: *Quantity* was treated in terms of [±QD], a property of the RP. *Aspect* was treated in terms of [±telic] in the feature matrix of the verb. Lastly, *Actuality* is a grammatically relevant category for Finnish and was treated as the opposition between non-actual clauses and actual clauses. Actuality is analyzed as a Finnish-specific form of the clausal operator *status*. In this case, partitive is the expression of a non-actual value of *status*. These distinctions function on the basis of binary oppositions. A negative choice on a decision node triggers the partitive on the object. A positive choice necessitates further specification (i.e. a further decision node) until the lowest level in the tree is reached. Assuming that the partitive is rather the unmarked and the accusative the marked choice is certainly justified from the point of view of the decision tree. It is this tree that I used to draw up priority relations between functions and, ultimately, the case assignment rules in section 7. I have also shown that, due to the general complexity of the partitive case, designing Finnish-specific case assignment rules is an apt approach to capture and explain the partitive's behavior. I have presented evidence that suggests that macrorole-hood of the second argument is irrelevant for the distinctions driving partitive marking. For example, the [-telic] feature (and its subsequent marking) seems to have priority over the non-referential status of a given argument. However, as Nielsen (2007) points out throughout his paper, such

distinctions or functions are not easily unified on the semantic level. Nielsen (2007: 60) hints at a general transitivity contrast behind the accusative-partitive opposition, with accusative expressing high transitivity and partitive expressing low transitivity. Answering such a question is beyond the scope of this paper. It is therefore a potential area of future research. Another question warranting further study is Nielsen's fourth function. Is downtoning of Illocutionary Force with the partitive describable in terms of modifying IF and if so, how can that be captured? There are also complex, accusative RPs that internally license partitive to lower-level elements. Kiparsky (1998: 267, also see Miljan & Cann 2013 for Estonian) refers to such instances as internally assigned partitives. In order to fully grasp the nature of the partitive, such RPs also require investigating.

References

Ackerman, Farrell & Moore, John. 1999. Telic Entity as a Proto-Property of Lexical Predicates. In *LFG99 Online Proceedings*, M. Butt & T. H. King (eds). Stanford: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG4-1999/lfg99-toc.html

Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory* 21(3): 435–483.

Blake, Barry J. 2001. *Case*. Cambridge: CUP.

Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

—. 1991. Differential object marking in Romance and beyond. In *New Analyses in Romance Linguistics, Selected Papers from the XVIII Linguistic Symposium on Romance Languages 1988*, D. Wanner & D. Kibbee (eds), 143–170. Amsterdam: Benjamins.

—. 1997. Le marquage différentiel de l'objet dans les langues d'Europe. In *Actance et Valence dans les Langues d'Europe*, J. Feuillet (ed.), 193–258. Berlin: de Gruyter.

De Hoop, Helen. 1992. *Case configuration and noun phrase interpretation*. Groningen: Grodil.

De Hoop, Helen & Malchukov, Andrej. 2008. Case-Marking Strategies. *Linguistic Inquiry* 39(4): 565–587.

Dik, Simon C. 1997. *The Theory of Functional Grammar*. Berlin: Mouton de Gruyter.

Karlsson, Fred. 2004. *Finnische Grammatik*. Hamburg: Helmut Buske Verlag.

Kiparsky, Paul. 1998. Partitive Case and Aspect. In *The projection of Arguments: Lexical and Compositional Factors*. M. Butt & W. Geuder (eds), 265–307. Stanford: CSLI Publications.

Kittilä, Seppo & Malchukov, Andrej. 2009. Varieties of accusative. *In The Oxford Handbook of Case*, A. Malchukov & A. Spencer (eds), 549–562. Oxford: OUP.

Malchukov, Andrej & de Swart, Peter. 2009. Differential case marking and actancy variations. In *The Oxford Handbook of Case*, A. Malchukov & A. Spencer (eds), 339–355. Oxford: OUP.

Miljan, Merilin & Cann, Ronnie. 2013. Rethinking case-marking and case alternation in Estonian. *Nordic Journal of Linguistics* 36(3): 333–379.

Nielsen, Peter Juul. 2007. Finnish object case. A unified description of multifunctionality in a sign contrast. *Acta Linguistica Hafniensia: International Journal of Linguistics* 39(1): 37–68.

Van Valin, Robert D., Jr. 1991. Another Look at Icelandic Case Marking and Grammatical Relations. *Natural Language and Linguistic Theory* 9: 145–194.

—. 2001. *An Introduction to Syntax*. Cambridge: CUP.

—. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

—. 2009. Case in Role and Reference Grammar. In *The Oxford Handbook of Case*, A. Malchukov & A. Spencer (eds), 102–120. Oxford: OUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax. Structure, meaning and function*. Cambridge: CUP.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca: Cornell University Press.

The Syntax-Semantic Interface in Italian Result-Oriented Argument Structures

Anna Riccio
University of Naples "L'Orientale"

Abstract

Resultatives are a productive linguistic phenomenon in a number of languages (e.g., in English or German), but they are limited in Romance languages (e.g., in Italian or French). This paper presents an analysis of Italian constructions which receive a result-oriented interpretation, in particular adjectival phrases (APs). Relatively few studies have focused on investigating such constructions in Italian literature. We mention the well-known work by Napoli (1992), entitled "Secondary resultative predicates in Italian". Among the more recent studies are those by Cattaneo (2008) and Ramchand (2008).

The aim of this study is to show that the syntax-semantics interface developed by Role and Reference Grammar [RRG] (Van Valin & La Polla 1997; Van Valin 2005) provides a useful descriptive and theoretical framework for an investigation of result-oriented APs. The resulting analysis reveals the syntactic and semantic properties of Italian result-oriented APs that equate them with a specific set of English resultative APs.

1 Introduction

Resultative constructions are a productive cross-linguistic means to indicate the result of an action described by the main verb, often accompanied by an extension of argument structure. A set of resultatives are illustrated in (1a–f) from English:

(1) a. I wiped the USB drive clean.
b. The joggers ran the pavement thin.
c. John danced himself breathless.
d. The river froze solid.
e. The cook cut the tomatoes into small pieces.
f. The cup broke into many pieces.

The examples in (1a–c) and (1e) present the fixed word order, i.e., V + NP + XP, in which the latter constituent is expressed by resultatives, such as adjectives and a prepositional phrase, following the direct object. Differently, both the verbs in (1d) and (1f) have the intransitive structure, i.e., V + XP.

Most of the scientific literature on resultatives argues that canonical constructions, such as those exemplified in (1a–d), are unavailable in Romance languages.[1] Consider example (2) from Italian, one of most commonly cited in literature:

(2) * Gianni martellò il metallo piatto.
John hammered the metal flat
'John hammered the metal flat.' (Broccias 2003: 2)

The Italian transitive sentence corresponding to the word-by-word translation from English is ungrammatical, but the alternative construction consisting of the manner verb *appiattire* 'flatten' followed by the prepositional phrase *a colpi di martello* 'by hammering it' is accepted, as illustrated in (3):

(3) Gianni appiattì il metallo a colpi di martello.
John flattened the metal with strokes of hammer
'John flattened the metal by hammering it.' (Broccias 2003: 2)

The unaccusative construction in (1d) is also ungrammatical in Italian, as shown in (4):

(4) * Il fiume è ghiacciato solido.
'The river froze solid.' (Napoli 1992: 72)

Froze solid means 'completely hard'. In the Italian construction, the concept of 'the river gets very cold and changes into ice' is expressed by *è ghiacciato*. As a matter of fact, *solid* in (4) may also be optional in English, whereas in other cases the presence of the resultative AP (or PP) is required in order for the utterance to be grammatical (e.g., *James ran *(his feet sore)*, in Boas (2003: 7)).

Italian allows constructions which correspond to those in (1e–f), i.e., resultative prepositional phrases, as in (5a–b) and (6a–b):

(5) a. Ho spinto il pianoforte nella sala da pranzo.
'I pushed the piano into the dining room.'

[1] See Green (1973), Grevisse (1980), Rothstein (1985), Talmy (1985), Jackendoff (1988), Merlo (1988, 1989), Bosque (1990), Demonte (1991), Pustejovksy (1991), Tenny (1994), Snyder (1995), Legendre (1997), Washio (1997), Morimoto (1998), Demonte and Masullo (1999), Folli (2000), Mateu (2000), Folli and Harley (2005, 2006), Folli and Ramchand (2001, 2002, 2005), Cattaneo (2008), Ramchand (2008), Shibagaki (2010), among others.

b. Ho calciato la palla nell'angolo.
'I kicked the ball into the corner.'

(6) a. Sabato ho dipinto una parete di blu.
'Saturday I painted a wall blue.'

b. Ho tagliato il pane a fette sottili.
'I cut the bread into thick slices.'

Examples in (5) are locational PPs and those in (6) are state PPs; each one is perceived as a result of the previous action.

Italian generally lacks adjectival phrases like those in (1a–d). Nevertheless, a web-based research allowed for the identification of Italian APs, especially in substandard varieties, as illustrated in (7a–f):

(7) a. Sabato ho dipinto le pareti bianche, ad eccezione di una parete di blu
Saturday have.1SG painted the walls white except for a wall of blue
profondo.
deep
'Saturday I painted the walls white, except for a wall [I painted] deep blue.'

b. Gianni ha tosato l'erba corta.
John has mowed the.grass short
'John mowed the grass short.'

c. Cuocere i ravioli in abbondante acqua salata; scolarli ben asciutti.
boil the ravioli in abundant water salted drain.them well dry
'Boil the ravioli in abundant salted water; drain them dry.'

d. L'erba cresce alta.
the.grass grows high
'The grass grows high.'

e. Maria ha tagliato i capelli corti.
Mary has cut the hair short
'Mary cut her hair short.'

f. Il mio macellaio taglia la carne sottile.
the my butcher cuts the meat thin
'My butcher cuts the meat thin.'

The data in (7a–f) have been obtained by using Italian corpora and questionnaires testing native speakers' intuitions, who agree on grammaticality judgments, although they show a great deal of judgment variation concerning acceptability.[2]

One of the reasons for the unacceptability/ungrammaticality of the examples in (7) is that Italian is classified as a *verb-framed language*, and, thus, it systematically lacks verb-particle

[2] The corpora used for obtaining data are ItTenTen10, CORIS/CODIS, and the WebCorp. Forty Italian native speakers, aged 20–25, were tested individually on the grammaticality and acceptability of the data, i.e., 20 Italian constructions formed by [SN + V + NP + XP] and [SN + V + XP].

constructions and other constructions that are commonly analyzed as complex predicates in *satellite-framed languages* (Snyder 2001). Namely, Mateu (2000) attributes the contrast between languages without resultatives and those with them to the fact that the former are 'verb-framed' in the sense of Talmy (1991), while the latter are 'satellite-framed.' For instance, in the cases of directed motion events, verb-framed languages generally express some meaning associated with a path on the verb, whereas satellite-framed languages like English have the tendency to express the path by means of a PP (including particles). The latter constructions are not typical of Romance languages. But the situation is more complex than it appears. Simone (2008) argues that Italian is the only language in which the motion path can be encoded in a separate particle (adverb) associated to the verb (*satellite*).[3] This claim challenges the previous hypothesis, and could then account for the result-oriented interpretation for the APs in Italian, even if they are still "spurious" resultative constructions in any case (Washio 1997).

The existence of result-oriented adjectival phrases in Italian can be justified by the application of the Manner and the Predicative tests. Let us consider the following sentence in (8):

(8) I ragazzi camminarono veloci.
the boys walked quick
'The boys walked quickly.'

We can paraphrase it into *The walking event was quick*, i.e. 'there is an event which consisted of walking and was quick' (cf. Levinson 2010). The Manner test is passed. The appropriate answer to the question "How was the walking event?" is "The walking event was quick.". We cannot say the same for example (7e) (repeated below for convenience):

(9) Maria ha tagliato i capelli corti.
Mary has cut the hair short
'Mary cut her hair short.'

If the Manner test were successful, then *corti* would refer to the event of cutting, i.e. 'there is an event which consisted of cutting and which was short(?)' (cf. Levinson 2010). However,

[3] See the following example from Quaderni di Serafino Gubbio operatore (The Notebooks of Serafino Gubbio cinematograph operator), Pirandello (1916): "C'è una carrozzella che corre davanti. Pò, popòòò, pòòò. Che? La tromba dell'automobile la tira indietro ? Ma sì! Ecco pare che la faccia proprio andare indietro, [...] seguita a dare indietro, indietro, con le case, gli alberi, i rari passanti, [...]" [There is a one-horse carriage in front. "Peu, pepeeeu, peeeu" What? The horn of the motor-car is pulling it back? Why, yes! It does really seem to be making it run backwards, [...] continues to fall behind, far behind, with the houses, the trees, the occasional pedestrians, [...] Transl. by C.K. Scott Moncrieff, 2012], in Simone (2008).

we cannot paraphrase example (9) into *The cutting event is short*.[4] Unlike *veloci*, *corti* passes the Predicative test, i.e. *The hair became short*.[5]

The examples in (7) show transitive result-oriented AP constructions, excluding (7d) which represents an unaccusative argument structure. The verb *cresce* 'grows' is an incremental verb which may identify an accomplishment or achievement structure (but not necessarily). All verbs in (5), (6), and (7) (usually *change-of-state verbs*) convey information about the end result state of the postverbal NP, but without the result predicate this end remains vague in some regard. In this sense, without the result-oriented phrase or the proper contextual background information in a sentence like *Maria ha tagliato i capelli*, it is not clear what the exact final state of *capelli* is. Thus, the added predicate (i.e., *corti* 'short') causes the vague endpoint of the event to become more precise or intensifies the action of the verb by denoting the metaphorical end state of the Patient/Theme argument. Napoli (1992) claims that these constructions are acceptable in Italian, even if the activity has to be iterative or durative before the desired effect on the object is achieved.

The paper is structured as follows. Sections 2 compares the semantic and syntactic features of Italian result-oriented APs with those of English APs. Section 3 describes the semantic and syntactic features of Italian result-oriented APs within the RRG framework. Section 3.1. proposes the logical structures (LSs) that reflect an activity causing change of state with results. Section 3.2 illustrates the layered structure and the operator projection of the Italian

---

[4] The Manner test also fails with the following constructions taken from Campanian dialects, as in (ia) and (ib):

i a. Comm sta vestut bell
how is dressed nice
'How well he dressed!'
Manner test: ?The dressing event was beautiful
Predicative test: He became nice.

b. Tu stai vestu' serie' e sistimat.
you are dressed serious and orderly
'You are dressed neat and tidy.'
Manner test: ?The dressing event was serious/unsmiling and tidy.
Predicative test: He became serious.

[5] As a matter of fact, the verb *cut* belongs to the domain of action, and then it lexicalizes both manner (how the movement/activity is carried out) and result. For instance, both the constructions in (iia) and (iib) involve an activity such that a particular manner of action directly leads to a particular state:

ii a. (En) John cut the bread thin.

b. (It) Gianni ha tagliato il pane sottile.

result-oriented APs. Section 3.3. shows how the macrorole assignment proceeds in such constructions and illustrates their semantics-to-syntax linking. Section 4 concludes this paper.

2 Semantic and syntactic features of resultatives in Italian and English

Resultatives are semantically distinguished into *strong* vs. *weak* constructions (Washio 1997). In strong resultatives "the meaning of the verb and the meaning of the adjective are completely independent of each other" (Washio 1997: 7), whereas in weak resultatives the meaning of the predicate is lexically entailed in the meaning of the verb, and it is possible to predict the result of the action from the semantics of the verb. The distinction between weak vs. strong resultatives is not clearcut. It is likely that the same construction can be interpreted both as weak and strong resultative. This is the case with the verbs of removing in (10a–b). According to Washio (1997), both the examples represent weak resultatives, whereas they are strong resultatives in Kaufmann and Wunderlich's (1998) terms:

(10) a. The girl scrubbed the floor clean.
b. John has wiped the table clean.

What unifies both the sentences in (10) is the event-argument homomorphism theory driving from the observation of the interaction between verbal semantics and the meaning of resultative predicates (Wechsler 2001). Homomorphism means that parts of the verbal event must correspond to parts of the change event and vice versa. Chao Li (2008) refers to such constructions as *subcategorized weak resultatives*, whose objects are subcategorized by the causing predicate. When the object is not subcategorized, resultatives are unsubcategorized (e.g., *Bryan drank Shannon under the table*). The subcategorized weak resultatives are also distinguished from *subcategorized strong resultatives*. See examples (11a–b):

(11) a. The waiter wiped the table clean. (subcategorized weak resultative)
b. The boy hammered the metal flat. (subcategorized strong resultative)

According to Chao Li (2008), the subcategorized weak resultatives are more acceptable than the subcategorized strong resultatives, which in turn are more acceptable than the unsubcategorized resultatives (cf. *implicational hierarchies*).

Referring to the examples in (7a–f) in Section 1, they could be considered as subcategorized weak resultatives consisting of verbs which denote change and imply a result state. In (7a), the resultative reading is accounted for by the event-argument homomorphism theory. In

other words, the adjective *bianche* is predicated of the NP which already exists in the verbal frame (i.e., 'to paint'). *Dipingere* 'paint' is a verb of coloring followed by a (non-) gradable adjective. So, parts of the event of painting correspond to parts of the event of the wall's becoming white, i.e., the path metonymically signaled by the adjective *white* (which denotes its final point). The meaning of the verb *dipingere* 'paint' may then imply that the color of the object *le pareti* 'the walls' is changed. This also occurs for example (7d), in which the process involves an 'underlying object', since the meaning can be interpreted as 'the grass became high as a result of the process of growing'. The same phenomenon takes place for the English sentences in (11a–b). Since the theme argument is shared, *metal* in (11b) is a subcategorised object of the verb, then homomorphism between the verbal event and the change event is required. Such constructions are then formed by verbs which express an activity which implies that a patient undergoes a change in state as a result of the agent's action. Thepkanjana and Uehara (2009) call them 'implied-result' verbs. [6] It is therefore acceptable to claim that the resultative reading in the Italian constructions is also accounted for by some semantic and syntactic similarities with English resultatives.

3 An RRG analysis of Italian Result-Oriented Constructions

Role and Reference (RRG) provides a valuable framework for analyzing Italian result-oriented constructions semantically and syntactically. The semantic representation of the clause deals with the logical structures (LS) of verb classes and the argument positions that come with each verb class (i.e., predicate-argument structure). The syntactic representation involves the layered structure of the clause (LSC) consisting of the core argument/s and the periphery. The linking system provides the bidirectional mapping procedure from semantics to syntax and vice versa, which is applied to some discussed grammatical phenomena (e.g. case marking, voice assignment).

3.1 Semantic Representation

The term "resultative" is used in two different ways in cross-linguistic studies. It can be distinguished in terms of Vendler's (1967) verbal aspect categories as stative resultative in (12a) and accomplishment (or achievement) resultative in (12b):

(12) a. The wall is painted.
b. John painted the wall blue.

[6] In this regard, if a situation or action implies another thing, the other thing is likely to exist or be true (i.e., implied-result verbs, such as *wipe*). On the contrary, if a situation or action entails a particular thing, it involves having or doing that thing (i.e., entailed-result verbs, such as *kill*) (Thepkanjana and Uehara 2009).

In (12a), we observe an inherently unbounded, non-dynamic state of affairs, while in (12b) there is a bounded process with a non-punctual/punctual change of state.

Result-oriented constructions like *Maria ha tagliato i capelli corti* in (9) have a causative meaning, since the verb specifies an argument that is both patient and theme. The agent by acting on a patient/theme effects some change of state in it. Referring to example (9), when *Maria ha tagliato i capelli*, the *capelli* 'hair' must have a cut in it, the cut must have been affected by her acting on it in a certain way (viz., by moving an object into a contact with it; see Levin, 1985; Hale and Keyser, 1986, 1987). The logical structure that accurately reflects an activity causing change of state with results, as given in (13):

(13) **do'**(Maria, [**cut'**(Maria, capelli)]) CAUSE [BECOME **short'**(capelli)]

The LS shows a complex structure consisting of a predicate indicating the causative action or event (usually an active predicate), linked to a predicate indicating the resulting state of affairs by an operator-connective CAUSE. We exactly deal with a causative active accomplishment meaning by the addition of a resultative that specifies the state achieved by NP. Thus, the causative active accomplishment combined with a resultative phrase can be viewed as resulting in an achievement (Mateu 2000): BECOME **short'** is an achievement, whose anterior state holds when the cutting process is going on.

Recently, Van Valin (this volume) proposes the decomposition of the logical operator BECOME into PROC[ess] & INGR in order to capture the 'measuring out' properties of what Dowty (1991) calls the 'incremental theme'. The incremental nature of cutting in *Maria ha tagliato i capelli corti* is simultaneous ('∧') to the process of shortening, and both the processes constitute a single complex event which achieves a termination expressed by the RESULT part of the template ('INGR **be-short'**), as illustrated in (14a–b):

(14) a. **do'**(x, [**cut'**(x, y)]) ∧ PROC **shorten'**(y)] CAUSE [INGR **be-short'**(y)]

b. [**do'**(Maria, [**cut'**(Maria, capelli)]) ∧ PROC **shorten'**(capelli)] CAUSE [INGR **be-short'**(capelli)]

The argument *capelli* 'hair' is assigned two main properties: (1) "having been cut by Maria" (which includes the simultaneous process of shortening), and (2) "becoming/being short". The former event is described by the primary predicate, and the latter by the secondary one; both the predicates denote a single event.

3.2 Syntactic Representation

The layered structure of example (9) is illustrated in Figure 1:

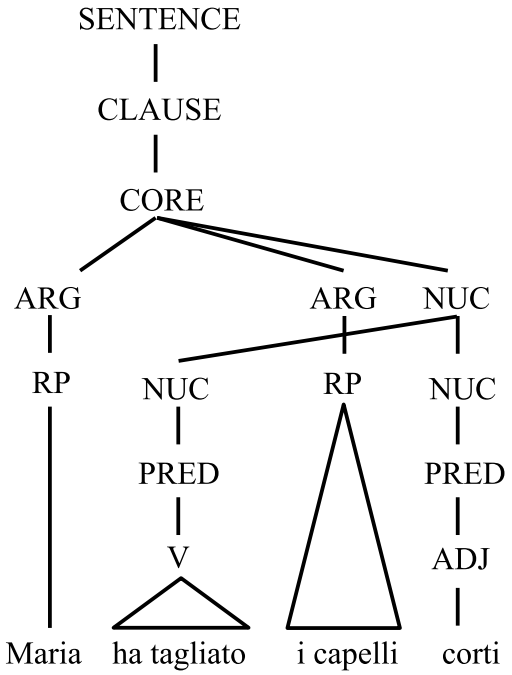


Fig. 1: The constituent structure for (9)

The figure represents a nuclear juncture in which the two nuclei (i.e., the verb *ha tagliato* 'cut' and the adjective *corti* 'short') form a single, complex nucleus with a single set of core arguments. The verb can thus be said to consist of two subevents, i.e., an activity and a resulting state acting together in a nuclear juncture.

The RRG theory of clause linkage allows for a cosubordination nexus relation in which the two nuclei of equivalent sizes are layered together in a coordinate-like relation with no marker of syntactic dependency between and among them (Olson 1981). The linked nuclei are "dependent upon the matrix unit for expression of one or more of the operators" (Van Valin 2005: 201). The semantic representation of (9) is illustrated in (15):

(15) ‹$_{\text{IF}}$DECL‹$_{\text{TNS}}$PRES‹$_{\text{ASP}}$‹PERF [**do'**(Maria, [**cut'**(Maria, capelli)])
∧ PROC **shorten'**(capelli)] CAUSE [INGR **be-short'**(capelli)] $_{\text{ASP}}$›$_{\text{TNS}}$›$_{\text{IF}}$›

The aspect operator has scope over the cutting action (*ha tagliato* = NUC) and the resulting event (*corti* = NUC), and indicates that the cutting and the shortening are completed. The operator projection is represented in Figure 2.

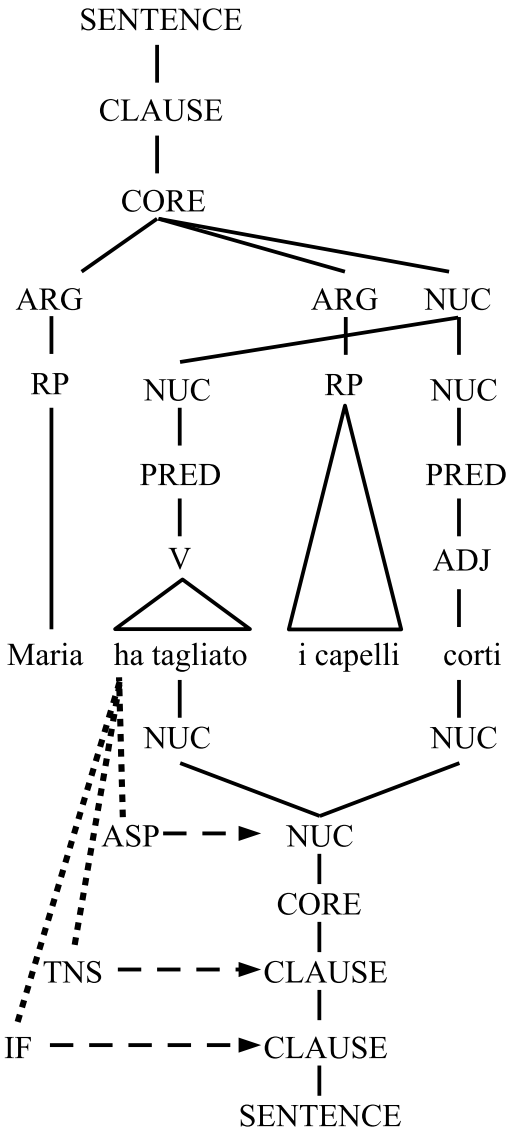


Fig. 2: Nuclear cosubordinate structure for (9)

3.3 Linking Semantics and Syntax

The mapping between the decompositional semantic representation of verbs and the morphosyntactic expressions of their arguments is mediated by macroroles [MRs] (actor and undergoer). RRG assigns MRs to verbal arguments on the basis of the actor-undergoer hierarchy [AUH] and the default macrorole assignment principle (see Van Valin & LaPolla 1997).

Figure 3 presents a detailed view of how the semantics-to-syntax linking proceeds in example (9).

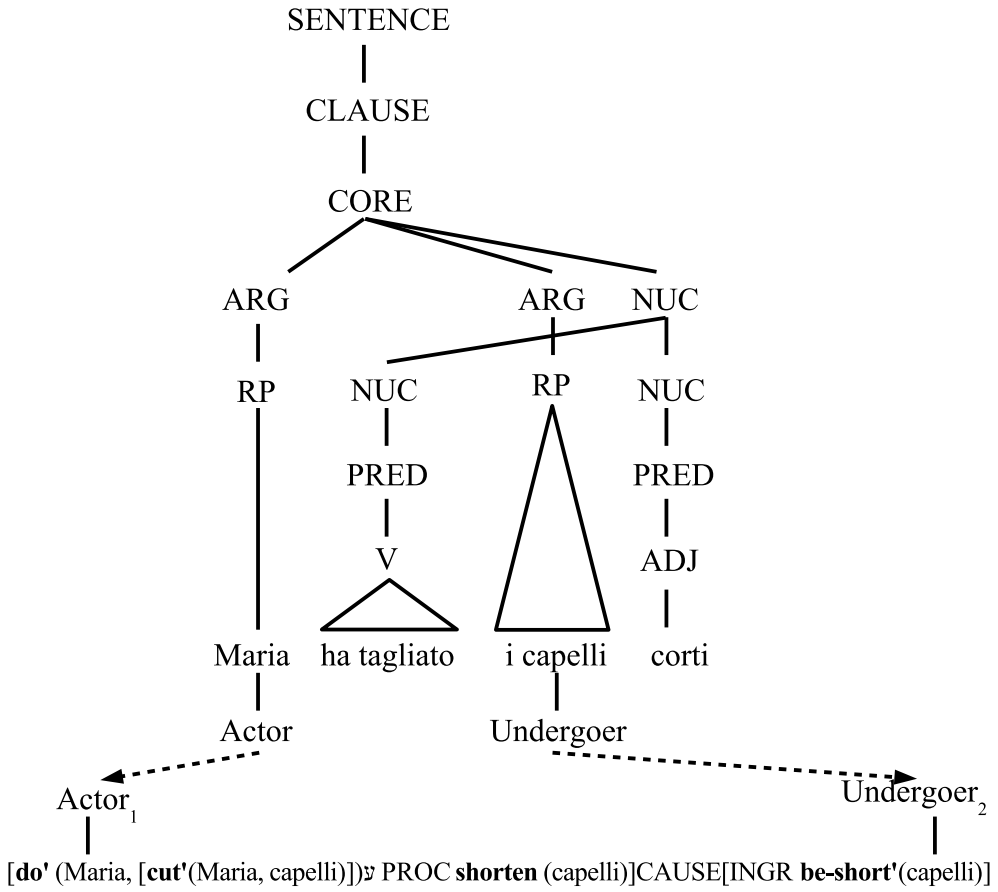


Fig. 3: The semantics-to-syntax linking in example (9)

The argument *Maria* is selected as actor of the first PRED (i.e., the activity predicate *ha tagliato*). It is the privileged syntactic argument because it participates in the initiation of the causal chain. The argument *capelli* is selected as Undergoer of the second PRED (i.e., the result predicate *corti*), and it participates in the endpoint of the causal chain. Because *Maria* and *capelli* are assigned two and three argument positions in the logical structure, respectively, all these arguments are not distinctive.

4 Conclusion

This paper has provided an overview of characteristic features of Italian result-oriented APs, and has shown that the application of the RRG framework to the analysis of such constructions reveals their semantic and syntactic properties. By exploiting the parallelism with the English resultatives, semantic and syntactic similarities have been shown between English weak resultatives and Italian result-oriented constructions (Section 2). In both languages, (subcatego-

rized) weak resultative APs are specifiers, intensifiers or degree modifiers of the action of the verb. Therefore, the lack of strong resultatives in Italian wouldn't prevent us from speaking of resultative semantics in Romance languages.

The semantic and syntactic analysis of Italian result-oriented constructions APs within the RRG framework (Section 3) has revealed logical structures that reflect an activity causing change of state with results (Section 3.1), and the RRG theory of the clause linkage has described their syntactic status consisting of nuclear cosubordinate junctures (Section 3.2).

Finally, an RRG account of the semantics-to-syntax linking in Italian result-oriented APs has been described on the basis of the semantic and syntactic properties. It has been shown that such constructions with nuclear-level juncture involve an actor, which participates in the initiation of the causal chain, and an Undergoer, which participates in the endpoint of the causal chain (Section 3.3.)

References

Boas, Hans C. 2003. *A constructional approach to resultatives*. Stanford: Center for the Study of Language and Information.

Bosque, Ignacio. 1990. Sobre el aspecto en los adjetivos y en los participios. In *Tiempo y aspecto en español*, I. Bosque (ed.), 177–214. Madrid: Cátedra.

Broccias, Cristiano. 2003. The cognitive basis of adjectival and adverbial resultative constructions. *Annual Review of Cognitive Linguistics* 12/2(1): 103–126.

Cattaneo, Andrea. 2008. L'interazione tra clitici soggetto e oggetto. Talk given at the *XIV Giornata di dialettologia*, June 20, Padua.

Demonte, Violeta. 1991. Temporal and aspectual constraints on predicative AP's. In *Current Studies in Spanish Linguistics*, H. Campos & F. Martínez-Gil (eds), 1–34. Washington: Georgetown University Press.

Demonte, Violeta & Masullo, Pascual José. 1999. La predicación: los complementos predicativos. In *Gramática Descriptiva de la Lengua Española, vol. 2*, I. Bosque & V. Demonte (eds), 2461–2523. Madrid: Espasa.

Folli, Raffaella. 2000. *Constructing telicity in English and Italian*. ESSLI workshop on *Paths and telicity in event structure*, Ms., Oxford University.

Folli, Raffaella & Harley, Heidi. 2004. Consuming results: flavors of little-v. In *Aspectual enquiries*, P. Kempchimsky & R. Slabakova (eds), 1–25. Dordrecht: Kluwer.

—. 2005. Event-path homomorphism and the accompanied-motion reading in motion causatives. In *Proceedings of WECOL 2004*. http://www.fresnostate.edu/artshum/linguistics/documents/WECOL%202004%20VOLUME%2016.pdf

—. 2006. On the licensing of causatives of directed motion: Waltzing Matilda all over. *Studia Linguistica* 60: 121–155.

Folli, Raffaella & Ramchand, Gillian. 2001. Getting results: motion constructions in Italian and Scottish Gaelic. In *Proceedings of WCCFL 20*, K. Megerdoomian & L. A. Bar-el (eds), 101–114. Somerville, MA: Cascadilla Press.

—. 2002. Event structure composition: the case of goal of motion and resultative constructions in Italian and Scottish Gaelic. In *Proceedings of the Perspectives on Aspect Conference*, Utrecht.

—. 2005. Prepositions and results in Italian and English: an analysis from event decomposition. In *Perspectives on aspect*, A. Van Hout, H. De Swart & H. J. Verkuyl (eds), 81–105. Dordrecht: Springer.

Green, Georgia M. 1973. A syntactic syncretism in English and French. In *Issues in Linguistics*, B. B. Kachru, R. B. Lees, Y. Malkiel, A. Pietrangeli & S. Saporta (eds), 257–278. Urbana, Ill.: University of Illinois Press.

Grevisse, Maurice. 1980. *Le bon usage* (11th edition). Paris: Duculot.

Hale, Ken & Keyser, S. Jay. 1986. Some transitivity alternations in English. In *Lexicon Project Working Papers 7*, Center for Cognitive Science. Cambridge, MA: MIT Press.

—. 1987. *Some Transitivity Alternations in English* [Lexicon Project Working Papers 10] Center for Cognitive Science at MIT. Cambridge, MA: MIT Press.

Jackendoff, Ray S. 1988. Why are they saying these things about us? *Natural Language and Linguistic Theory* 6: 435–442.

Kaufmann, Ingrid & Wunderlich, Dieter. 1998. *Cross-linguistic patterns of resultatives*. Ms., University of Düsseldorf.

Légendre, Geraldine. 1997. Secondary predication and functional projections in French. *Natural Language and Linguistic Theory* 15: 43–87.

Levin, Beth. 1985. Lexical semantics in review: an introduction. In *Lexical semantics in review. Lexicon Project Working Papers 1*, B. Levin (ed.), 1–62. Cambridge, MA: MIT Press.

Levinson, Lisa. 2010. Arguments for pseudo-resultative predicates. *Natural Language &. Linguistic Theory* 28: 135–182.

Li, Chao. 2008. *Mandarin Resultative verb compounds*. München: LINCOM Europa.

Mateu, Jaume. 2000. Why can't we wipe the slate clean? A lexical-syntactic approach to resultative constructions. *Catalan Working Papers in Linguistics* 8: 71–95.

Merlo, Paola. 1988. Secondary predicates in Italian and English. In *Proceedings of the 5. Eastern States Conference on Linguistics*, J. Powers & K. de Jong (eds), 338–348. Columbus: Ohio State University.

—. 1989. Risultative in italiano e in inglese. *Rivista di Grammatica Generativa* 14: 29–53.

Morimoto, Yuko. 1998. *Los verbos de movimiento en español: una aproximación léxico conceptual*. PhD dissertation, Universidad Complutense de Madrid.

Napoli, Donna Jo. 1992. Secondary resultative predicates in Italian. *Journal of Linguistics* 28: 53–90.

Olson, Michael Leon. 1981. *Barai clause junctures: toward a functional theory of interclausal relations*, PhD dissertation. Canberra: Australian National University.

Pirandello, Luigi. 1925. *Quaderni di Serafino Gubbio operatore*. Milano: Garzanti. (*Shoot! The Notebooks of Serafino Gubbio, Cinematograph Operator*, transl. by C.K. Scott Moncrieff. 2012. Pisa: Aonia edizioni)

Pustejovsky, James. 1991. The syntax of event Structure. *Cognition* 41: 47–82.

Ramchand, Gillian Catriona. 2008. *Verb meaning and the lexicon. A first phase syntax*. Cambridge: CUP.

Rothstein, Susan Deborah. 1985. *The Syntactic Forms of Predication*. Bloomington, Indiana: Indiana University Linguistics Club.

Shibagaki, Ryosuke. 2010. Mandarin secondary predicates. *Taiwan Journal of Linguistics* 8(1): 57–93.

Simone, Raffaele. 2008. I verbi sintagmatici come costruzione e come categoria. In *I verbi sintagmatici in italiano e nelle varietà dialettali. Stato dell'arte e prospettive di ricerca. Atti delle giornate di studio Torino, 19–20 febbraio 2007* M. Cini (ed) 13–30. Frankfurt am Main: Peter Lang.

Snyder, William. 1995. A neo-davidsonian approach to resultatives, particles, and datives. *NELS* 25: 457–471.

—. 2001. On the nature of syntacic variation: Evidence from complex predicates and complex word-formation. *Language* 77: 324–342.

Talmy, Leonard. 1985. Lexicalization patterns: semantic structures in lexical forms. In *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon*, T. Shopen (ed.), 57–149. Cambridge: CUP.

—. 1991. Path to realization: a typology of event conflation. *Proceedings of the Berkeley Linguistic Society* 17: 480–519.

Tenny, Carol. 1994. *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.

Thepkanjana, Kingkarn & Uehara, Satoshi. 2009. Resultative constructions with "implied-Result" and "entailed-Result" verbs in Thai and English: a contrastive study. *Linguistics* 47(3): 589–618.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Van Valin, Robert D., Jr. & La Polla, Randy J. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: CUP.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.

Washio, Ryuichi. 1997. resultatives, compositionality and language variation. *Journal of East Asian Languages* 6: 1–49.

Wechsler, Stephen. 2001. An analysis of English resultatives under the event-argument homomorphism model of telicity. *Proceedings of the 3rd Workshop on Text Structure*. University of Texas at Austin. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.9.958&rep=rep1&type=pdf

Activity Hierarchy and Argument Realization in (R)RRG

Rolf Kailuweit
Heinrich-Heine-Universität Düsseldorf

Abstract

The actor-undergoer hierarchy hits its limits when it comes to describing the macrorole assignment of several verb classes. This chapter deals with four problematic cases: causation, three-place state predicates, two-place object-experiencer verbs, and finally, transitive versus intransitive verbs of motion. Working out ideas in Kailuweit (2013), I propose the activity hierarchy as an alternative to the actor-undergoer hierarchy. Starting from a set of two actor features (c = causation/control, m = mental) and one undergoer feature (r = resultative), the activity hierarchy brings into play semantic criteria that have dominated the discussion on generalized semantic roles in the last decades. The activity hierarchy allows for a sufficient number of activity degrees to describe the macrorole assignment in the four problematic cases.

1 Introduction

The actor-undergoer hierarchy is the masterpiece of Role and Reference Grammar. In 1984, Foley and Van Valin introduced two levels of semantic roles: an open list of specific thematic relations and two generalized semantic roles, the so-called macroroles: actor and undergoer. Specific thematic relations were arranged in a graded continuum of activity based on *Aktionsart*. Standard RRG (Foley & Van Valin 1984: 59; Van Valin & LaPolla 1997: 127, 146; Van Valin 2005: 61, 126; Van Valin 2013: 79) distinguishes five degrees of activity. The two more active degrees imply an activity component (**do'** [x, ...]) in the semantic description (logical structure = LS) of the predicate. The prototypical actor is a controlling human being acting as an AGENT,[1] i.e. the first argument of *to murder*. In contrast, the first argument of *to kill* need not have the value +hum. Even a natural force (e.g. *hurricane*) or an abstract concept or

[1] This is formally represented by DO' instead of **do'**.

state of affairs (e.g. *disease*, *drinking too much*) could be the first argument of the predicate. Hence, the first argument of *kill* is not a prototypical actor (an AGENT), but an EFFECTOR.

The other three grades are related to a stative component in the LS (**pred'**(x) and **pred'**(x, y)). RRG distinguishes between one-place and two-place states. The argument of a one-place state is a prototypical undergoer, e.g. an entity undergoing a change of state, being destroyed or killed, or coming into existence. Incidentally, not all one-place states in English, German, or Romance languages are represented in the lexicon by monovalent verbs. On the contrary, one-place states appear as part of the semantic representation of causative verbs of destruction. These verbs show a high activity contrast between the arguments and a complex semantic representation with both a cause and an effect component. The effect component consists of the resultative one-place state. As an effect of a breaking event, for example, an entity remains broken. Note that the lexical entries could be even verbs with a valency of three: *break the teacup against the window*. The entity that undergoes a change of state and remains broken is always selected for undergoer and linked to the direct object position in an active clause.

Two-place states (**pred'**(x, y)) formalize the semantics of bivalent stative verbs such as *own* or *like*. In line with a localist approach to thematic relations (cf. Gruber ([1965] 1976); Jackendoff (1972)), the first argument of these verbs could be considered a LOCATION. Therefore, the x argument (*Mary*) of *Mary owns an old car* is interpreted as a sphere of rights of possession and the y-argument (*car*) falls into this sphere. In *Peter likes pizza*, the x-argument (*Peter*) represents a space where the liking of the y-argument (*pizza*) is located. The second argument of the two-place state corresponds to a THEME in terms of Gruber ([1965] 1976). However, we also find two-place states as part of the semantic representation of three-place verbs of transfer (*give*), putting (*load*, *spray*), or removal (*drain*, *empty*). Depending on the language, the semantic subclass, and in the end the individual lexeme,[2] some of these verbs allow for a LOCATION argument in the direct object position and, therefore, for a marked undergoer choice, e.g. *Mary gave Peter a book*, *Harry sprayed the wall with paint*. Hence, standard RRG considers the first argument position of a two-place state a candidate for actor or undergoer. It will be the actor of transitive bivalent states, but the undergoer of the LOCATION as direct object construction of (three-place) verbs of transfer, putting, or removal.

[2] The so-called Dative alternation does not appear in German or in the standard varieties of Romance language (but see Abreu Gomes (2003) for colloquial Brazilian Portuguese). In Kailuweit (2005a), I showed that Pinker's semantic subclass approach to locative alternation (Pinker 1989) does not hold for German or Romance languages. Some of the English subclasses that, according to Pinker, allow for locative alternation do not have any instances in French, Spanish, or German, while others of his classes that do not permit alternation in English do so in German, French, or Spanish. However, we find verbs that do not alternate in every subclass allowing for locative alternation (see Kailuweit 2005a; 2008).

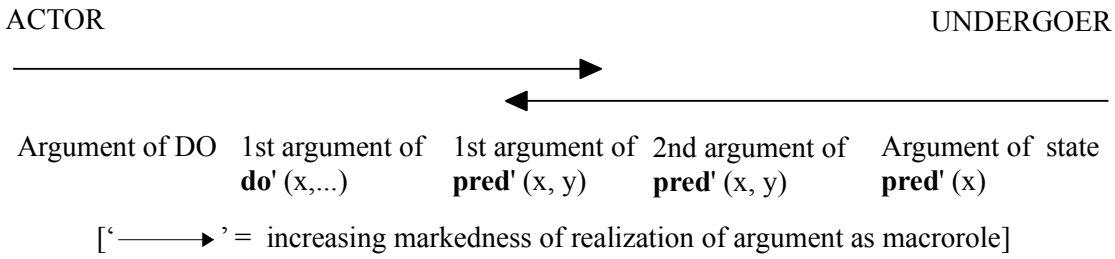


Table 1: Actor-undergoer hierarchy (cf. Foley & Van Valin 1984: 39; Van Valin & LaPolla 1997: 127, 146; Van Valin 2005: 61; Van Valin 2013: 79)

This brief overview has shown that macrorole assignment in RRG only takes five degrees of activity into consideration. In addition, there are no uniform criteria to grade the five degrees. The criteria are quite complex, ranging from lexical semantics (argument of DO versus argument **do'**(x, ...)), to *Aktionsart* (activity versus state), and to the number of arguments (one-place states versus two-place states).

2 Exploring the Limits of the Actor-Undergoer Hierarchy

Independently of the fact that the heterogeneous criteria for macrorole assignment seem to be problematic from a theoretical point of view, the actor-undergoer hierarchy hits its limits when it comes to describing the macrorole assignment of several verb classes. In this section, I will deal with four problematic cases: causation, three-place-predications, two-place object-EXPERIENCER verbs, and finally, verbs of motion.

2.1 Causative Predicates

The first problem concerns causativity. When Foley & Van Valin first worked out the actor-undergoer hierarchy in 1984, they followed the prevailing view at that time that all accomplishments are causative and that all causatives are accomplishments (Dowty 1979: 186; Foley & Van Valin 1984: 39). Later work in RRG (cf. Van Valin & LaPolla 1997: 97) has also shown that for states, activities, and achievements, too, there is a corresponding causative class while accomplishments such as *melt* also occur in a non-causative construction: *the hot water melted the ice* ⇔ *the ice melted*. Hence, causativity is formalized in RRG as "α CAUSE β where α β are logical structures of any type" (VanValin 2005: 45). Nonetheless, macrorole assignment for causative states and activities does not follow straightforwardly from the actor-undergoer hierarchy. In (1) and (2) the position of the highest ranking argument in the α-part of the con-

struction is the same as the position of the highest ranking argument in the β-part. In addition, in (2) the undergoer argument is an EFFECTOR, a position that is not accessible for undergoer selection in accordance to the actor-undergoer hierarchy.

(1) Bill's owning a gun frightens Mary. (causative state)
[**have'**(Bill, gun)] CAUSE [**feel'**(Martha, [**afraid'**])]
(Van Valin & LaPolla 1997: 97)

(2) The girl bounced the ball around the room. (causative activity)
[**do'**(girl, ...] CAUSE [**do'**(ball, [**bounce'**(ball)])]
(*ibid.*: 97)

A solution would be to flag the [**do'**(x, ...] position (thematic relation: EFFECTOR) as accessible for undergoer assignment. An additional rule could stipulate that the actor argument of causative constructions is always located in the α-part and the undergoer argument in the β-part of the complex construction. However, the question remains open as to whether in causative constructions the whole α-part is the actor or a single argument within the α-part. The first interpretation seems to be appropriate for (1), the second for (2).

2.2 Non-Causative Three-Place Predicates

As far as non-causative three-place predicates like *talk (to somebody about something)* or French *envier* – **envy'**(x, y, (z)) ('envy somebody for something') are concerned, the actor-undergoer hierarchy does not provide a formalism that accounts for the correct macrorole assignment. Standard RRG does not explicitly deal with these predicates. Van Valin & LaPolla (1997: 116–118) give a detailed description of English verbs of saying, but they do not indicate a LS for *talk*. In standard RRG, all three-place verbs are considered causative, but it is clearly the absence of causativity that distinguishes verbs of saying from verbs of telling (cf. Van Valin & LaPolla 1997: 118). Although the type frequency of non-causative three-place predicates is low, the token frequency of verbs of talking is high. However, if verbs of talking raise a problem for the actor-undergoer hierarchy, the problem naturally seems to be resolvable. Standard RRG allows for several classes of two-place activities, i.e. **do'**(x, [**sing'**(x, (y))], **do'**(x, [**see'**(x, (y))], **do'**(x, [**tap'**(x, (y))] etc. (cf. Van Valin & LaPolla 1997: 115).[3] The second argument represents an instance of **pred'**(x, y) (cf. *ibid.*: 127), although the predicate embedded in activities is not a state. Hence, RRG could allow for three-place activities with

[3] Two other classes of two-place activities are verbs of consumption and creation with an unspecified object: *eat pizza*, *write letters*. Notice, that the objects of these predicates are not affected undergoers as the specified objects of the active accomplishments *eat a pizza*, *write a letter*. They are instances of **pred'**(..., y), because they are rather the theme of eating and writing than the actually consumed or created object.

an argument in each of the three medium positions of the actor-undergoer hierarchy: **do'**(x), **pred'**(x, ...), and **pred'**(..., y). In Staudinger, Hartung, & Kailuweit (2008), we suggested the following LS for *parler*, the French counterpart of English *talk*:

(3) parler à qn de qc = **do'**(x, [**talk'**(x, y, z)])

Notice that the predicate **talk'**(x, (y), (z)) is not a state just as little as the **sing'**(x, (y)) component is a state in **do'**(x, [**sing'**(x, (y))]) in standard RRG. Activities do not consist of states embedded under a **do'** operator. Hence, we could associate the two arguments y and z with the positions **pred'**(x, ...) and **pred'**(..., y) of the actor-undergoer hierarchy without pretending that they are arguments of a state predicate. Standard RRG associates the y-argument of **sing'** with the **pred'**(..., y) position as well. However, this is not a solution for three-place non-causative states, such as Romance verbs of envy. Take, for example, French *envier*:

(4) ...cette belle figure creuse [...] il la lui enviait
...this pretty face haggard [...] he ACC.3SG;F DAT.3SG envie-PST.3SG
'It was this pretty haggard face that he envied him for.'
(FRANTEXT: ETCHERELLI)

The possible LS for French *envier* – **envy'**(x, y, (z)) – does not correspond to the three stative argument positions in the actor-undergoer hierarchy. The three positions **pred'**(x, ...), **pred'**(..., y) and **pred'**(x) distinguish two-place states (**pred'**(x, y)) from one-place states (**pred'**(x)). The position **pred'**(x), i.e. the prototypical undergoer position, cannot be added to the positions **pred'**(x, y), since the structures **pred'**(x) and **pred'**(x, y) exclude each other logically. In addition, the semantics of the OBJECT OF ENVY is different from those of the prototypical undergoer. The OBJECT OF ENVY is not affected by the event as the argument of **dead'** or **broken'** is; it does not undergo a change of state.

2.3 Non-Causative Object-EXPERIENCER Verbs

Nonetheless, it is not only three-place static verbs of emotion which raise problems as far as macrorole assignment is concerned. As I have shown in Kailuweit (2005b, 2013), there is no adequate representation for non-causative two-place episodic verbs of emotion. Standard RRG distinguishes between two types of episodic verbs of emotion, transitive verbs of the *upset*-type, and copular constructions such as **be.angry'**.[4]

[4] In line with Nissenbaum (1985) and Pesetsky (1995), Van Valin & LaPolla (1997) distinguish between non-episodic verbs of emotion, i.e. verbs of liking which denote a general subjective judgment, and episodic verbs of emotion, i.e. verbs of anger or fear which refer to a change of an emotional state of the experiencer in a concrete situation (cf. Koch 2001; Kailuweit 2005).

(5) episodic verbs of emotion (Van Valin & LaPolla 1997: 156, 402)

a. The photo in the newspaper upsets James.
[**be-in'**(newspaper, photo)] CAUSE [**feel'**(James, [**upset-about'**(**be-in'**(newspaper, photo))])]

b. Pat is angry at Kelly.
[**feel'**(Pat, [**angry.at'**(Kelly)])]

As Grimshaw (1990) and Pesetsky (1995) have argued, standard RRG considers the transitive variant causative. According to the LS given in (5), the non-transitive variant that is not only represented by copula constructions, but also by full verbs such as *to worry about*, could be described as a lexicalized anticausative variant of the transitive causative construction. Episodic emotions select the thematic relations EXPERIENCER and SENSATION. Standard RRG represents the SENSATION-argument as a predicate with an internal argument, e.g. **angry.at'**(y). This internal argument is not accessible for macrorole assignment. Therefore, verbs such as *to worry about* should be considered as macrorole intransitive states. Their only argument, the EXPERIENCER, assumes the undergoer macrorole.

There are several problems with this analysis (cf. Kailuweit 2005b, 2013). In the present paper, I shall condense the discussion to verbs of interest. The English verb *interest*, as well as its counterparts in German or in Romance languages, seems to follow the pattern in (4):

(6) verbs of interest

a. The photo in the newspaper interested James.
[**be-in'**(newspaper, photo)] CAUSE [**feel'**(James, [**interested-in'**(**be-in'**(newspaper, photo))])]

b. James is interested in the photo.
[**feel'**(James, [**interested-in'**(**be-in'**(newspaper, photo))])]

However, the causative analysis of *interest* is problematic. As a two-place verb, *interest* does not meet the tough-construction test.

(7) James is easy to upset.

(8) * James is easy to interest.

In Kailuweit (2005: 188), I have shown that this test – among others – is a good indicator for an active-causative interpretation of transitive object-EXPERIENCER verbs (OE-verbs).

While the majority of these verbs allows for the so-called tough-construction and licenses the imperative, the passive construction, and adverbials such as “deliberately”, a smaller group

of these verbs do not.[5] Verbs of interest are among those that do not permit a controlling AGENT as the subject argument in English, German, and Romance languages. In addition, all these languages allow for three-place constructions of verbs of interest. These contructions, acceptable to different degrees in the mentioned languages, seem to be causativations. Note that there is a parallel analysis of transitive *march* being the causativation of intransitive *march* in standard RRG (Van Valin & LaPolla 1997: 100).

(9) a. The troops marched to the barracks.

b. The sergeant marched the troops to the barracks.

(10) a. French: Il m' intéressait à ce qui l'
he ACC.1SG interest-PST.IMPF.3SG in that what ACC.3SG
intéressait
interest-PST.IMPF.3SG
'He got me interested in what he was interested in'
(FRANTEXT: BEAUVOIR, S. de)

b. ? Uncle Henry interested James in history.

As three-place constructions, verbs of interest meet the tough-construction test:

(11) a. * James is easy to interest.

b. James is easy to interest in anything wholesome.

In conclusion, two-place verbs of interest as well as a considerable subgroup of transitive OE-verbs are not causative. If this were the case, then a natural LS for two-place and three-place verbs of interest would be:

(12) a. **interest'**(x, y)

b. [**do'**(x, ∅)] CAUSE [**interest'**(y, z)]

However, these LS predict the wrong macrorole assignment. In the three-place construction, the y argument (EXPERIENCER) would be a marked choice for undergoer. In a parallel way, the x argument (EXPERIENCER) of the two-place construction would be selected as undergoer. This instance of marked undergoer choice would entail marked actor choice: the y-argument of **interest'**(x, y) has to be selected as actor to assume the subject position. Hence, a LS **interest'**(x, y) turns the principles of macrorole assignment upside down.

[5] Prototypical non-agentive-causative object experiencer verbs in French are *affecter* ('move'), *attirer* ('attract'), *emporter* ('carry away'), *frapper* ('strike'), *intéresser* ('interest'), *obséder* ('obsess'), *préoccuper* ('preoccupy'), *remuer* ('move') and *révolter* ('disgust'), in Italian *allietare* ('delight'), *attirare* ('attract'), *attrarre* ('appeal'), *colpire* ('strike'), *consumare* ('exhaust'), *divorare* ('devour'), *interessare* ('interest'), *preoccupare* ('preoccupy'), *rivoltare* ('disgust'), *sconvolgere* ('upset'), *trascinare* ('enthuse') (Kailuweit 2005b).

2.4 Motion Verbs as Active Accomplishments

Motion verbs, such as Italian *correre* ('run') or *slittare* ('slide'), constitute a problem for linking theories. In non-directional usage, they behave like activities and select HAVE in compound tenses, while they seem to be accomplishments selecting BE in compound tenses in combination with a goal-adverbial.

(13) La macchina è slittata nel fosso.
'the car slid into the ditch'

(14) La macchina ha slittato in modo pericoloso.
'the car slid dangerously'
(Schwarze 1996: 12)

RRG provides an elegant analysis of this phenomenon. Verbs of directed motion are basically activities that could be converted into active accomplishments via a lexical rule (Van Valin & LaPolla 1997: 111–13; Van Valin 2005: 47; Van Valin 2013: 85).

(15) activity [motion] ⇒ active accomplishment:
do'(x, [**pred'**(x)]) ⇒ **do'**(x, [**pred'**(x)]) & INGR **be-LOC'**(y, x)

The linking of the active accomplishment construction seems to follow straightforwardly from the general rules. As macrorole-intransitive constructions, they select one macrorole. Due to the activity component, this MR is an actor (Van Valin 2013: 83). However, in Italian the selection of an actor seems to be in contradiction with the auxiliary chosen in compound tenses not only in the activity construction, but also in the active accomplishment construction. The use of auxiliaries in Italian is due to the fact that intransitive activities select HAVE while intransitive states, achievements, and accomplishments select BE (Van Valin 1990). Notice that the macrorole of intransitive non-activities is an undergoer and not an actor. This is in line with approaches that consider "agentlessness" the prevailing factor of inaccusativity (Schwarze 1996). Hence, *correre* and *slittare* as active-accomplishments behave like activities as far as macrorole assignment is concerned, but like accomplishments regarding auxiliary selection.

There are additional problems with motion verbs. They could be intransitive as Italian *correre* ('run') or *slittare* ('slide') and their English counterparts or transitive as English *enter* or *reach*. Some verbs, e.g. English *climb*, even allow for both an intransitive and a transitive construction.

(16) a. Chris climbed Mont Blanc.
b. Chris climbed up Mont Blanc.

Van Valin & LaPolla (1997: 153) start from Talmy's (1985, 1991) distinction of satellite-framed and verb-framed languages. While English as a Germanic language preferably codes the manner of motion lexically and the GOAL on a satellite (*run into the room*), it is the other way around for Romance languages (i. e. French *entrer au salon en courant* ['enter the room running']). From a typological point of view, English is considered a satellite-framed language; Romance languages are verb-framed. However, English possesses verb-framed predicates such as *enter* while Romance languages express directed motion exceptionally with a satellite-framed construction as we have seen for Italian *slittare* ('slide').[6]

As far as the LS for active accomplishments is concerned, it is not clear how standard RRG would apply this structure to transitive constructions:

(17) a. Chris ran to the park.
do'(Chris, [**run'**(Chris)]) & INGR **be-at'**(park, Chris)

b. Chris entered the room.
do'(Chris, [**enter'**(Chris)]) & INGR **be-in'**(room, Chris)
or
do'(Chris, [**enter'**(Chris, room)]) & INGR **be-at'**(room, Chris)

Nonetheless, starting from any of two LS proposed in (17b), Chris would not only be the first candidate for actor, but also for undergoer. Hence, an additional MR-assignment rule would be necessary: "Assign actor first, then assign undergoer to the remaining direct core argument." The choice of macroroles would not only depend on the internal syntax of the LS, but also on temporal order inside the processing algorithm.

Van Valin (this volume) introduces the following modification to the LS of active accomplishments of motion to account for the incrementality of these predicates: [**do'**(x, [**run'**(x)]) ∧ PROC **cover.path.distance'**(x, (y))] & INGR **be-at'**(z, x). This LS is applicable to transitive constructions of the type *Chris ran two miles*, where the y-argument *two miles* constitutes the Undergoer.

[6] Van Valin & LaPolla (1997: 153) claim that verbs of motion in verb-framed languages tend to be macrorole transitive. This cannot be confirmed for Romance languages and seems to be incorrect from a typological point of view (Choi 2009). Intransitive French verbs of directed motion are among the most frequent verbs of the language. According to the frequency dictionary of Julliand, Brodin, and Davidovitch (1970) *venir* ('come') is the fourth most frequent verb, *aller* ('go') is the tenth, *rester* the 14th, *passer* the 16th, and *arriver* the 21st most frequent. *Revenir* ('come again') is the 25th, *partir* ('leave') the 35th, *entrer* ('enter') the 36th, and *sortir* ('go out') the 40th most frequent. Van Valin & LaPolla (*ibid.*) explain the transitivity of English *enter* by the fact that the verb is of Romance origin. However, French *entrer* as well as Spanish *entrar* or Italian *entrare* are intransitive verbs. English *climb*, allowing a transitive construction, is of Germanic origin.

(18) Chris ran two miles.
[**do'**(Chris, [**run'**(Chris)]) ∧ PROC **cover.path.distance'**(Chris, two miles] & INGR **be-at'**(path.endpoint, Chris)

However, in constructions with a GOAL, e.g. *to the park*, it remains unclear why the leftmost argument not chosen for actor, the z-argument *park*, is ruled out for undergoer selection. An explanation would be that according to Van Valin (2005: 65) actor and undergoer are never oblique arguments within the core.

(19) Chris ran two miles to the park.
[**do'**(Chris, [**run'**(Chris)]) ∧ PROC **cover.path.distance'**(Chris, two miles] & INGR **be-at'**(park, Chris)

On a side note, PROC **cover.path.distance'**(Chris, two miles) appears to be a rather unusual formalization. Note that the transitive meta-predicate **cover.path.distance'** is not transitive in English **Chris covered the path distance two miles*. A grammatically correct, but still semantically odd version would be *Chris covered the path distance of two miles*. Of course, one could argue that the meta-predicates – in the present example the incremental component – need not correspond to a direct paraphrase in the language. However, this is the principle of evidence in the semantic descriptions (LS) in standard RRG. The ordering of the arguments, especially for two-place states, follows from the internal syntax of the (embedded) state-predicate. Therefore, the correct macrorole assignment and linking of local relations with the LS **be-LOC'**(x, y) could be retrieved from the syntax of English be + preposition, for example (Van Valin 2005: 46–47; 58–60).

(20) Kim is in the library.
be-in'(library, Kim)

Furthermore, if we start from the syntactically more correct formalization PROC **cover.path.distance.of'**(two miles, Chris), *Chris* would be the best candidate for Undergoer. Of course, this argument already bears the actor macrorole as being the leftmost argument in the activity part of an active accomplishment.

Transitive verbs of motion, such as *reach*, raise another puzzling problem that the active accomplishment approach does not account for in a satisfying way.

(21) The climbers reached the summit.
INGR **be-at'**(summit, climbers)

In line with Dowty (1979: 68), Van Valin (this volume) lists *reach* as a prototypical achievement.[7] This seems to be true from a logical point of view. Nonetheless, starting from a LS INGR **be-at'**(summit, climbers), *the climbers* should be the undergoer of this transitive predicate and *the summit* the actor.

In personal communication, Van Valin reconsidered *reach* an active achievement adducing the following evidence. On the one hand, *reach* seems to be controllable, allowing for adverbs of the type *deliberately*. On the other hand, *reach* combines with *in x time* adverbials:

(22) a. Chris deliberately reached the room just after Mary started speaking.
b. Chris reached the meeting room in ten minutes.

The difference between *reach* and *run* lies in the fact that the first is not incremental (Van Valin, personal communication). In my opinion, it is problematic to make an argument out of rather marginal uses of *reach* with a controlling agent. In fact, the argumentation would not cover cases with an inanimate moving object:

(23) The bullet reached the target.

In addition, there is no linguistic evidence for the fact that *reach* is not incremental. If we focus on the non-specific activity part, there seem to be no specific differences between *run to the park* and *reach the park*. They both imply a gradual approach. In fact, *reach* seems to differ from an achievement such as *find*. *Find* refers to a facultative previous activity of searching. However, while the previous activity of searching is not necessarily an incremental approach to the GOAL of finding, we cannot reach a place without an approaching movement. In conclusion, the *Aktionsart* of *reach* as well as macrorole assignment remains an open question.

3 An Alternative to the Actor-Undergoer Hierarchy

Standard RRG retrieves macrorole assignment and therefore linking from *Aktionsart* information at the lexical level. However, as we have seen in the last sections, additional information is necessary to predict the correct assignment of macroroles. Logical structures show an internal syntax. They are ordered lists of arguments where the position, e.g. first or second argument of **pred'**(x, y), is decisive for macrorole assignment. At the meta-linguistic level,

[7] Van Valin (2013: 83) considered *reach* a transitive counterpart of an intransitive active accomplishment, such as *run*, with a LS "basically the same for both verbs", although *reach* has "an unspecified verb of motion in the activity part of the LS". However, in personal communication, he rejected this analysis, which contradicts the common assumption that *reach* is an achievement.

the position is determined by the syntax of the English predicates used in the formal descriptions. We have seen that at least for some causative constructions the basic rules of macrorole assignment are overridden by event structure information. Since causativity is not represented in the actor-undergoer hierarchy, it has to be stipulated that the actor is always assigned to the α-part of a causative construction (α CAUSE β), while the undergoer is located in the β-part. Hence, even an EFFECTOR, i.e. the first argument of **do'**(x, ...), is accessible for undergoer selection if it is located in the β-part of α CAUSE β.

At the level of the lexical entries, the syntax of the object languages comes into play. If a two- or three-place predicate is macrorole intransitive in a given language this has to be coded in the lexicon. Therefore, English macrorole assignment and linking for English *like* follow straightforwardly from the LS, while Italian *piacere* has to be marked as macrorole intransitive:

(24) lexical entry for Italian *piacere* ('like')
MR1 **like'**(x, y)

The lexical entries of Standard RRG do not respect the ideal separation of a lexical-semantical level and a syntactical level of representation mapped onto each other in both directions by linking algorithms. We already find a lot of syntactic information in the lexicon. In addition, Van Valin (2006: 285) proposes logical structures enhanced with macrorole and case assignment information "in order to expedite interpretation". Therefore, what was supposed to be the output of a linking process could at best be described as a result of lexical precompiling driven by lexical rules.

In Kailuweit (2013), I went one step further. My sketch of a Radical Role and Reference Grammar (RRRG) is inspired by unification grammar approaches and provides lexical structures with semantic, syntactic, and pragmatic information organized in feature-value pairings. At the heart of this approach lies the activity hierarchy, i.e. a formalism that represents different degrees of activity for each argument of a predicate. In previous research on verbs of emotion (Kailuweit 2005b, 2007a, 2007b, 2013, 2015), I ordered decomposed thematic relations according to their activity degree. The main sources of inspiration were the feature-based lists of thematic relations found in Rozwadowska (1988) and in Reinhart (2001, 2002). They represent a more formalized alternative to Dowty's (1991) proto-role approach.

Following Rozwadowska (1988), I consider three features: causative and/or control [c], mental (sentient) [m], and resultative (change of state) [r]. In line with Reinhart (2002),[8] my approach allows the features to assume three values +, –, and ±. While Rozwadowska (1988) allows for $2^3 = 8$ combinations and Reinhart (2001, 2002) for $3^2 = 9$, in my approach $3^3 =$

[8] Reinhart (2001, 2002) operates with only two features: C and M.

27 combinations are possible. This may seem like a complicated approach, but it is by far a more manageable system in comparison to the proto-role approach. Dowty (1990) lists five proto-agent and five proto-patient properties that are theoretically combinable, yielding 2^{10} = 1024 possible combinations.

In addition, the number of activity degrees decreases due to the fact that the features are weighted (see Table 2). The feature [c] is a strong actor feature, [m] is a weak actor feature, and [r] is a strong undergoer feature. The presence [+] of a strong feature will duplicate the value of the presence of a weak feature. If an argument is underspecified for one feature [±], the value will be half of the [+] value.

+c	±c	–c	+m	±m	–m	–r	±r	+r
4	2	0	2	1	0	0	-2	-4

Table 2: Features and Values for Activity Clusters

Weighting the features reduces the 27 possible combinations to eleven different degrees of activity. The combination [+c +m –r] represents the prototypical actor with the value 4+2+0 = 6, while the prototypical undergoer corresponds to a combination with the value [–c –m +r] 0+0-4 = -4. Nine intermediate summations are mathematically possible. Once the degree of activity of each argument of a predicate is determined, the assignment of macroroles and the linking of the argument to syntactic functions follow straightforwardly. As in Standard RRG, a transitive construction selects the most active argument for the actor macrorole and the most passive one for the undergoer macrorole. An active construction assigns the PSA-function (subject) to the actor and the direct object function (marked by accusative case in an accusative language) to the undergoer. If the construction is intransitive, it takes only one macrorole. Intransitive predicates denoting a stative subevent select an undergoer, which is realized as the PSA. In the rest of Section 3, I will come back to the four problematic cases I dealt with in Section 2.

3.1 Causative Predicates Revisited

As far as causative states and activities are concerned, there should be a clear contrast of activity between the two arguments, the CAUSER and CAUSEE. The most active argument of causative emotional states[9] is the argument called the CORRELATE in Kailuweit (2005b) following Ruwet (1993), i.e. the state of affairs at which the emotion is directed, and that is

[9] The state of affairs that episodic experiencer-object verbs denote is more complex. It consists not only of a state, but of a punctual change of the emotional state of the experiencer in a given situation caused by the

perceived as the CAUSER of the emotion. *Owning a gun* in (1), repeated for convenience in (25), is such an argument:

(25) Bill's owning a gun frightens Mary. (causative state)

The corresponding feature-value-cluster is [+c –m –r]. The EXPERIENCER, *Mary* in example (25) undergoes a change of state in a concrete emotional episode. As I have shown in earlier research (Kailuweit 2005b, 2013), the feature-value-cluster of this type of EXPERIENCER corresponds to [–c +m +r]. The macrorole assignment follows straightforwardly:

(26) frighten (Mary, Bill's owning a gun)
(Mary [–c +m +r]EXP = -2 ⇒ undergoer, Bill's owning a gun [+c –m –r]COR+CAU = 4 ⇒ actor)

Note that in the RRRG-approach *frighten*(x, y) is not a logical structure. The list of arguments in *frighten*(x, y) is not a syntactically ordered list. Semantics constitutes the hierarchy of the arguments, not syntax.

Causative activities such as (2), repeated for convenience as (27), are slightly different. Their CAUSER-argument is a prototypical actor bearing the feature-value-cluster [+c +m –r]. The CAUSEE-argument seems to be what Rozwadowska (1988) calls an affected AGENT bearing actor- and undergoer-features. Nonetheless, *the ball* can be classified as a non-animate object. Hence, the value for the [m]-feature is [–]. The only applicable actor-feature is [c]. From a logical point of view, inanimate moving objects have no control over their movements, but their physical properties are, to a certain extent, responsible for the way in which they move. We can bounce a ball around the room, but not a stone or a table. This fact licenses the attribution of a [±] value to the [c]-feature. The *room* as a location-argument is represented by the cluster [–c –m –r]. Notice that in the non-causative construction the *ball* is not affected. Due to a feature-value cluster of [±c –m –r], the activity degree is 2. Hence, *ball* is more active than *room* and will be selected as actor.

(27) The girl bounced the ball around the room. (causative activity)
bounce (girl, ball, room)
(girl [+c+m-r] = 6 ⇒ actor, ball [±c –m +r] = -2 ⇒ undergoer, room [–c –m –r]= 0)

(28) The ball bounced around the room.
(ball [±c –m –r] = 2 ⇒ actor, room [–c –m –r]= 0)

non-experiencer argument, the CORRELATE. In addition, they also denote the emotional state in its temporal extension as the parallel mental state of awareness of the CORRELATE. Therefore, they do not properly fit into any of the Vendler-Dowty classes of *Aktionsart* (Kailuweit 2015).

3.2 Non-Causative Three-Place Predicates Revisited

Non-causative three-place predicates show the following activity contrasts. Verbs of talking attribute the thematic relations SPEAKER, ADDRESSEE, and TOPIC OF CONVERSATION to their arguments. The SPEAKER controls his or her activity without being a CAUSER, so the value for [c] will be ± and the cluster is [±c +m –r]. The ADDRESSEE is supposed to hear the talking [+m] but not to be specifically affected [–r]. Hearing somebody talk is not a controlled activity [-c]. Notice that talking does not imply controlled listening: we are not always listening to someone who is talking to us. Hence, the cluster for the ADDRESSEE is [–c +m –r]. Finally, the TOPIC OF CONVERSATION as a state of affairs corresponds to the non-causative CORRELATE of verbs of emotion bearing a neutral cluster [–c –m –r]. Thus, the SPEAKER is the most active argument, followed by the ADDRESSEE. Prototypically, a dative is assigned to an argument showing a middle degree of activity. French *parler* corroborates this rule:

(29) J' ai parlé à ma mère de notre nouvelle vie
I have-PRS.1SG talk-PTCP to my mother about our new life
'I have talked to my mother about our new life.'
(FRANTEXT: MOTHERLAND, H.)

(30) parler (1sg, mother, new life)
1sg [±c+m-r] = 4 ⇒ actor, mother [–c +m –r] = 2 = Dat, new life [–c –m –r] = 0.

In Kailuweit (2005b), I describe Romance verbs of envy as a non-causative class of verbs of emotion. Their EXPERIENCER is rather active as the EXPERIENCER of *love* or *hate* is, expressing a subjective judgment.[10] This fact is coded by a value ± for [c] yielding the cluster [±c +m –r]. The other two arguments correspond to the CORRELATE and to a role, which I introduced with the label POINT OF REFERENCE. This thematic relation corresponds to the one Pesetsky (1995) named, rather obscurely in my opinion, a SUBJECT MATTER OF EMOTION. The POINT OF REFERENCE appears as the non-EXPERIENCER argument of *to fear for somebody*. Notice that the CORRELATE of your anxiety is not the person you fear for, but the fact that this person is in danger. Hence, the CORRELATE is unexpressed with *to fear for*, but combines with the POINT OF REFERENCE, i.e. the OBJECT OF ENVY, for three-place verbs of envy. The POINT OF REFERENCE is represented by the feature-value-cluster [–c ±m –r]. The CORRELATE is the most passive argument and assumes the undergoer macrorole. The activity clusters for French *envier* ('envy') are given under (31), where I repeat example (4) in a modified version:

[10] In standard RRG, this EXPERIENCER is called an EMOTER.

(31) ...Paul lui enviait cette belle figure creuse
Paul DAT.3SG envie-PST.IMPF 3SG this pretty face haggard
'Paul envied him for this pretty haggard face.'
(Paul [±c +m –r]EXP = 4 ⇒ actor, pretty face [–c –m –r]COR = 0 ⇒ undergoer, DAT.3SG [–c ±m –r]POR = 1)

Again, the argument of a middle degree of activity receives the dative.

3.3 Non-Causative Two-Place Object experiencer Predicates Revisited

As far as non-causative two-place object EXPERIENCER verbs are concerned, the activity degree of the EXPERIENCER is decisive. With episodic emotions, the emotional state begins at the moment the CORRELATE is perceived. The EXPERIENCER undergoes an uncontrollable change in his or her emotional state, although she or he does not consider the CORRELATE the CAUSER of the emotion. According to the proposal in Kailuweit (2005b, 2013), the difference between causative (e.g. *frighten*, see example (25) above) and non-causative transitive OE-verbs should be coded at the CORRELATE. I consider the activity degree of the experience the same in both constructions. The EXPERIENCER undergoing a change of state in a concrete emotional episode corresponds to a feature-value-cluster [–c +m +r]. Hence, the macrorole assignment follows straightforwardly for both classes of episodic accusative-EXPERIENCER predicates.

(32) causative OE-verbs, e.g. *frighten* (x, y)
([–c +m +r]EXP = –2 ⇒ undergoer, [+c –m –r]COR+CAU = 4 ⇒ actor)

(33) non-causative OE-verbs, e.g. *interest* (x, y)
([–c +m +r]EXP= -2 ⇒ undergoer, [–c –m –r]COR = 0 ⇒ actor)

For causative OE-verbs there is a considerable difference in activity between the CORRELATE perceived as the CAUSER of the emotion and the EXPERIENCER undergoing a change in his or her emotional state. For non-causative OE-verbs the CORRELATE is still the most active argument. Therefore, the CORRELATE receives the actor macrorole. However, there is only a slight contrast in activity between the two arguments. In Kailuweit (2015), I argue that this fact explains the well-known psych-properties of transitive OE-verbs (Belletti & Rizzi 1988, Landau 2009).[11] My claim is that psych-properties prototypically appear with non-causative transitive OE-verbs.

[11] Bouchard (1995) refuted Belletti & Rizzi's (1988) psych-properties one by one. However, he adduced causative OE-verbs as counter-evidence. What makes things puzzling is the fact that causative and non-causative OE-verbs do not form two clear-cut classes, but rather a continuum around prototypes such as *frighten* at the causative pole and *interest* at the non-causative pole (Kailuweit 2005, 2015 for a discussion of Spanish and Italian data).

3.4 Verbs of Motion Revisited

Finally, yet significantly, the activity hierarchy could help readjust the description of transitive verbs of motion. We have seen that standard RRG would fail to predict the undergoer selection of *run* with a DISTANCE-argument and a GOAL-argument. In addition, there is no description of variation shown for the locative argument of *climb*. Finally, macrorole assignment for transitive *reach* remains unclear, if we consider *reach* an achievement. I repeat the examples (16), (18), and (19) for convenience:

(34) a. Chris climbed Mont Blanc.
b. Chris climbed up Mont Blanc.

(35) Chris ran two miles to the park.

(36) The climbers reached the summit.

Starting from the frame-analysis of motion events by Talmy (1985), we could distinguish the following components (Choi 2009: 170). Since DISTANCE is a relevant participant in our examples, it should be added to the list:

MOTION	presence of motion
FIGURE	the moving object
GROUND	the reference object
PATH	the course followed by the Figure object with respect to the GROUND object.
DISTANCE	PATH length covered by the MOVER

Table 3: Components of Non-Causative Motion Events

One idea to deal with the difference between the two versions of *climb* could be substituting the *Gestalt* psychology conception of 'ground' – opposed to 'figure' – with a more specific participant role corresponding to the motion event:

(37) GROUND: The material basis in or on which the motion takes place.

The thematic relations realizing these four participant roles with specific predicates may correspond to three different degrees of activity. The most active argument is the 'figure', or in more specific terms, the MOVER with a feature-value-cluster [±c –m –r] as we have already seen for the MOVER-argument of the non-causative construction of *bounce*. If the MOVER is animate, controlling the movement, the cluster is [±c +m –r].

I propose a passive cluster for the DISTANCE argument [–c –m ±r]. The reason to attribute a ± value of the feature [r] to the DISTANCE-argument lies in its incrementality. In line with

Dowty (1991), I consider incrementality an undergoer property. If we like to put it like this, the DISTANCE is metaphorically "consumed" the more the motion proceeds.

In between the two lies the activity degree of the PATH-argument. The PATH as a LOCATION neither controls the movement, nor is it affected by it. Therefore, the PATH-argument shows a neutral degree of activity corresponding to the cluster [–c –m –r].

Should we consider the GROUND to be different from a LOCATION with a neutral cluster [–c –m –r]? Actually, we could ponder a more passive feature-value-cluster [–c –m ±r], in fact, the same we attributed to the DISTANCE-argument. The GROUND is at least to some extent affected by the motion: WALKERS, RUNNERS, RIDERS, or BIKERS leave traces on the GROUND they are moving on. Even SWIMMERS produce waves in the water they swim through. Yet still, the stairs we step on undergo usual wear and tear. We might code this less prototypical affectedness – if we compare it to objects that a specific activity consumes or destroys – by the ± value of the feature [r].

However, from a semantic point of view, this does not seem to be the right solution. Applied to our examples, the difference between *climb Mont Blanc* and *climb up Mont Blanc* does not consist in the considerable impairment of the mountain's landscape by the act of climbing. It is the completion of the action that comes into play, and to an extent, an inversion of 'figure' and 'ground'. The climbers are no longer the 'figure' that appears on the mountain as a 'ground', but the mountain seems to be the 'figure' situated in the trophy collections of the climbers.

As a consequence, I propose that the motion scenario includes a participant role that I will call the TARGET. The TARGET does not undergo a physical change of LOCATION or state, but when it is achieved it falls into the MOVERS domain. It is inscribed, so to speak, into the MOVERS logbook. Hence, an appropriate feature-value-cluster would be [–c –m ±r]. Note that the cluster is identical for DISTANCE and TARGET. In the end, both types of arguments belong to the same incremental process. While the DISTANCE describes the precise length of the PATH covered by the motion, the TARGET focuses on the endpoint of this incremental process. Dowty (1991) pointed out that it is not the MOVER who undergoes an incremental process, but the PATH. This seems to be correct, but it is not the PATH as an external LOCATION that is affected, but the subjective PATH appropriated by the MOVERS. While the DISTANCE from the mountain hut the climber starts from, to the summit of Mont Blanc remains objectively the same and could be covered by other climbers, the climber who has climbed Mont Blanc has incorporated this DISTANCE in an incremental way.

If we apply the set of activity clusters determined for verbs of motion to our examples, macrorole assignment and linking seem to follow straightforwardly:

(38) a. Chris climbed Mont Blanc.
Chris [±c +m –r] = 4 ⇒ actor, Mont Blanc [–c –m ±r] = -2 ⇒ undergoer

b. Chris climbed up Mont Blanc.
Chris [±c +m –r] = 4 ⇒ actor, Mont Blanc [–c –m –r] = 0

(39) Chris ran two miles to the park.
Chris [±c +m –r] = 4 ⇒ actor, two miles [–c –m ±r] = -2 ⇒ undergoer, park [–c –m –r] = 0

(40) The climbers reached the summit.
climbers [±c +m –r] = 4 ⇒ actor, summit [–c –m ±r] = -2 ⇒ undergoer

Note that there is no difference in the activity clusters between transitive *climb* and *reach*. *Reach* selects a TARGET argument as transitive *climb* does. In my opinion, there is no evidence against an incremental interpretation of the state of affairs denoted by *reach*. The verb implies an unspecific directed motion covering path distance.[12] Therefore, *reach* combines perfectly with *in x time* adverbials.

(41) a. The climbers reached the summit in five hours.

b. Peter reached the conference room in 10 minutes.

There is still an unsolved problem concerning the varying auxiliary selection of Italian verbs of motion such as *correre* ('run') or *slittare* ('slide'). Fundamentally, motion has to be distinguished from change of LOCATION. A theme-argument undergoing change of LOCATION corresponds to the cluster [–c –m +r]. At the lexical level, the prototypical verb of change of LOCATION that does not combine with adverbs such as *deliberately* is *arrive*.

The activity hierarchy approach would explain the varying auxiliary selection of Italian *slittare* ('slide') as a mismatch of the lexical and the constructional level. At the lexical level, *slittare* is a verb of motion focusing on the manner of motion. Only if the verb *slittare* is constructed with a GOAL-argument can it become an active accomplishment denoting a change of LOCATION. The constructional schema overrides the activity degree of the MOVER that becomes an argument undergoing a change of location ([±c +m –r] ⇒ [–c –m +r]).

However, such an approach has consequences for the interpretation of *Chris ran two miles to the park*. If we consider *run* in this construction a three-place predicate denoting a change of LOCATION, then macrorole assignment and linking would clash. *Chris* ([±c +m –r] ⇒ [–c –m +r]) = -4) would be the most passive argument and therefore the first choice for undergoer,

[12] This is not surprising from an etymological point of view. The etymology of *reach* is "stretch out, extend" (in various lit. and fig. uses). OE. *ræċan* (pt. *rǣhte*, *rāhte*) http://www.oxfordreference.com. This meaning component is not only present as a metaphor *in reach for*, but also licences the combination with *in x time* that would be ruled out for activities and achievements.

and *the park* ([–c –m –r] = 0) would be the most active argument and therefore the actor. Nonetheless, a closer look at the semantics reveals that *run two miles to the park* is not just a three-place version of *run to the park*. While *run to the park* focuses on the change of location, *the park* in *run two miles to the park* functions as the endpoint of the path incrementally covered by the MOVER. Hence, *the park* in *run two miles to the park* is not the third argument of *run*, but depends on the DISTANCE-argument as a specification of the path with the length of two miles.

(42) During sports day, the two miles from the schoolyard to the park were run by each pupil.

In conclusion, *run to the park* and *run two miles to the park* are both two-place constructions. The latter is a motion construction with a DISTANCE-argument, the former a change of location construction with a GOAL-argument.

4 Conclusion and Outlook

In this paper, I have shown that the activity hierarchy is a powerful tool that helps to solve difficult cases of macrorole assignment where RRG's actor-undergoer hierarchy hits its limits. On the one hand, the activity hierarchy allows for a finer grained semantic analysis. On the other hand, it is less complex, avoiding questions of syntactic order and even temporality in the process of macrorole assignment. Starting from a set of two actor features and one undergoer feature, the hierarchy brings into play semantic criteria that have dominated the discussion on generalized semantic roles in the last decades. The semantic analysis of different verb classes starts from a holistic approach to human communication in line with Van Valin's (1980: 229) claim in the early days of RRG. The activity degrees of different participants in a specific state of affairs are determined by interpretation, but not in an arbitrary way. They are not always based on morpho-syntactic tests. However, the tests that motivate the *Aktionsart* classification in RRG might not falsify the assumed feature-value clusters. Hence, the clusters are as objective as the logical structures in RRG, but allow for a sufficient number of activity degrees to describe the macrorole assignment in complex verb classes.

In the present paper, I dealt with 11 different feature-value clusters (out of 27) and 7 different activity degrees (out of 11). The following table summarizes the results.
Note that the participant roles listed in the third column have no theoretical states. In this respect, the activity hierarchy does not differ from the actor-undergoer hierarchy. In the fourth column, examples of English verbs are given restricted to the verb classes discussed in this paper.

[+c +m –r]	6	AGENT-CAUSER	bounce (**x**, y, z), interest (**x**, y, z)
[+c –m –r]	4	INANIMATE CAUSER	frighten (x, . . .)
[±c +m –r]	4	CONTROLLING MOVER/SPEAKER/EMOTER	talk (x, . . .), climb, love(x, . . .)
[c –m –r]	2	INANIMATE MOVER	bounce(**x**, y)
[–c +m –r]	2	ADRESSEE	talk(. . . y . . .)
[–c ±m –r]	1	POINT OF REFERENCE OF EMOTIONS	fear for (. . . y)
[–c –m –r]	0	LOCATION/PATH/CORRELATE OF EMOTIONS/TOPIC OF CONVERSATION	run (. . . y), interest (**x**, y), talk (. . . y)
[–c –m ±r]	-2	DISTANCE/TARGET	run (. . . , y), climb (. . . y) 2MR
[±c –m +r]	-2	AFFECTED MOVER	bounce (x, y, **z**)
[–c +m +r]	-2	AFFECTED EXPERIENCER	interest (x, **y**)
[–c –m +r]	-4	PATIENT (undergoing a change of state)/THEME (undergoing a change of location)	arrive (. . . y)

Table 4: Activity Hierarchy of Arguments of Selected Predicates.

Further research will have to show whether the activity hierarchy is able to describe a broad field of verb classes in a satisfying way. In RRRG as a more radical version of RRG (Kailuweit 2013), the activity hierarchy would replace the actor-undergoer hierarchy. However, RRRG does not aim to be a completely new theory. It is a project of theory building that – deeply inspired by the work of RRG's founder, Robert Van Valin – will eventually develop into an independent version of RRG, or just give an impulse to people working in the RRG framework to remodel central parts of the theory, e.g. the formalism of semantic representation and the mechanisms of macrorole assignment.

References

Abreu Gomes, Cristina. 2003. Dative alternation in Brazilian Portuguese: typology and constraints. *Language Design* 5: 67–78.

Belletti, Andrea & Rizzi, Luigi. 1988. Psych Verbs and Theta Theory. *Natural Language and Linguistic Theory* 6: 291–352.

Bouchard, Denis. 1995. *The Semantics of Syntax: A Minimalist Approach to Grammar.* Chicago/London: The University of Chicago Press.

Choi, Sonja. 2009. Typological differences in syntactic expressions of Path and Causation. In *Routes to Language: Studies in Honor of Melissa Bowerman*, V. M. Gathercole (ed.), 169–194. Mahwah, NJ: Lawrence Erlbaum Associates.

Dowty, David. 1979. *Word Meaning and Montague Grammar. The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.

—. 1991. Thematic proto-roles and argument selection. *Language* 67(3): 547–619.

Foley, William A. & Van Valin, Robert D, Jr. 1984. *Functional Syntax and Universal Grammar.* Cambridge, MA: CUP.

Grimshaw, Jane. 1990. *Argument Structure.* Cambridge, MA/London: MIT Press.

Gruber, Jeffrey S. (1965/1976): *Studies in lexical relations.* PhD Dissertation, MIT. Published as: *Lexical structures in syntax and semantics.* Amsterdam: North-Holland.

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

Juilland, Alphonse, Brodin, Dorothy & Davidovitch, Catherine. 1970. *Frequency Dictionary of French Words.* The Hague/Paris: Mouton.

Kailuweit, Rolf. 2005a. Lokativalternanz bei transitiven Verben – Englisch, Französisch, Spanisch und Deutsch im Vergleich. In *Beiträge zum romanisch-deutschen und innerromanischen Sprachvergleich, Vol. 2*, B. Wotjak, C. Schmitt (eds), 183–196. Bonn: Romanistischer Verlag.

—. 2005b. *Linking: Syntax und Semantik französischer und italienischer Gefühlsverben.* Tübingen: Niemeyer.

—. 2007a. *El enlace de los verbos de sentimiento – un cálculo de rasgos*. In *Actas del VI Congreso de Lingüística General, Santiago de Compostela, 3–7 de mayo de 2004, Vol. 2, Tomo 1, Las lenguas y su estructura (IIa),* P. Cano López (ed.): 1699–1708.

—. 2007b. *La classe P dans Les Verbes français et les verbes de sentiment*. *Langue Française* 153: 33–39, J. François, D. Le & D. Leeman (eds).

—. 2008. A RRG description of locative alternation verbs in English, French, German and Italian. In: *New Applications of Role and Reference Grammar*, R. Kailuweit, B. Wiemer, E. Staudinger, R. Matasović (eds), 328–355. Newcastle upon Tyne: Cambridge Scholars Publishing.

—. 2013. Radical Role and Reference Grammar (RRRG): A sketch for remodelling the Syntax-Semantics-Interface. *Linking Constructions into Functional Linguistics. The role of constructions in grammar*, B. Nolan & E. Diedrichsen (eds). 103–141. Amsterdam: John Benjamins.

—. 2015: Romance object-experiencer verbs: From aktionstart to activity hierarchy. In: *Verb Classes and Aspect* [IVITRA Research in Linguistics and Literature 9], E. Barrajón López, J. Cifuentes Honrubia & S. Rodríguez Rosique (eds), 312–333: Amsterdam: John Benjamins.

Koch, Peter. 2001. As you like it. Les métataxes actantielles entre expérient et phénomène. In: *La valence. Perpectives romanes et diachroniques*, L. Schøsler (ed.), 59–81. Stuttgart: Steiner.

Landau, Idan. 2009. *The Locative Syntax of experiencers.* Cambridge, MA: MIT Press.

Nissenbaum, Helen Fay. 1985: *Emotion and focus.* Stanford: CSLI

Pesetsky, David. 1995. *Zero Syntax. Experiencers and Cascades*, Cambridge, MA: MIT-Press.

Pinker, Steven. 1989. *Learnability and cognition: The acquisition of Argument Structure.* Cambridge, MA: MIT Press.

Primus, Beatrice. 1999. *Case and Thematic Roles. Ergative, Accusative, Active.* Tübingen: Niemeyer.

Reinhart, Tanja. 2001. *A synopsis of the Theta System.* http://www.let.uu.nl/~tanya.reinhart/personal/Papers/pdf/overview-final-with%20new%20intro.pdf

—. 2002. The Theta system – an overview. *Theoretical Linguistics* 28(3): 229–290.

Rozwadowska, Bożena. 1988. Thematic restrictions on derived nominals. In *Thematic Relations* [Syntax and Semantics 21], Wendy Wilkins (ed.), 147–165. New York: Academic Press.

Ruwet, Nicolas. 1993. Les verbes dits "psychologiques". Trois théories et quelques questions. *Recherches Linguistiques de Vincennes* 22: 95–124.

Schwarze, Christoph. 1996. The syntax of Romance auxiliaries. In *LFG-Workshop. Proceedings of the First LFG Conference. Rank Xerox Research Centre, Grenoble, August 26–28, 1996*. M. Butt & T. Holloway (eds). CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/

Staudinger, Eva, Hartung, Matthias & Kailuweit, Rolf. 2008. Linking syntax to semantics: Template selection and PP-attachment ambiguities, in: *New Applications of Role and Reference Grammar.* R. Kailuweit, B. Wiemer, E. Staudinger & R. Matasović (eds), 389–413. Newcastle upon Tyne: Cambridge Scholars Publishing.

Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. In *Language typology and syntactic description: Vol. 3. Grammatical categories and the lexicon*, T. Shopen (ed.), 58–76. Cambridge: CUP.

—. 1991. Path to realization: A typology of event conflation. In *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, L.A. Sutton & C. Johnson (eds), 480–519. Berkeley, CA: Berkeley Linguistics Society.

Van Valin, Robert D., Jr. 1980. Meaning and interpretation. *Journal of Pragmatics* 4: 213–231.

—. 1990. Semantic parameters of split intransitivity. *Language* 66(2), 221–260.

—. 2005. *Exploring the Syntax-Semantics Interface.* Cambridge: CUP.

—. 2006. Semantic macroroles and language processing. In *Semantic Role Universals and Argument Linking – Theoretical, Typological, and Psycholinguistic Perspectives*, I. Bornkessel, M. Schlesewsky, B. Comrie & A. D. Friederici (eds), 163–301. Berlin: Mouton de Gruyter.

—. 2013. Lexical representation, co-composition, and linking syntax and semantics. In *Advances in generative lexicon theory*, J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki & C. Lee (eds), 67–107. Dordrecht: Springer.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax. Structure, Meaning and Function.* Cambridge, MA: CUP.

Commercial Event Verbs in Spanish. A Corpus Based Study

Sergio Ibáñez
IIFL – Universidad Nacional Autónoma de México

Abstract

This work presents the analysis of Spanish verbs related to commercial events – *vender* 'to sell', *comprar* 'to buy', etc. –. We follow Fillmore's proposal (1977) that the different constructional patterns of verbs related to the same semantic domain have the same cognitive frame as the background. By means of a corpus based study we show that the four participants of the commercial event frame – the seller, the buyer, the goods and the money –, are semantically and syntactically productive. We claim that they are verbal arguments and should be part of a semantic level different from the cognitive frame and from the syntactic frames that only include the obligatory arguments. That level is equivalent to RRG logical structures.

1 Introduction

Fillmore's pioneer proposal on frame semantics in *The case for case reopened* (1977) established that the different constructional patterns of a verb, or of a set of predicates related to the same semantic domain, can be seen as relativized scenes that have an organized set of notions as their background; a *frame*, which in cognitive terms functions as the complete picture relative to which each scene is understood in the context of a communicative process. Fillmore exemplified this idea with the commercial event, which includes as participants a seller, a buyer, the goods which are acquired and the money which is exchanged for the goods. In a language like English, this frame can be activated by a set of different predicates, among which are *to sell, to buy, to cost, to pay,* and *to spend*. Each of these verbs brings in a different perspective on the event, and with this perspective a focus on one or some of its participants:

(1) a. I bought a dozen roses.
b. I paid five dollars.

c. I bought a dozen roses from Harry for five dollars.

d. I paid Harry five dollars for a dozen roses.

As can be seen in the clauses in (1), Fillmore's original examples, the verb *to buy* puts the buyer and the goods into perspective, as in (1a), while the verb *to pay* focuses on the seller and the money – (1b) –. Nevertheless, both predicates can appear with more complements: in (1c) the seller and the money get coded along with the goods and the buyer, and in (1d) the goods and the seller are also included, in addition to the other two participants.

In the same work, Fillmore pointed out that the case frame of predicates in the same domain does not need to comprise all the participants that are postulated as part of the relative frame, but just those which are put into perspective by each verb. This claim raises the question, relative to the examples in (1), of whether we should include, as part of the case frame, argument or logical structure of predicates such as *to buy* and *to pay*, just the participants that are usually considered obligatory, that is, the ones that get coded as subject and direct object, or whether we should also include the optional participants. In this last case, the argument structure of both verbs would equal the number of participants of the frame. This question is solved in various ways by different frameworks: On the one hand, syntactically oriented theories tend to include just the obligatory arguments; on the other hand, semantically driven frameworks allow for the inclusion of more participants, even if they are coded as oblique complements and are optional. Nevertheless, this topic is still a matter of debate.

RRG (Van Valin 2005) falls into the second category, as it takes semantics as the starting point in the analysis of clause structure. So, although RRG's logical structures (LS) are usually postulated to handle the syntactically most important arguments (there are just two macroroles) of a predicate, in principle, there is no restriction in the number of arguments that can be postulated as part of an LS. And in fact, in the last ten years, there have been, within the same theory, several proposals that posit semantically 'expanded' LSs for a diverse range of predicates, for example, for speech act verbs (Van Valin & LaPolla 1997; Ibáñez 2008), for cutting verbs (Mairal & Faber 2001) and for impact verbs (Mairal 2003, 2004), among others.

In this work, I want to continue on this same path and bring more positive arguments into the discussion for considering that the LS or argument structure of predicates should be semantically determined and should include all the participants that are relevant to the denoted event. The basic idea is that to explain the whole range of the constructional patterns of a predicate, as it is used by the speakers in real communicative processes, one needs to take into account all the participants that are related to its event structure and that can be coded in a 'significant' frequency in its different contextual uses. In order to show this, I present a corpus based analysis of the Spanish verbs related to the commercial event: *vender* 'to sell', *comprar*

'to buy', *pagar* 'to pay' and *cobrar* 'to charge'. The goal is to determine the constructional patterns of these predicates in terms of the number of participants that are coded, in terms of the semantic identity of these participants, and in terms of their coding properties. The data come from the Corpus de Referencia del Español Actual (CREA) and include 600 cases, 150 sentences for each of the 4 analysed predicates.

Although we positively show that the LS of predicates should have more information than the syntactic frames that only include the more prominent syntactic participants, the so-called obligatory arguments, we also show that the LS of a predicate is different from their relative frame: the number and semantic values of the arguments in the former can be different from the ones in the frame. In this sense, we argue that in order to completely explain the syntactic behaviour of predicates one needs recourse to three different instances of analysis: (1) The actual syntactic construction which conceptually activates a certain state of affairs; (2) the cognitive frame, which functions as the whole scenario against which the conceptual scene activated by the clause is interpreted and understood; and (3) the LS or argument structure, which provides the number of arguments and their semantic and syntactic identity.

2 The commercial event frame

Returning to Fillmore's proposal, the commercial event frame essentially has four participants: a seller, a buyer, the goods and the money. The frame can roughly be described as follows: someone, a buyer, gives money to another person, a seller, and this last one, in exchange, gives some goods to the former. Fillmore's point of view is that this frame functions as the cognitive background that permits the speakers to understand sentences related to it, but at the same time he seems to think that the frame, as a whole, plays a limited role in the identity of the particular syntactic frames of verbs such as *to sell* and *to buy*. These syntactic frames are supposed to imply focalization of a particular 'segment' of the frame. In Fillmore's words:

> A prototypic commercial event involves all of these things (the four participants already mentioned), but any single clause that we construct in talking about such an event requires us to choose one particular perspective on the event. In the usage that I suggest, any verb identifying any particular aspect of the commercial event will constrain us to bring one or more of the entities in the event into perspective, the manifestation of this choice for English being the selection of grammatical functions corresponding to the notions of underlying subject and direct object. (1977: 72)

Fillmore's view, then, seems to be that the syntactic frames operate only with the participants that are brought into perspective with each predicate, those that are nuclear in grammatical terms:

> Within the view that meanings are relativized to scenes, however, it may not be necessary to believe that everything that is included in our understanding of a sentence is necessarily a part of the underlying grammatical structure of that sentence; it seems preferable to say that a word like *buy* or *pay* activates the scene of the commercial event; that everybody who understands the word knows what the various components and aspects of such an event are; and that a speaker's linguistic knowledge of a verb includes the knowledge of the grammatical ways in which various parts of the event can be realized in the form of the utterance. (1977: 75)

So the question I want to address in this work is what the actual participants which have to be postulated as part of the argument structure or the logical structure (LS) of the commercial event predicates are: are they only those that are coded in nuclear grammatical functions? Are they all those that are part of the frame? In other words, the question posed investigates the relation between the cognitive frame and the language semantic-syntactic frames.

Our particular view is that to answer this question one needs to see, in real corpus data, what the actual participants are that are significantly coded, in frequency terms, with each verb. That is, besides acknowledging what the obligatory syntactic participants are, one needs to determine which ones are brought into perspective, that is, which are productively coded, in any formal way. So, the proposal is that arguments must be 'extracted', by way of a factorization process, from the whole range of constructional patterns of a verb.

3 Constructional horizons of Spanish commercial event verbs

First, it has to be said that, indeed, there is no iconic relation between the frame and the LS of verbs. Although the seller and the buyer are equally engaged as agents in the frame (and simultaneously as receivers and sources), in the verbal semantics of the predicates under study, they are treated as different notions: On the one hand, *vender* 'to sell' has the seller as an agent and the buyer as a recipient or goal argument, and *pagar* 'to pay' has the buyer as an agent and the seller as a recipient. On the other hand, *comprar* 'to buy' has the buyer as an agent and the seller as a source, and *cobrar* 'to charge' has the seller as an agent and the buyer as a source argument. Other predicates, like *intercambiar* 'to interchange' *permutar* 'to permute' or *canjear* 'to exchange', which also have two participants that are simultaneously 'givers' and 'receivers', don't show this difference; they imply both of these participants as agents and treat them alike grammatically: for example, both can be conjoined in a single complex NP and appear as the subject of the clause – *Jorge y Carmen intercambiaron regalos* / 'George and Carmen exchanged gifts' –, where both Jorge and Carmen are givers and receivers. This behavior is not allowed for the verbs we are analyzing here: **Jorge y Carmen compraron dulces por $10 dólares* 'George and Carmen bought candies for $10 dollars' doesn't imply that

one of them is the seller; both of them are buyers. This difference is important, since in the conceptual frame, the actions of 'selling' and 'charging', on the one side, are notions that are aligned together, and, on the other side, the actions of buying and paying are the ones that go together. Just the other way around, in semantic terms *vender* and *pagar* are the ones that go together, as they both imply an agent and a recipient or goal, while *comprar* and *cobrar* are similar in that they have an agent and a source. This accounts for the existence of two different sets of constructional patterns: one for *vender* 'to sell' and *pagar* 'to pay', and one for *comprar* 'to buy' and *cobrar* 'to charge'. In what follows I present all the constructions registered for each pair of verbs and their frequency, in what I call their constructional horizons:

		Vender 'to sell'	*Pagar* 'to pay'
4 argument constructions		**9 cases 6.0%**	**19 cases 12.7%**
1. S (A) + V + DO (T) + OA ($) + IO (R)		8 cases 5.3%	4 cases 2.7%
2. S (A) + V + DO ($) + OA (T) + IO (R)		1 case 0.7%	15 cases 10%
3 argument constructions		**55 cases 36.6%**	**64 cases 42.7%**
Direct constructions:		**51 cases 33.9%**	**60 cases 40%**
3. S (A) + V + DO (T) + IO (R)		21 cases 14%	5 cases 3.3%
4. S (A) + V + DO (T) + OA (G)		13 cases 8.6%	0 cases 0%
5. S (A) + V + DO (T) + OA ($)		11 cases 7.3%	9 cases 6%
6. S (A) + V + DO ($) + IO (R)		2 cases 1.3%	11 cases 7.3%
7. S (A) + V + DO ($) + OA (T)		3 cases 2.0%	28 cases 18.7%
8 .S(A) + V + OA ($) + IO (R)		1 case 0.7%	0 cases 0%
9. S(A) + V + OA (T) + OI (R)		0 cases 0%	4 cases 2.7%
Passive constructions:		**4 cases 2.7%**	**4 cases 2.7%**
10. S(T) + V + IO (R) + OA (A)	(PP)	2 cases 1.3%	1 case 0.7%
11. S(T) + (*se*)V + IO (R) + OA ($)	(RP)	2 cases 1.3%	0 cases 0%
12. S($) + (*se*)V + IO (R) + OA (T)	(RP)	0 cases 0%	3 cases 2%
2 argument constructions:		**63 cases 42.0%**	**53 cases 35.3%**
Direct constructions:		**40 cases 26.7%**	**37 cases 24.7%**
13. S (A) + V + DO (T)		34 cases 22.6%	21 cases 14%
14. S (A) + V + DO ($)		2 cases 1.3%	5 cases 3.3%
15. S (A) + V + IO (R)		1 case 0.7%	3 cases 2%
16. S (A) + V + OA (T)		0 cases 0%	5 cases 3.3%
17. S (A) + V + OA (G)		1 case 0.7%	0 cases 0%
18. S (A) + V + OA ($)		0 cases 0%	1 case 0.7%
19. S (T) + V + DO ($)		2 cases 1.3%	0 cases 0%
20. S ($) + V + DO (T)		0 cases 0.0%	2 cases 1.3%
Passive constructions:		**23 cases 15.3%**	**16 cases 10.7%**
21. S (T) + V + OA (A)	(PP)	2 cases 1.3%	1 case 0.7%
22. S (T) + V + IO (R)	(PP)	5 cases 3.3%	1 case 0.7%
23. S (T) + V + OA (G)	(PP)	2 cases 1.3%	0 cases 0%
24. S ($) + V + OA (A)	(PP)	0 cases 0%	1 case 0.7%
25. S (T) + V + OA ($)	(PP)	1 case 0.7%	1 case 0.7%

26. S (T) + (*se*)V + IO (R)	(RP)	6 cases 4%	0 cases 0%
27. S (T) + (*se*)V + OA (G)	(RP)	3 cases 2%	0 cases 0%
28. S (T) + (*se*)V + OA ($)	(RP)	3 cases 2%	5 cases 3.3%
29. S ($) + (*se*)V + OA (G)	(RP)	1 case 0.7%	0 cases 0%
30. S ($) + (*se*)V + OA (T)	(RP)	0 cases 0%	7 cases 4.7%
1 argument constructions:		**23 cases 15.3%**	**14 cases 9.3%**
Direct constructions:		**5 cases 3.3%**	**6 cases 4.0%**
31. S (A) + V		3 cases 2%	5 cases 3.3%
32. S (T) + V		2 cases 1.3%	0 cases 0.0%
33. S ($) + V		0 cases 0%	1 case 0.7%
Passive constructions:		**18 cases 12%**	**8 cases 5.3%**
34. S (T) + V	(PP)	2 cases 1.3%	2 cases 1.3%
35. S ($) + V	(PP)	0 cases 0%	1 case 0.7%
36. S (T) + (*se*)V	(RP)	13 cases 8.6%	2 cases 1.3%
37. S ($) + (*se*)V	(RP)	1 case 0.7%	2 cases 1.3%
38. (*se*)V + IO (R)	(IM)	2 cases 1.3%	1 case 0.7%

Table 1: Constructional schemes *vender, pagar*

		Comprar 'to buy'	*Cobrar* 'to charge'
4 argument constructions		**11 cases 7.3%**	**16 cases 10.7%**
1. S (A) + V + DO (T) + OA ($) + OI (S)		10 cases 6.6%	1 case 0.7%
2. S (A) + V + DO ($) + OA (T) + OI (S)		1 case 0.7%	15 cases 10.0%
3 argument constructions		**100 cases 66.7%**	**60 cases 40.0%**
Direct constructions:		**99 cases 66.0%**	**59 cases 39.3%**
3. S (A) + V + DO (T) + IO (S)		24 cases 16%	6 cases 4%
4. S (A) + V + DO (T) + OA (S)		16 cases 10.6%	1 case 0.7%
5. S (A) + V + DO (T) + OA ($)		36 cases 24%	3 cases 2%
6. S (A) + V + DO ($) + IO (S)		0 cases 0%	17 cases 11.3%
7. S (A) + V + DO ($) + OA (S)		0 cases 0%	5 cases 3.3%
8. S (A) + V + DO ($) + OA (T)		6 cases 4%	23 cases 15.3%
9. S (A) + V + OA ($) + IO (S)		1 case 0.7%	1 case 0.7%
10. S (A) + V + OA (T) + IO (S)		0 cases 0%	3 case 2%
11. S (A) + V + DO (T) + IO (B)		16 cases 10.6%	0 cases 0%
Passive constructions:		**1 case 0.7%**	**1 case 0.7%**
12. S (T) + V + OA (S) + OA (A)	(PP)	1 case 0.7%	0 case 0%
13. S ($) + V + OA (T) + OA (A)	(PP)	0 case 0%	1 case 0.7%
2 argument constructions:		**5 cases 23.3%**	**60 cases 40.0%**
Direct constructions:		**0 cases 13.3%**	**50 cases 33.3%**
14. S (A) + V + DO (T)		15 cases 10%	10 cases 6.7%
15. S (A) + V + DO ($)		0 cases 0%	25 cases 16.7%
16. S (A) + V + IO (S)		2 cases 1.3%	5 cases 3.3%
17. S (A) + V + OA (T)		0 cases 0%	8 cases 5.3%
18. S (A) + V + OA (S)		1 case 0.7%	1 case 0.7%
19. S (A) + V + OA ($)		1 case 0.7%	1 case 0.7%

20. S ($) + V + DO (T)		1 case 0.7%	0 cases 0%
Passive constructions:		**15 cases 10%**	**10 cases 6.7%**
21. S (T) + V + OA (A)	(PP)	2 cases 1.3%	0 cases 0%
22. S (T) + V + IO (S)	(PP)	4 cases 2.7%	0 cases 0%
23. S ($) + V + OA (A)	(PP)	0 cases 0%	1 case 0.7%
24. S (T) + (*se*)V + IO (S)	(RP)	3 cases 2%	2 cases 1.3%
25. S (T) + (*se*)V + OA (S)	(RP)	2 cases 1.3%	0 cases 0%
26. S (T) + (*se*)V + OA ($)	(RP)	4 cases 2.7%	0 cases 0%
27. S ($) + (*se*)V + OA (T)	(RP)	0 cases 0%	7 cases 4.7%
1 argument constructions:		**4 cases 2.7%**	**14 cases 9.3%**
Direct constructions:		**3 cases 2.0%**	**5 cases 3.3%**
28. S (A) + V		2 cases 1.3%	5 cases 3.3%
29. S ($) + V		1 case 0.7%	0 cases 0%
Passive constructions:		**1 case 0.7%**	**9 cases 6.0%**
30. S (T) + V	(PP)	1 case 0.7%	2 cases 1.3%
31. S ($) + V	(PP)	0 cases 0%	2 cases 1.3%
32. S (T) + (*se*)V	(RP)	0 cases 0%	2 cases 1.3%
33. S ($) + (*se*)V	(RP)	0 cases 0%	2 cases 1.3%
34. (*se*)V + IO (S)	(IM)	0 cases 0%	1 case 0.7%

Table 2: Constructional schemes *comprar, cobrar*

The constructional schemes in Tables 1 and 2 have a semantic-syntactic identity. On the one side, the initials in capital letters out of the brackets account for each of the grammatical functions that are present in the different type of clauses: S is subject, DO direct object, IO indirect object, OA is oblique argument, V is verb and (*se*)V is a verb marked as passive with the reflexive *se*. On the other side, the letters and symbols inside the brackets are the initials for the semantic role of the participant that is coded in each of the grammatical functions. Thus, (A) is agent, (T) is theme, ($) is the money participant, (R) is recipient, (S) is source, (G) is goal, and (B) is benefactive, a non-argument participant that appears coded in the core function of IO. The money and the goods have the same status as themes, as they function partially alike in semantic and syntactic terms. However, I am labelling the goods as theme and I am retaining the label 'the money' for this participant in order to make the differentiation between them easier. The conjoined letters (PP), (RP) and (IM), inside the brackets, stand, respectively, for a periphrastic passive construction, *se* reflexive passive construction and an impersonal construction. It is important to know that the ordering of the elements in the constructional schemes doesn't reflect the actual syntactic order inside the clauses.

As an example, the constructional pattern No. 1, in Table 1, can be read as follows: it is a 4 argument clause where the agent is the subject, the theme the direct object, the money is an oblique argument, a PP, and the recipient is an indirect object. An example for each type of construction in both tables is presented below:

Examples for each constructional type in Table 1

4 argument constructions

(1) En la primera mitad del 2000 México **vendió** mercancías al mundo por 79 mil millones de dólares,
'In the first half of 2000, México sold goods to the world for 79 thousand million dollars,'

(2) Los ambulantes aseguran deben **pagar** a su líder hasta mil 200 pesos mensuales por un local.
'Ambulant workers assure they have to pay their leader around $1200 pesos each month for a commercial venue.'

3 argument constructions
Direct Constructions

(3) . . . nuestro ya conocido Moe le **vendió** el Stardust a otro hombre de negocios.
'. . . our well-known Moe sold the Stardust to another businessman.'

(4) . . . desvió un embarque recién llegado de España, para **venderlo** en Jamaica.
'. . . (he) altered the course of a shipment, which had just arrived from Spain, in order to sell it in Jamaica.'

(5) . . . es un hombre que **vende** su cuerpo por dinero.
'. . . (he) is a man that sells his body for money.'

(6) Chapa **pagó** a Ramiro Aguilar, el autor del "anónimo", todavía prófugo, 2 millones 500,000 pesos.
'Chapa paid $2,500,000 pesos to Ramiro Aguilar, the author of the anonymous (letter), who is still on the run.'

(7) Tan sólo por él, el equipo europeo **pagó** millón y medio de dólares.
'In exchange for him, the European team paid one million and a half dollars.'

(8) La ANDI aceptó esta propuesta, y **pagó** a los intérpretes a 10,000 pesos cada uno.
'The ANDI accepted the proposal, and paid the performers in quantities of 10,000 pesos each'

(9) Buñuel. . . **habría pagado** también al pastor por una cabra despeñada a balazos.
'Buñuel. . . also would have paid the shepherd for a goat that fell from the cliff after a gunshot.'

Passive constructions

(10) . . . el camino que lleva a las lagunas **ha sido vendido** por autoridades municipales corruptas al grupo Santa Bárbara,
'. . . the road to the lagoons has been sold by the corrupt authorities to the Santa Bárbara Group,'

(11) . . . la producción excesiva **se le vende** a los estados unidos por un precio más bajo,
'. . . the excessive production is sold to the United States for a lower price,'

(12) . . . **se pagaron** grandes cantidades a los terratenientes por la compra de predios afectados tras el surgimiento del EZLN.
'. . . big amounts of money were paid to the landlords for purchasing the properties, which were affected by the rising of the EZLN.'

2 argument constructions
Direct constructions

(13) Hewlett Packard se ha convertido no solamente en una compañía que **vende** equipo, sino en una empresa que brinda servicios,
'HP has become, not only a Company that sells equipment, but a Company that offers services.'

(14) Este pueblo contó por algún tiempo con el dinero que **pagó** la Azufrera Panamericana,
'For some time this town had at its disposal the money that the Azufrera Panamericana paid,'

(15) Entra al rescate el gobierno; les va a volver a **pagar**.
'The government comes in to the rescue; (it) will pay them again.'

(16) Esto quiere decir que todos los mexicanos tendrán que **pagar** por los quebrantos de los 27 ingenios.
'This means that all Mexicans will have to pay for the downfall of the 27 sugar mills.'

(17) En la carrera para **vender** en el mercado interno, muchos negocios están descuidando factores. . .
'In the race for selling on the internal market, a lot of businesses are neglecting certain factors. . . '

(18) Al reservar en el restaurante típico, nos avisaron que deberíamos **pagar** en dólares.
'When we made the reservation at the typical restaurant, we were told we would have to pay in dollars.'

(19) Las computadoras de escritorio **vendieron** este año millones de pesos.
'Personal computers sold for millions of pesos this year.'

(20) Al final de cuentas, esos 150 000 dólares **pagaron** sus deudas.
'At the end, those 150 000 dollars paid his debts.'

Passive constructions

(21) Imagínense cuanto orgullo da el saber que sus estudios **fueron pagados** por nosotros.
'Imagine how proud we feel knowing that his studies were paid by us.'

(22) El volumen que **fue vendido** al mercado nacional fue 5.5% inferior al de 1999.
'The volume that was sold to the national market was 5.5% inferior to the one in 1999.'

(23) . . . y reetiquetan la ropa para **ser vendida** en el mercado negro a precios aún más baratos.
'. . . and (they) relabel the prices, in order to get the cloth sold on the black market at even lower prices.'

(24) En mi libro menciono que hasta mi sueldo **era pagado** por esa persona.
'In my book, I mention that even my salary was paid by that person.'

(25) Todo ello **fue pagado** con los menguantes ingresos de los mexicanos que deben trabajar para vivir.
'All this was paid with the low incomes of the Mexicans that have to work for a living.'

(26) ...han supervivido unos pocos recuerdos autóctonos que **se venden** a los turistas foráneos.
'...a few native souvenirs have survived, ones that are sold to the foreigner tourists.'

(27) ...esos alimentos orgánicos que desde hace tiempo **se venden** en el mercado de la salud.
'...those organic foods that for quite some time have been sold on the health market.'

(28) La actividad tuvo un ligero repunte, y **se pagó** el kilo a $14 promedio.
'The activity had a little increase, and a kilo was paid $14 on average.'

(29) ...miles de millones **se venden** en el mercado europeo.
'...thousands of millions are sold in the European market'

(30) Mucho dinero **se ha pagado** por los intereses en los últimos años.
'A lot of money has been paid for interest in the last years.'

1 argument constructions:
Direct constructions:

(31) ...con la finalidad de que los clientes conozcan el disco que quieren adquirir, antes de **pagar**.
'...in order to let the clients listen to the record they want to buy, before paying.'

(32) ...un producto que en realidad no seduce, no **vende**, no tiene buena posición en el mercado.
'...a product that really doesn't seduce, doesn't sell, that doesn't have a good position on the market.'

(33) ...porque se trata de ofertas que valen la pena. Usted no se preocupe, su sueldo **paga**.
'...because it's about bargains that are worth it. Don't worry, your salary pays.'

Passive constructions

(34) Al final no se produjo tan terrible crimen y la joven **fue vendida** como esclava.
'At the end, that horrible crime didn't take place and the young girl was sold as a slave.'

(35) ...los montos de los créditos solicitados al Banrural y al Fira **fueron pagados** totalmente en el caso de la primera institución.
'...The amounts of credits requested from Banrural and from Fira were completely paid in the case of the former.'

(36) Los libros de ovnis, por poner una muestra, **se venden** como pan caliente.
'Books about UFOs, just as an example, sell like hot bread.'

(37) ...un incremento de 55% sobre el año anterior donde **se vendieron** 67 millones de euros.
'...an increase of 55% over the last year, in which 67 million euros were sold.'

(38) ...no se nos **pagó** al término de las etapas acordadas y pese a ello nosotros proseguimos...
'...we weren't paid at the end of the agreed upon stages, and yet we did continue...'

Examples for each constructional type in Table 2

4 argument constructions

(1) La compañía continuó con su fiebre de adquisiciones, al anunciar que va a **comprar** las acciones que aún no posee a Indigo NV por 629 millones de dólares.
'The Firm continued with its purchasing fever, announcing that it is going to buy from Indigo the shares it does not yet possess, for 629 million dollars.'

(2) ...porque Talsud ofertó **cobrarles** 43 pesos por usuario (unos cinco dólares).
'...because Talsud offered to charge 43 pesos per user (about five dollars).'

3 argument constructions
Direct constructions

(3) El dueño de El Universal **compró** al gobierno de Sonora el periódico El Sonorense...
'The owner of the Universal bought the newspaper El Sonorense from the government of Sonora...'

(4) ...ambos equipos **habían comprado** sus respectivas cartas de transferencia en el pasado Mercado de Transferencia de Jugadores.
'...Both Teams had bought their respective transference cards in the past Players Transference Market.'

(5) ...Si usted **compró** una impresora Epson Stylus Color 740 hace un año por 199 dólares...
'...if you bought a Epson Stylus Color 740 printer a year ago for 199 dollars...'

(6) A las 8,000 ó 10,000 personas que asisten a las tribunas populares, les **cobramos** entre 5 y 7 pesos.
'We charged the 8,000 or 10,000 people in the cheap stands between 5 and 7 pesos.'

(7) Y los sueldos también se dispararon: Romario **cobraba** en el Valencia 3 millones 500,000 dólares.
'And the salaries also went high: Romario was paid 3 million 500 000 dollars with the Valencia Team.'

(8) ...tenían cargos de "gestores sociales" y **cobraban** por ello 490.20 pesos mensualmente.
'...They held positions as 'community agents' and charged 490.20 pesos monthly.'

(9) Al término de la discusión, les **compró** en dólares, como ellos querían.
'At the end of the discussion, (he) bought them (something) in dollars, as they wanted.'

(10) Los mensajes empezarán a trasmitirse por Televisa, Televisión Azteca y Multivisión, las cuales no **cobrarán** a la arquidiócesis por sus servicios.
'The messages will begin to be transmitted by Televisa, Televisión Azteca y Multivisión, and they won't charge the Archdiocese for their services.'

(11) ...cuando la besamos o cuando le **compramos** los zapatos que tanto necesita.
'...when we kiss her or when we buy her the shoes she badly needs.'

Passive constructions

(12) ...productos naturales que **son comprados** en esas comunidades por empresas que los envasan y los **venden**.
'...the natural products that are bought in those communities by companies that pack them and sell them.'

(13) Los montos de endeudamiento **fueron cobrados** por los bancos nacionales por otorgar préstamos en divisas estadounidenses.
'The amounts of the debts were collected by the national banks in exchange of granting them loans in US currency.'

2 argument constructions:
Direct constructions:

(14) ...porqué el anuncio de que HP **comprará** Compaq no inyectó alegría en los inversionistas bursátiles.
'...because the announcement that HP is going to buy Compaq didn't inject happiness in the stock market investors.'.

(15) La Selección *cobra* una cantidad fija, sin importar si se registra un lleno.
'The National Team charges a fixed fee, it doesn't matter if the game is sold out or not.'

(16) ...convenios con 60 países a quienes **cobramos** y nos cobran.
'...trades with 60 countries to which we charge and which charge us.'

(17) ...serán los portales que provean el contenido los que **cobren** por sus servicios.
'...The sites that supply the contents will be the ones that are going to charge for their services.'

(18) ...los interesados puedan consultar sus catálogos y **comprar** en el portal.
'...interested people can consult its catalogues and buy on the website.'

(19) ...su fabricación cuesta en pesos mexicanos pero ellos **cobran** en dólares.
'...its manufacture costs Mexican pesos but they charge in dollars.'

(20) ...porque es falso que el dinero todo lo compra. A mí no me **compran** con nada.
'...because it's not true that money buys everything. Nothing can buy me'

Passive constructions

(21) La disputa surgió después de que el banco mexicano Confía, **fuera comprado** por Citibank en 1997.
'The dispute arose after the Mexican Bank Confía was bought by Citibank in 1997.'

(22) . . . tierras y huertas que **habían sido compradas** a los indios de Tacuba.
'. . . lands and orchards that had been bought from the natives of Tacuba.'

(23) . . . el manejo de las cuotas sindicales que **fueron cobradas** desde 1978 por la administración. . .
'. . . the handling of the union fees that were collected since 1978 by the administration. . . '

(24) . . . la satisfacción de los servicios adquiridos y por producto que **se ha comprado** a Hewlett.
'. . . satisfaction with the acquired services and with the product bought from Hewlett.'

(25) Es muy probable que pronto estas máquinas mágicas puedan **comprarse** en la tienda de la esquina.
'It is highly probable that these magical machines can soon be bought in the store on the corner.'

(26) El 15 de noviembre de 1989, **se compró** la empresa en 18,520 millones de pesos.
'On November 15th1989, the Company was bought for 18,320 million pesos.'

(27) . . . se construirán vialidades de segundo piso y **se cobrará** una cuota por circular.
'. . . second floor roadways will be built and a fee will be charged for driving on them.'

1 argument constructions:
Direct constructions:

(28) Marcia **cobrará** a más tardar el próximo 4 de enero.
'Marcia will collect (a salary) by the upcoming fourth of January at the latest.'

(29) Sabe muy bien que el dinero puede **comprar** y no duda en utilizarlo para eso.
'(He/she) knows very well what money can buy, and (he/she) doesn't hesitate in using it for that.'

Passive constructions

(30) . . . porque **fueron comprados** algunos empleados de los aeropuertos.
'. . . because some airport employees were bribed.'

(31) Ese dinero ya no tenía que **ser cobrado**, puesto que la empresa ya no estaba dentro del contrato de compra-venta.
'That money didn't have to be collected anymore, because the Company no longer figured in the sales contract.'

(32) . . . otra factura que **se ha cobrado** tiene que ver con la seguridad.
'. . . another bill that has been charged relates to security.'

(33) Los primeros días **se cobrarán** 280 pesos, después la tarifa aumentará.
'280 pesos will be charged in the first days, later the fee will be raised.'

(34) En este país **se le cobra** a todos, menos a los que sí tienen que pagar.
'In this country everybody is being charged, except those who do have to pay.'

What can be preliminarily observed in Tables 1 and 2 is that these constructional horizons have a very broad range of clause types. They imply huge constructional richness. The one in

Table 1 includes 38 constructions and the other one, in Table 2, has 34. It is necessary to point out that all these constructional patterns are very general, in the sense that they don't show important coding differences, like the type of preposition that appears in each of the oblique argument PPs.

The presence in the corpus of clauses with four, three, two or just one participant shows the broad structural possibilities that speakers have to use the analysed predicates. For example, passive constructions account for 18.3% of the total cases. Beyond this, it is important to notice the extended interplay each of the four principal participants has in the constructional horizons. All appear in very different combinations of 4, 3 or 2 of them, in different syntactic configurations, all of them with the possibility of performing different syntactic functions (S, DO, IO or OA).

Next, I first report some general characteristics of these semantic-syntactic configurations and then, in Section 4, I will detail the syntactic performance of each of the participants.

Although these verbs are commonly treated as two place transitive predicates (in the sense that they seem to need no more than two obligatory arguments), we can see that they do appear, in real use data, in 4 argument constructions: *pagar* in 12.7% of the cases, *comprar* in 7.3%, *cobrar* in 10.7% and *vender* in 6.0% of the cases. These are not high percentages, but they certainly are significant, if we consider that pragmatic conditions are supposed to impose a limited preferred argument structure of one or two participants (Du Bois 2003). More importantly, these 4 argument constructions appear in two different configurations: one with the theme as the DO and the money as an OA, and the other one with the money as the DO and the theme as an OA (constructions (1) and (2), respectively, in Tables 1 and 2). This variability shows the productivity of the 4 argument pattern beyond its frequency.

As for the 3 argument constructions, they appear in 46.5% of the total cases, in a range that goes from 66.7% with *comprar* to 36.6% with *vender*, *pagar* and *cobrar* fall in-between with something around 40% of the cases. It is important to notice that this 3 argument pattern doesn't imply the loss of one specific participant, as any combination of them can be attested. We have combinations of agent, theme and recipient/source; agent, theme and money; agent, money and recipient/source, etc. All of these combinations also appear in different combinations of grammatical relations: S, DO and IO; S, DO, and OA, and S, OA and IO. But if we consider the fact that some of the arguments can fulfil two or three of these grammatical relations (for example the theme and the money can be the DO, OA or S) we have a total of at least 9 different active or direct three argument constructions.

So considering the higher frequency of this type of pattern and its argument variability in semantic and syntactic terms, one can claim that these verbs we are analysing are three argument predicates. But which three of the four participants do we have to postulate as part

of the LS? Clearly the four of them are needed to explain the broad range of 3 argument constructions the speakers use with these verbs. The potential for the appearance of all the participants remains even in the passive construction, as there are also clauses of this type with 3 arguments in different combinations (although they represent just 1.7 of the total cases).

Moving on to 2 argument constructions, they also represent, as it is expected, a large part of the cases: 35.2%; but as we have just seen, they are less frequent than the cases of 3 argument clauses (46.5%). They go from 42% of the cases with *vender* to 23.3% with *comprar*. As can be seen, these two verbs, *comprar* and *vender* are complementary in terms of the frequencies in which they appear with 3 or 2 arguments. In my corpus, *comprar* is clearly more oriented to three argument constructions as *vender* is to two argument clauses. Curiously, *pagar* and *cobrar* show almost equal percentages of these two types of constructions.

Again, different combinations of participants are possible: agent and theme, agent and recipient, agent and money, even money and theme, in a direct construction, as in the examples (20) and (21) for Table 1, and in the example (20) for Table 2. All the same, we can find in the corpus examples of different combinations of grammatical relations: S and DO, S and IO and S and OA, all of which can be fulfilled by different participants: again, at least the theme and the money can be the S, DO and OA. The total number of direct clause patterns with 2 arguments is 8.

Interestingly, passive constructions with two participants are equally well represented; they are almost 11% of the total cases and they account for 30.3% of the 2 argument constructions. *vender* is the predicate that is more oriented to this kind of clause (15.3% of the cases with this verb). There are examples in the corpus of both types of passive clauses that are possible in Spanish: the periphrastic passive and the *se* reflexive passive. This shows, in part, the structural flexibility that the predicates under study have. Speakers can easily focus on each of the participants by way of different structural options. It also shows that although passive constructions typically are, in typological terms, a structural mechanism for reducing the syntactic valence to one or none arguments, in Spanish, in the case of the verbs under study, in real use data, they seem to be a structural way for changing the prominence on the participants but not necessarily to completely reduce the syntactic combinatorial possibilities of these predicates.

As seen before, one argument constructions are also possible. They are in a total of 9.2% of the cases and go from 15.3% of the cases with *vender* to just 2.7% with *comprar*. *Pagar* and *cobrar* have the same percentage (9.3%). Most of these cases correspond to passive clauses (67%), so direct constructions are few (33%), but their presence in the corpus shows that also *bona fide* transitive predicates can be used, if the communicative context allows it, with just one argument. Interestingly, this one argument can be any of the four participants: the

agent, the theme, the recipient or source, and the money. This, again, as with the case of 4 argument constructions, shows that this kind of structural behaviour, is not an accident or an anomaly but one more open structural option for the speaker. Ultimately, this indicates that the 'obligatorily coded' criterion for defining what an argument is, is not a valid one, when confronted or tested against the broad constructional possibilities predicates show in real use data.

4 Syntactic performance of the four arguments

4.1 The case of the recipient/goal and source arguments

Now, I will show what the syntactic performance is for each of the participants individually. Let's begin with the data corresponding to the recipient or goal argument, in the case of *vender* and *pagar*, and to the source argument, in the case of *comprar* and *cobrar*. The recipient argument appears with the former ones in 39.3% of the total cases for both predicates. It is usually coded as an IO (83% of the cases), that is, as a core argument, but it can also appear, when its referent is inanimate, as an OA introduced by the preposition *en* (17%), as in the example 17 for Table 1; in such cases, this argument is semantically a goal, and syntactically it functions as an oblique argument in the core.[1]

Vender has this argument in 42.6% of the cases, in 16 of the 29 structural configurations

[1] Despite these differences in their animacy and in their syntactic coding, we can say that the goal and the recipient still fulfil the basic role, as both are instances of the same frame participant. For example, with *vender* 'to sell' they stand for the buyer participant, and as such, they both can also alternatively be the subject of *comprar* 'to buy', as in the examples (ib) and (id):

i. a. Pedro vendió un carro a María.
'Pedro sold a car to Mary.'
b. María compró un carro a Pedro.
'Mary bought a car from Peter.'
c. Este año México vendió muchos carros en el mercado europeo.
'This year Mexico sold a lot of cars in the European market.'
d. Este año el mercado europeo compró muchos carros a México.
'This year the European market bought a lot of cars from Mexico.'

This alternation in the coding of the goal-recipient arguments in terms of the animacy feature is very common in Spanish and in a lot of languages. In Spanish, this differential marking also shows up, for example, with predicates of change of location as *poner* 'to put':

ii. a. Pedro puso la medalla en un cuadro.
'Peter put the medal in a frame.'
b. El presidente puso la medalla al general.
'The president put the medal on the general.'

it presents in Table 1 (55%). The recipient/goal appears in clauses with four, three or two arguments and even in sentences where it is the sole argument, as in example (38) for Table 1. It is coded as the IO in 69% of the cases and 31% as an OA.

As for the recipient of *pagar*, it appears in 36% of the cases, in 11 of the 29 (38%) semantic-syntactic patterns it shows up; and, again, it can be projected in clauses with four, three, two or one argument, but it always gets coded as an IO (100% of the cases).

In sum, we can say that this participant, the recipient, has an important presence in the corpus, which indicates that the speaker tends to include it when talking about selling or paying. Getting coded in one third of the cases, in different combinations of participants, and mostly as an IO, which in Spanish is a core function, clearly is not the expected behavior for a non-argument.

Moving on to the source participant of *comprar* and *cobrar*, we can see that it is projected in 36.3% of the total cases of these two predicates, and it is coded as an IO in 81% of the cases and as an OA in 19%. With *comprar* it appears in 34% of the cases, in 11 of the 22 (50%) clause patterns this verb has. It shows up in sentences with four, three and two arguments, but not in those with just one argument (although this is perfectly possible). The source can be coded as an IO (72.5%) or as an OA (27.5%), introduced by the preposition *en*, just as the recipient is. On the other hand, *cobrar* presents this argument in 38.6% of the cases, 88% of them as an IO and 12% as an OA. 12 of the 26 semantic-syntactic configurations this verb appears in (46%) have a source. And this participant shows up in all kind of constructions: with four, three, two and one argument.

In sum, it can be said that, although they appear less frequently in the corpus, with percentages of coding that go from 42.6% to 34% of the cases, and being mostly coded as an IO, both the recipient and source are very productive in the constructional horizons of these verbs. We can say that they certainly are arguments and part of the LS of these predicates.

4.2 The case of the money argument

Now, what about the status of the money participant? Well, it also has high rates of projection with all the predicates under study. In total, it gets coded in 49% of the cases, more than the ones the source and the recipient have.

Of course, as expected, it is with *pagar* 'to pay' and *cobrar* 'to charge' that this participant has more presence: it appears in 68% of the cases with this last one predicate and 65.3% with the former, but the percentages that *vender* (27.3%) and *comprar* (35.3%) show are also significant. It gets projected in 13 of 29 (45%) of the clause patterns that *vender* has; in 15 of 29 (52%), in the case of *pagar*; in 13 of 26 (50%) in the case of *cobrar*, and in 8 of 22 (36%) in the case of *comprar*.

As for the grammatical relations this participant can fulfil, it can be a DO (56.1%), an OA (34.3%), and the S (9.6%), mostly in passive constructions in this last case, but as well in direct ones, as in the example (20) in Table 1, although in very limited quantities. All the same, as it is expected, *pagar* 'to pay' and *cobrar* 'to charge', have this argument mostly coded as a DO, 64.4% and 84.3% of the cases, respectively. On the other hand, *vender* and *comprar* mostly code it as an OA, in 63.4% and 90.6% of the cases, respectively.

So, taking into account that it appears in 49% of cases and that it has very productive variability to be coded in different nuclear grammatical relations, we can say that the money is an argument of these predicates. In syntactic terms, it is a core argument, when coded as a DO, and an oblique argument in the core when coded as an oblique.

4.3 The case of the theme argument (the goods)

As for the theme argument, the goods, it is projected in 79.5% of the total cases. It is a very productive participant in syntactic terms. Of course, *comprar*, with 95.3% of the cases, and *vender* with 90%, are the two predicates that have more examples with the presence of this participant. But *pagar*, with 76.6% of the cases, and *cobrar* with 56%, also show very high percentages. In most of the clauses in our corpus, the theme is coded as a DO (56.1%), which is exactly the same percentage for the money in this syntactic function; but it also appears as an OA (26.9%) and as a S (17%), mostly in passive constructions, but also, at least with *vender*, in direct clauses, as in the example (19) for Table 1. Complementarily to what can be seen in the case of the money argument, the theme appears mostly coded as a DO with *vender* (62.2%) and with *comprar* (86%). *Pagar* in 54.8% of the cases and *cobrar* in 69% have this argument mostly coded as an OA. It is also interesting to observe that *vender* has its theme coded as the subject in 34.8% of the cases.

So, as it is expected, this argument has a lot of prominence with these predicates. More than the money (49%), but both are equally important, as can be seen from the fact that they appear together, in the same clause, in 33.6% of the global cases, especially with *pagar* which in almost 50% of the cases is coded with the presence of both arguments. It is also important to note that, whether the theme or the money is being coded, a PP appears in almost 40% of the cases (38.16%). This frequency allows us to consider it, in syntactic terms, a core PP.

4.4 The case of the agent argument

Of course, the agent is expected to have the higher frequency of coding. It does. It appears in 82% of the total cases. But the difference with the percentage that the theme shows is not that big, as this last argument is coded in 79.5% of the cases. In most of the clauses where

it appears, in 98% of the cases, the agent is the subject. As said before, there is an 18% of passive cases that justly accounts for the 18% of the clauses where the agent is not present.

There is nothing more to say about the agent. But what is important to note here is that, although passive clauses suppose the intervention of an additional constructional scheme, the presence of 18% of passive cases shows that when using these predicates in communicative routines, the speakers, for different pragmatic motivations, can leave out of the coding process this so called obligatory participant, the agent. And they do so with a certain frequency (in almost 35% of the cases with *vender*). In fact, the speakers can leave out any semantic participant they want to. So obligatory coding vs. optional coding cannot be an appropriate criterion for determining an argument status. What one can see in the broad constructional horizons the verbs under study have, is that there are some participants that are recurrent. Four in this case. They show important percentages of coding, but none of them is strictly necessary in syntactic terms. Nevertheless, from a semantic point of view, they all are needed by the speakers, in a different proportion, to adequately complete the representation of the events they are talking about.

5 Conclusions

So going back to our working questions: which are the actual participants that have to be postulated as part of the LS of the predicates under study? And what is the relation between the cognitive frame and the language semantic-syntactic frames? We can say that, at least in this case, the four participants in the frame, the seller, the buyer, the money and the goods, have real semantic and syntactic value. They are part of the event denoted by these predicates; they are productively coded, that is, they have significant frequencies of projection; they appear in different configurations or combinations of participants, and they can be coded in different syntactic functions; they are all arguments, although they are not syntactically obligatory.

At the same time, the frame participants have a particular identity in the LS (or the argument structure of other theories) of the verbs; an identity that is conditioned not by cognitive principles, but by semantic and syntactic ones. As it was pointed out before, on the one hand, the seller has an agent identity with *vender* and *cobrar*, but it functions as a source with *comprar* and as a goal/recipient with *pagar*. On the other hand, the buyer is an agent with *comprar* and *pagar*, but a goal-recipient with *vender* and a source with *cobrar*. So, although the LS and the cognitive frame have the same number of participants, these have a different identity in each of those structures. In this sense, both the cognitive frame and the LS are needed to explain the behavior of the verbs under study.

As a way of capturing these facts and as a way of reflecting the constructional possibilities these predicates have, we can posit, in a preliminary fashion, something like the following mega logical structure:

([(**do'**(x, ∅)) CAUSE ((BECOME **have'** (z, **goods'**(w))) ∧ (BECOME **have'** (x, **money'**(y)))
∧ (BECOME **Not have'** (z, **money'**(y))) ∧ (BECOME **Not have'** (x, **goods'**(w))))]
∧ [(**do'**(z, ∅)) CAUSE ((BECOME **have'** (z, **goods'**(w))) ∧ ((BECOME **Not have'**(x,
goods'(w))) ∧ (BECOME **have'** (x, **money'**(y))) ∧ (BECOME **Not have'** (z, **money'**(y))))])

From the first causative segment,[2] selecting one, two or three of the BECOME modified predicates, we have the projection of both *vender* and *cobrar*. By the same procedure, the second causative segment gives us *comprar* y *pagar*.

Now, how does the linking process operate? Let's take the case of *pagar* as an example. Taking the second causative segment from the LS above, we have this next one:

Pagar LS:
(**do'**(z,∅)) CAUSE (BECOME **have'** (x, **money'**(y))) ∧ (BECOME **have'**(z, **goods'**(w)))
∧ (BECOME **Not have'**(z, **money'**(y))) ∧ (BECOME **not have'**(x, **goods'**(w)))

The selection of at least two of the BECOME modified predicates yields the following clauses. Both are perfectly possible, but, as we have seen, they are not favored in the same frequency:

(2) a. Pedro pagó 2000 pesos a Julia (por la camisa).
'Peter paid Julia 2000 pesos (for the shirt).'

b. Pedro pagó la camisa a Julia (en/por/a 2000 pesos).
'Peter paid Julia the shirt (for 2000 pesos).'

For getting the first clause, *Pedro pagó 2000 pesos a Julia (por la camisa)*, the speaker has to select the following LS:

(**do'**(z, ∅)) CAUSE (BECOME **have'** (x, **money'**(y))) ∧ (BECOME **have'** (z, **goods'**(w)))

In consonance with the RRG linking system and the actor-undergoer hierarchy, the (first) argument of **do'**, the agent, is the unmarked choice for actor, and as such is chosen as the privileged syntactic argument (PSA) and gets a subject coding. As for the undergoer, the first BECOME segment that is selected, the one more to the left, is the one that has priority of

[2] I'm considering commercial event verbs as a special type of transfer verbs, as they imply that 'someone gives money to someone else that in exchange gives some goods'. In this sense, I'm following Van Valin and La Polla (1997) in treating them as causative predicates of change of possession.

syntactic projection, hence, its second argument is the one selected as undergoer, and as such, it gets coded as a DO; in the same fashion, its first argument is coded as an IO, as recipients generally does in Spanish with transfer verbs. Then, if coded, the second argument of the second BECOME segment is projected as an oblique argument, a pp. From this same LS come other valence reduced constructions as *Pedro pagó 2000 pesos* ('Peter paid 2000 pesos'), *Pedro pagó* ('Peter paid') and *Pedro le pagó a Julia por la camisa* ('Peter paid Julia for the shirt'), which are possible in discourse contexts where the missing arguments are recoverable or have a somehow generic interpretation.

For the second clause, *Pedro le pagó la camisa a Julia (en/por/a 2000 pesos)* the speaker selects this LS:

> (**do'**(z, ∅)) CAUSE (BECOME **Not have'**(x, **goods'**(w))) ∧ (BECOME **have'** (x, **money'**(y)))

Following the same procedure as above, the argument of **do'** is selected as actor and it is coded as subject. From the first BECOME segment, the second argument, the goods, is selected as undergoer and then projected as the DO, and the first argument, the recipient, as an IO. All the same, the second argument of the second BECOME segment, the money, if projected, is coded as an oblique argument. As above, clauses with less arguments are projected from this same LS if context allows it. As can be seen, the four arguments are possible choices for projection, even in different syntactic functions, but following the same actor-undergoer hierarchy, there are some linking possibilities that are ruled out:

(3) a. *Julia pagó la camisa (Julia is the seller)
'Julia paid the shirt' (Julia is the seller)

b. *La camisa pagó los 2000 pesos
'The shirt paid the 2000 pesos'

The example (3a) is ungrammatical as the first argument of a stative predicate is always overridden for the actor macrorole by the first or the sole argument of a **do'** predicate. For the same reason, the second argument of the stative predicate cannot be selected as PSA. [3]

[3] Nevertheless, in our corpus data there exist clauses like *Los 2000 pesos pagaron la cuenta* 'the 2000 pesos did pay (for) the bill', which implies that a different linking has taken place. There are just a few examples in the corpus, but speakers tend to find them possible in some contexts, though they consider them very rare. And there is a good reason for this: in the first place, this kind of constructions would have to be ruled out for the same reason that examples like (3b) are ungrammatical. Interestingly, the same happens with *comprar*: Clauses that have the goods argument as subject are ungrammatical, but the ones with the money argument

Now, what is the status of the particular syntactic frames each verb tends to favor? As previously noted, Fillmore, as probably most of the syntacticians do, believes these to be the crucial linguistic information to explain the verbs behavior. They are related to the cognitive frame, yes, but the later remains in the background and it seems that the important thing to consider is which participants are the ones coded in the nuclear functions of subject and direct object with each verb. As the data offered in this work show, the syntactic frames favored by each verb are only a portion of a set of broader possibilities, that is, what we call here the constructional horizon of a verb. Many constructional patterns are attested and each one plays an important part in the actual use those verbs have in real communicative contexts. In this sense, the favored syntactic frames don't have a special status, they are just the more frequent ones. Of course, the relative frequencies that each predicate shows for every type of construction permit us to determine the differences among the predicates, which we can call the specific semantic and syntactic vocation of each predicate, but for reasons of space, we leave the analysis of these differences for further work.

So summing up, three levels of analysis are needed to describe the syntactic behavior of a predicate: (1) The actual syntactic constructions which conceptually activate a certain state of affairs; (2) the cognitive frame, which functions as the whole scenario against which the conceptual scene activated by a certain clause is interpreted and understood; and (3) the LS or argument structure, which provides the number of arguments and their semantic and syntactic identity. Of course, the most important and linguistically significant level is the LS as it encapsulates all the actual constructional possibilities of a predicate and as it specifies the arguments needed to produce these potential syntactic possibilities, arguments that don't need to correspond in number and type to the participants in the frame.

---

coded as subject are possible. The same happens with *vender* and *cobrar* but the other way around: constructions with the money participant as subject are not allowed, but clauses with the goods argument coded as subject are grammatical.
One possible explanation for the existence of this type of clauses is that they can be the result of the implementation, during the linking, of a constructional scheme that precludes the agent argument of being projected from a LS as the following:

do'(z, (**give'**(z, **money'**(y)))) CAUSE (BECOME **Not have'** (z, **money'**(y))) ∧ (BECOME **have'** (z, **goods'**(w)))

This is a somehow different LS from the ones previously presented. This new one has the predicate **give'** (with its respective arguments) as the second argument of **do'**. This is not an arbitrary addition as a commercial event can be viewed as a double transference process: someone gives money to somebody else, and in exchange, this somebody else gives some goods back to him. So what the agent argument does, in the case of *pagar* and *comprar* is to give some money.

Now, if the speaker chooses that LS as a starting point and then she applies the constructional scheme, which prevent the z or agent argument from getting a syntactic realization, then the y argument, the money, has to be coded as subject and the w argument is projected as a DO. Of course, this suggestion needs a more detailed exploration, an issue that we leave for further research.

References

Du Bois, John W. 2003. Argument structure. Grammar in use. In *Preferred Argument Structure*, J. W. Du Bois, L. E. Kumpf & W. J. Ashby (eds), 11–60. Amsterdam/Philadelphia: John Benjamins.

Fillmore, Charles. 1977. The case for case reopened. In *Syntax and Semantics* 8, P. Cole & J. M. Sadock (eds), 59–81. New York: Academic Press.

Ibáñez, Sergio. 2008. Saying verbs in Spanish. Deepening the lexical semantics. In *Investigations of the Syntax-Semantics-Pragmatics Interface*, R. D. Van Valin Jr. (ed.), 3–21. Amsterdam/Philadelphia: John Benjamins.

Mairal, Ricardo. 2003. Why the notion of lexical template? *Anglogermanica on Line* 2002(1). https://www.uv.es/anglogermanica/2002-1/mairal.htm

Mairal, Ricardo. 2004. Las reglas léxicas en la Gramática del Papel y la Referencia. In *Algunos Problemas Específicos de la Descripción Sintáctico-Semántica*, G. Wotjak & J. Cuartero (eds), 175–196. Berlin: Frank Timme Verlag.

Mairal, Ricardo & Faber, Pamela. 2002. Functional grammar and lexical templates. In *New Perspectives on Predicate Structure in Functional Grammar*, R. Mairal & M. J. Pérez (eds), 41–97. Berlin/New York: de Gruyter.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax, Structure, Meaning and Function*. Cambridge: CUP.

French and Spanish 'MAKE/GIVE + FEAR'-type LVCs – an RRG Constructional Account

Eva Staudinger
Albert-Ludwigs-University Freiburg

Abstract

This paper analyses Spanish *dar miedo* ('to scare'; lit. 'give fear') and French *faire peur* ('to scare, to frighten'; lit. 'make fear') type light verb constructions (LVCs) and provides a description of their syntactic characteristics. In both languages, the LVC types in question are based on a causative light verb (LV) and an emotion nominal which are tightly connected. The predicates are semi-compositional and show limited syntactic flexibility in both languages, but French LVCs are syntactically less flexible than their Spanish counterparts. These differences in syntactic flexibility are attributed to general language-specific constraints on word order flexibility. A description of LVCs which uses constructional schemas is proposed. The constructional schemas include information on the general principles that apply as well as the specific properties and constraints of LVC types. The constructional schemas are assumed to be linked with the lexical entries of specific LVCs.

1 Introduction

Light verb constructions (LVCs) – also referred to as compound verbs or support verb constructions – are complex predicates consisting of a light verb (LV) and a nominal predicate (NPRED), which specifies the type of event designated.[1] The term LVC is a cover term for complex predicates whose properties vary both within an individual language and across languages. For English, for example, Brinton & Traugott (2005) distinguish the *lose sight of* class and the *take a look at* class. LVCs of the *lose sight of* class are based on a comparably specific verb. They often have a highly idiomaticized meaning and a low pattern of productivity. LVCs of the *take a look at* class are based on light verbs with a rather general semantics

[1] There are many terms to designate complex predicates of different types. In French, for example, LVCs of the LV + NPRED type are discussed under the term "locution verbale" (e.g. Rohrer 1967),"expression à verbe support" (e.g. Marchello-Nizia 1996), "verbe complexe" (Valli 2007) or "construction à verbe support" (Schøsler 2008; Vivès 1993).

such as *make, take, give, have* and *do*. These LVs form a fairly stable set and occur with a range of different nominal predicates while exhibiting recurring and productive patterns (cf. Brinton & Traugott 2005: 131; Brinton 2011: 567–568). The number of nominal predicates used with these LVs increases over time (cf. Brinton 2011: 568).

Likewise, French and Spanish have LVC classes with specific and general verbs, which in turn can be subdivided into different categories according to their structural properties (NP with determiner or bare noun), the semantic class of the NPRED, or according to the productivity of the pattern they follow. The present article deals with particular subgroups of LVCs in French and Spanish consisting of the semantically general LVs *faire* ('do, make') and *dar* ('give') and a nominal predicate denoting an emotional state. Examples from French are *faire envie/honte/mal/peine/pitié/peur/plaisir* ('to whet (so.'s) appetite/'to embarrass'/'to hurt'/'to scare'/'to please'), and examples from Spanish include *dar asco/envidia/miedo/rabia/vergüenza* ('to make (so.) happy'/'to disgust'/'to make (so.) envious'/'to scare'/'to enrage'/'to embarrass').[2] I will henceforth refer to these LVCs as the *faire peur* and the *dar miedo* type. These classes of LVCs are subclasses of the LV + bare noun type in the respective languages. They are based on patterns that resemble those of free constructions, but they have very limited productivity (if any).[3]

LVCs have been studied from various perspectives. Synchronic studies on LVCs in French and Spanish have focussed on the syntactic and collocational properties of LVCs (e.g. Alonso Ramos 2004; Cuervo 2010; Detges 1996; Gledhill 2008; Romero Méndez 2007; Rohrer 1967).[4] Within the context of French linguistics, the syntactic behaviour of LVCs was a major research topic in the theoretical framework of *Lexique Grammaire* (Danlos 1992, Giry-Schneider 1987, Gross 1981, Vivès 1993). Studies in the context of grammaticalisation theory addressed the grammatical status of light verbs (e.g. Ströbel 2010), and aspects of the diachronic development of light verb constructions (e.g. Alba-Salas 2007; Schøsler 2008).[5]

Recent work on LVCs highlights the fact that LVCs are difficult to handle in approaches where lexicon and syntax are considered separate components (Gledhill 2008; Family 2014; Schultze-Berndt 2006; Wittenberg *et al.* 2014). Starting out from this observation, this article

[2] Literally: 'to do/make desire/shame/pain/pity/fear/pleasure' and 'to give disgust/envy/fear/anger/shame'.

[3] 'Productivity' in this context means that the combination between LV and NPRED is not collocationally restricted. *Dar miedo/faire peur* type LVCS are collocationally restricted to a certain number of emotion nominals. There are, however, cases where the position of the NPRED is filled by the synonym of an established NPRED, e.g. *faire vergogne* ('to cause shame'; cf. *faire honte* – 'to cause shame').

[4] For a discussion of the structural properties of LVCs in Old and Middle French cf. Bischof (2007), Marchello-Nizia (1996) and Valli (2007). For a cross-linguistic discussion of LVCs cf. Butt (2010) and Schultze-Berndt (2006).

[5] The diachronic development of LVCs exhibits both universal and language-specific tendencies. For English, Brinton & Traugott (2005) and Brinton (2011) suggest that LVCs are a product of both grammaticalisation and lexicalisation.

will discuss the characteristics of French and Spanish 'MAKE/GIVE FEAR'-type LVCs and provide a constructional RRG description.

2 Preliminary remarks on LVCs and their definition

Given that this is an article about LVCs, I should start by defining the term and explaining how it is used within linguistics in general, and this article in particular. However, as Winhart (2005: 1) and Family (2014: 5) point out, there is no universally accepted definition of the term. The criteria proposed to identify LVCs only partially coincide. There are several reasons for this, but I will limit this section to three remarks on this issue.

First of all, LVCs exist in typologically distinct languages, and their status and structure depend on the properties of the language they belong to. In languages where simple verbs form a small or closed class like Persian, the productivity and functional load of LVCs will most likely be different from that in languages where verbs are an open class, such as French, German or English.[6]

Secondly, LVCs exist in languages with different degrees of word order flexibility (cf. section 3.3). This affects argument realisation, and the syntactic behaviour of LVCs. In languages with rigid word order such as Modern Standard French, the syntax of LVCs seems to be more restricted than in languages with relatively free word order (e.g. German, Spanish, Old French).

Thirdly, the term "LVC" has been applied to different types of complex predicates not only across languages, but also within languages. Some use the term to designate complex predicates of the [V + N]/[N + V] type, and others use it for complex predicates consisting of two verbs ([V + V]; cf. Butt 2010: 48–49). These morphologically distinct types are again only cover terms for a range of different LVCs. V+ N compounds appear as [LV + abstract noun] compounds, [LV + deverbal noun] compounds, or the even more specific class of [LV + *nomen actionis*] compounds.[7] Given this situation, there is little to gain in systematically collecting and contrasting definitions from previous work. This is why I will limit myself to describing the properties of the type of LVCs that are the subject of this article.

3 LVCs of the 'MAKE/GIVE + FEAR'-type

The observations presented in the following are based on an analysis of 'MAKE/GIVE + FEAR'-type LVCs in the *Corpus del Español* (CDE; Davies 2002–) and FRANTEXT (ATILF

[6] According to Schultze-Berndt (2006: 366) it has been argued that simple verbs form a small or even closed class in Persian. Family (2014: 1) estimates that Persian has less than 200 simple verbs, whereas languages like English or French have more than 5000.

[7] Note that the criteria for defining LVs are also subject to discussion.

et al.).[8] The search period in the CDE was limited to the 20th century (20.5 million words). In the FRANTEXT data, the LVC type in question is much more frequent, and a smaller corpus sample was chosen. The French data was gathered mainly from the tagged corpus version for the search period 1980–2010 (7.3 million words; henceforth: FRANTEXT-CAT). For less frequent construction types (e.g. *Ça* + LVC + *de*-INF) and LVCs (e.g. *faire honte*), additional examples from the non-tagged version of FRANTEXT were considered.

In terms of size and text types, the Spanish and the French data are not comparable, so the focus will be on qualitative differences between French and Spanish rather than on differences in frequency of certain construction types.[9] In the Spanish data, the most frequent LVC of the *dar* + $N_{EMOTION}$ type is *dar miedo*, which is attested 180 times in the 20th century texts of the CDE.[10] This is why the present study mainly focusses on this LVC. In the French data, the most frequent LVC is *faire mal* ('to hurt'; 225 examples), which does not have a Spanish equivalent, followed by *faire peur* (190 examples).

3.1 Semantics

LVCs of the 'MAKE/GIVE FEAR'-type in French and Spanish are complex lexicalised three place predicates that are syntactically intransitive. They consist of a basic and fairly semantically general verb – fr. *faire* ('do, make') and sp. *dar* ('to give') – and a bare nominal denoting an emotional state, such as *peur* (fr.)/*miedo* (sp.; 'fear'), *honte* (fr.)/*vergüenza* (sp.; 'shame'), *plaisir* (fr.; 'pleasure') or *asco* (sp.; 'disgust'). In both languages, the LV carries a causative component, and there is a non-causative counterpart formed with the LV *avoir* ('to have') in French and *tener* (lit. 'to hold') in Spanish. This illustrates that there is something systematic about the patterns we find with these LVCs in the respective languages.

The LVs *faire* and *dar* also have full verb counterparts. As a full verb, Spanish *dar* is ditransitive. As a LV, *dar* combines with different types of predicative nouns (action nouns, state nouns) and forms ditransitive and intransitive predicates (cf. Cuervo 2010: 145–146). Psych-preds formed with *dar* (e.g *dar miedo*) are analysed as intransitive predicates by Cuervo (*ibid.*), whereas *dar apoyo* ('to support'), for example, is analysed as ditransitive.[11] French *faire* is transitive as a full verb, and can form transitive and intransitive two- and three-place predicates with predicative nouns.[12] In addition to that, *faire* can be used as a causativiser,

[8] In addition, observations on *faire mal* ('to hurt') from a previous analysis of French LVCs (cf. Staudinger 2013) are included.

[9] The CDE contains literary texts, newspaper texts and transcripts of spoken language of different geographical varieties of Spanish. The FRANTEXT sample only contains literary texts.

[10] In total, there are 360 instances of *dar asco/envidia/miedo/rabia/vergüenza* in the CDE data.

[11] In fact, *dar apoyo* occurs in the passive form.

[12] Examples for two-place predicates are *faire les courses* ('to do the shopping'), *faire face* ('face up to sth.')

and in colloquial French it also has copula-like functions. [13] It is an extremely versatile and semantically flexible verb.

The interpretations of *faire* and *dar* as a LV are highly context dependent and only preserve part of the semantics of the corresponding full verbs. This seems to be a common property of light verbs. Nolan (2014: 143) observes that the "light verbs of Irish are not bleached just for argument structure when they appear as part of a complex predicate but also for aspect and event structure". The semantic properties of the NPRED and of the arguments highlight particular semantic facets of the LV, which affects the interpretation of the predicate as a whole.

As shown in Staudinger (2013) for the French LVC *faire mal*, the lexical aspectual properties of LV + $N_{EMOTION}$ LVCs do not coincide with any of the established *Aktionsart* classes. The reasons for this were attributed, firstly, to the lack of spatio-temporal anchoring of the situations represented by psych-predicates, and secondly, to the co-compositional character of LVCs (cf. Staudinger 2013: 160–162). In LVCs, the semantic properties of a light verb and a state of affairs noun blend together to create one complex predicate meaning. Therefore, both the verb and the nominal predicate contribute to the *Aktionsart* properties (cf. Detges 1996: 155–156).[14] In the case of 'MAKE/GIVE FEAR'-type LVCs, the nominal part of the predicate refers to an emotional state and thus contributes a stative component to the complex predicate.

The nominal element does not refer to a specific instance of an emotional state. It rather contributes predicative information and characterises the type of emotional state. The fact that the predicative nominal can combine with the intensifying adverb *très* ('very') highlights the predicative character of the noun (cf. ex. (1a)). In Spanish, the situation with respect to modifiers is less clear. Modifiers such as *mucho*, *tanto* ('many'; 'very much') and *bastante* ('plenty'; 'quite') have adjectival and adverbial functions. However, the absence of a determiner underlines the predicative character of the noun.

---

and *faire un rapport* ('make a report'). For a more detailed analysis of the structure and function of VN-complexes based on French *faire* see Gledhill (2008).

[13] Cf. the examples in (i), both from Lauwers & Tobback (2010: 95).

i. a. ça fait chic
DEM make:3SG chic
'this looks fancy'

b. avec ses lunettes il fait très prof de fac
with POSS.PL glasses 3SG.NOM.M make:3SG very university teacher
'with his glasses, he looks very much like a university teacher'

[14] Detges (1996) opts for a two-dimensional analysis of the semantic properties of predicate classes. Following Koch (1981), he categorises predicates according to two semantic dimensions: type of state of affairs representation ("Art der Sachverhaltsdarstellung") and constitutive conditions ("konsitutive Sachverhaltsbedingungen") of the represented state of affairs. The semantic properties considered in his analysis partially coincide with those considered relevant for the assignment of *Aktionsart* classes.

The predicative noun is realised as a bare noun in most cases. When the NPRED is modified by an adjective or a PP, however, it requires an indefinite article. But semantic modification of the NPRED is rare and the possibilities seem to be limited. French (*faire*) *un mal de chien*/*une peur bleue* in ex. (3) are fixed expressions. Modifying adjectives specify the intensity of the emotion rather than its quality (cf. examples (2) and (3)).[15]

(1) French

a. ...ça fait mal, très mal.
DEM make:3SG pain INT pain
'...that hurts, that hurts a lot.' (Frantext: Déon (1965), *Le rendez-vous de Patmos*)

b. Le terme de métier me faisait très peur.
DET.M term of business me.CLI.DAT make:3SG fear.N
'The term "business" scared me.'
(Frantext: Bienne (1986), *Le silence de la ferme*)

(2) Spanish

a. ...a mí me dió una rabia inmensa.
to me.PRO.DAT me.CLI.DAT give:3SG.PST ART.INDEF.F rage immense:F
'It made me extremely angry.'
(CDE: 19-OR Habla Culta: Santiago: M6)

b. Y nos dio una vergüenza terrible...
and us.CLI.1PL give:3SG.PST ART.INDEF.F shame terrible
'And it embarrassed us terribly.'
(CDE: 19-OR Habla Culta: Buenos Aires: ..)

(3) French

a. ...ça m'a fait un mal de chien.
DEM me.CLI.DAT make:3SG ART pain GEN dog.
'...that hurt me incredibly.'
(Frantext: Pouy (1988), *La clé des mensonges*)

b. ...je me suis fait une peur bleue.
1SG.NOM 1SG.DAT be.AUX.1SG make:PPTC ART.INDF.F fear.N blue:F
'I scared myself to death.'
(Frantext: Djian (1985), *37.2° le matin*)

[15] FRANTEXT-CAT contains 190 examples of *faire peur*, but only in three examples is the NPRED modified by an adjective or a PP and thus combines with an indefinite article. Among the 180 examples of *dar miedo* in the 20th century data of the CDE, there are cases where *miedo* occurs with intensifiers, but there are no cases where the NPRED is modified by an adjective or a PP. Consequently, there are no instances of *dar miedo* where the NPRED combines with an indefinite article. This is why other examples of *dar* $+N_{EMOTION}$ LVCs were chosen for the illustration of modification.

Given the degree of lexicalisation of LVCs, on the one hand, that is highlighted by the restrictions on modification, and the language-specific semantic range of French *faire* and Spanish *dar* on the other, the translation chosen for the morphological glossing in English – 'do/make' and 'give' + NPRED – might not be adequate. The glossing of the separate elements highlights the structural properties of the predicate.

3.2 Argument structure and (morpho-)syntactic properties

Syntactically, LVCs of the 'MAKE/GIVE FEAR'-type are three place predicates involving two participants, a trigger of the emotion in the first argument position and an experiencer in the third argument position. The position of the second argument is not filled by an RP, but by a predicative noun that constitutes part of the predicate, so that strictly speaking, there is no second argument. For the semantic role corresponding to the trigger of an emotion, Kailuweit (2005) introduces the label CORRELATIVE (COR). As Kailuweit (2005: 84–85, 90–92) points out, the participant with the COR role is not strictly causative. Although it contributes to the emotional experience, it does not function as the sole cause for the emotion lived by an experiencer. For an experience to occur, the experiencer has to be aware of the 'cause' and evaluate it.[16]

Given that the predicate consists of two components, the question arises how both elements contribute to the argument structure. Following Detges (1996: 157–160), it is argued that the verb and the nominal predicate do not only jointly contribute to the meaning of the LVC, but also to the number of argument slots and their selectional restrictions. In the case of *dar miedo*, the LV *dar* contributes three argument slots. The first one is filled by the COR-argument, the second one by the predicative noun, and the third one by an EXPERIENCER (EXP) argument with dative case. The semantic role of the third argument and the semantics of the nominal predicate complement one another.

The case of French *faire mal* is different despite the surface similarities. French *faire* only contributes two argument slots for the syntactic frame, but LVCs of the *faire peur* type are three place predicates as well. The number of arguments and their semantic roles correspond to that of *dar miedo* LVCs. In French, the third dative EXP argument must have been integrated by other means into the predicate structure of the LVC. Given its semantics, and its preference to occur as a clitic, it might have its origins in a free 'affected' dative that was incorporated and is now part of the argument structure.[17]

[16] Similarly, Kutscher (2009: 67) characterises the relationship between experiencers and stimuli of emotions as bidirectionally causative, i.e. both the stimulus and the experiencer contribute to the emergence of an emotion.

[17] Unlike lexicalised datives that are part of the argument structure, free datives always appear as clitics and preferentially occur in 1st and 2nd person form (cf. Jouitteau & Rezac 2007: 98–99).

The position of the first argument – the COR – allows for animate and state of affairs arguments in both French and Spanish.[18] As Spanish is a pro-drop language, the COR argument need not be realised overtly (cf. ex. (4a')), while in French, an overt subject is required (cf. ex. (5a')). The COR argument can be realised as a pronoun, an NP or an extended infinitive. In spoken data and fictional representations of orality, clausal realisations of the COR are also attested (cf. examples (4c) and (5c)).

In French, several restrictions apply to the subject position and pronominal subjects are strongly preferred. In cases where the first argument (i.e. COR) is realised as an extended infinitive or a clause, it appears postverbally and the subject position is filled by an additional pronoun *ça/cela* (cg. ex. (5b)). In Spanish, word order is more flexible and no such additional pronoun is required.

(4) Spanish

a. Me da miedo la muerte...
me.CLI.DAT give:3SG fear.N DET.F death
'Death scares me'.
(CDE: 19-OR España Oral: CENT014A)

a'. ∅ me da miedo.
me.CLI.DAT give:3SG fear.N
'It scares me'.

a". La muerte me da miedo.
death me.CLI.DAT give:3SG fear.N
'Death scares me'.

b. ...a mi mamá le dió miedo ir nuevamente
to POSS.1SG mum her.CLI.DAT give:3SG.PST fear.N go.INF again
al mismo Hospital.
to:ART.SG.M same:M hospital
'My mum was scared to go to the same hospital again'.
(CDE: 19-OR Habla Culta: San José (CR))

c. ...pero a mi mamá le dió miedo que
but to POSS.1SG mum her.CLI.DAT give.3SG.PST fear.N that
me casara allá con un gringo...
me.CLI.AKK marry:1SG there with ART.INDEF.M gringo
'...but my mum was afraid that I would get married with a gringo there.'
(CDE: 19-OR Habla Culta: Caracas: M34)

[18] In the French data, animate subject referents are more frequent than in the Spanish data (14% vs. 3%). Although it cannot be excluded that the figures are influenced by the fact that the French and Spanish data come from different text genres, note that there are also qualitative differences in the data that can account for this difference (cf. section 3.3).

(5) French

a. ...ma tante Solange me faisait peur...
POSS.F aunt Solange me.CLI.DAT make:3SG.PST fear.N
'...my aunt Solange scared me...'
(FRANTEXT: Ernaux (1981), *La femme gelée*)

a'. Elle me faisait peur.
3SG.NOM.F me.CLI.DAT make:3SG.PST fear.N

b. Ça me fait pas peur de te laisser
that me.CLI.DAT make:3SG NEG fear.N to you.CLI.AKK leave:INF
sur le carreau!
on DET.M tile
'I'm not scared to kill you.'
(FRANTEXT: Thérame (1985), *Bastienne*)

c. ...cela me faisait peur quand je voyais
DEM me.CLI.DAT make:3SG.PST fear.N when 1SG.NOM see:1SG.PST
les autres envisager autre chose.
ART.PL other:PL envisage:INF other thing
'It scared me when I saw the others consider something else.'
(FRANTEXT: Berr (2008), *Journal*)

3.3 'MAKE/GIVE + FEAR'-type LVCs and word order flexibility

As mentioned in section 2, the general syntactic characteristics of a language, such as its word order flexibility, have implications for the behaviour of LVCs. This becomes evident when we look at the examples below taken from the CDE and FRANTEXT.

The position of the arguments in relation to the predicate is strictly regulated in French, whereas in Spanish, only arguments realised as clitics have a fixed position. While in Spanish the position of the first argument can be pre- or postverbal, in French, it has to be realised preverbally in subject position (cf. examples (6) and (7)).

(6) Spanish

a. Le daba miedo esa casa.
him/her.CLI.DAT give:3SG.PST fear.N DEM house
'That house scared him.
(CDE: 19-F *Mujeres al teléfono y otros...*)

b. Esa casa le daba miedo.
dem house him/her.CLI.DAT give:3SG.PST fear.N
'That house scared him.

(7) French

a. ...la nuit me fait peur, monsieur Jean.
DET.F night me.CLI.DAT make:3SG fear.N mister Jean
'The night scares me, monsieur Jean.'
(Frantext: Salvayre (1995), *La puissance des mouches*)

b.* Me fait peur la nuit.
me.CLI.DAT make:3SG fear DET.F night
'The night scares me.'

c. Elle me fait peur, la nuit.
3SG.F me.CLI.DAT make:3SG fear DET.F night
'The night scares me.'

Table 1 shows that the position of the COR-argument in the Spanish data varies. In the CDE sample, there are 90 examples for *dar miedo* where the COR-argument is lexically realised and where its position is not determined by constructional restrictions, as would be the case for relative sentences, for example. The data show a preference for postverbal subject realisation. As the sample is rather small, we should not attach too much importance to the exact figures. However, the high percentage of postverbal realisation of the COR argument is striking, and it is supported by Cuervo's analysis for the default word order with Spanish *dar miedo* type LVCs (cf. Cuervo 2010: 141).

	Frequency	**Relative frequency**	
DECLARATIVE SENTENCES			
COR not realised overtly	51		28.3%
COR lexically realised	**90**		50.0%
preverbal COR	**21**	**33.4%**	(11.7% of total)
postverbal COR	**69**	**76,6%**	(38.3% of total)
OTHER CLAUSE TYPES			
relative clauses	17		9.4%
interrogative clauses	12		6.7%
other	10		5.6%

Table 1: Preverbal and postverbal realisation of COR for *dar miedo*

As shown by Mayoral Hernández (2007: 227), a high percentage of postverbal realisation of subject arguments is typical of unaccusative verbs in Spanish. In line with this, Cuervo (2010: 146) proposes that *dar miedo* type LVCs are unaccusative predicates. Typically, unaccusative verbs are defined as intransitive predicates (e.g. *llegar* 'to arrive') with one argument that

is a direct object rather than a subject. Van Valin (1990) provides a semantic explanation for unaccusativity and attributes the syntactic behaviour of unaccusative predicates to the presence of a state predicate in the LS. Unaccusative predicates code a result state predicated of an undergoer (Van Valin 1990: 233).

I argue that the preference for postverbal realisation of the first argument with *dar miedo* type LVCs can also be explained by the structure of the predicate. The first argument is also the argument of a state predicate, however in this case, not an argument of a result state. The COR-argument is a stative cause, i.e. the LS of *dar miedo* contains a state predicate as a first component (cf. the constructional schema in Figure 1).[19] Arguments of state predicates are attributed the undergoer MR in linking.

Dar miedo type LVCs are non-prototypical causative predicates that trigger non-agentive readings of the first arguments, even in cases where the first argument refers to an animate being (cf. ex. (8)).

(8) Ø Me das miedo cuando ponés esa cara.
me.CLI.DAT give:2SG fear.N when put.2SG that.F face
'You scare me when you make that face.'
(CDE: 19-F *Los hombres de Celina*)

The claim that animate subjects are not interpreted as agents is also supported by the fact that for *dar miedo*, modal verb constructions with 'want'-verbs and purposive constructions are not attested in the CDE. The case of French *faire peur* LVCs is different (cf. examples (9) and (10)).

(9) Elle avait voulu me faire peur.
3SG.NOM have:AUX.3SG want:PPTC me.CLI.DAT make:INF fear.N
'She had wanted to frighten me.'
(FRANTEXT: Grèce (1982), *La nuit du sérail*)

(10) ([un authentique anthropophage] qui poussait des cris, des grognements, si on s'approchait trop près des barreaux,)
pour faire peur aux enfants.
in order to make:INF fear.N P.DAT.ART.PL children
'([a real man-eater] who let out shrieks and grunts when people came too close to the grille) in order to frighten the children.'
(FRANTEXT: Boudard (1995), *Mourir d'enfance*)

In addition to differences in type and position of arguments, we can also find different restrictions regarding the relative position of LV and predicative noun. In Spanish, there are

[19] In Staudinger (2013) I discussed quite extensively the semantics of the readings of the French LVC *faire mal* ('hurt') and the varying degrees of agentivity and causativity associated with different COR-arguments and predicate readings. LS representation as proposed in RRG does not reflect 'degrees' of more and less prototypical causatives.

examples where the NPRED precedes the LV, whereas in French, the LV always has to precede the NPRED (examples (11) and (12)). Although the inversion of LV and NPRED is rare, the Spanish LVC appears more flexible.[20]

(11) Spanish

a. ¿Qué es lo que tanto miedo te da?
Q be.3SG N REL INT fear.N 2SG.DAT give:3SG
What is it that scares you so much?
(CDE:19-F *El destino, el barro y la c...*)

b'. ¿Qué es lo que te da tanto miedo?
Q be.3SG N REL 2SG.DAT give:3SG INT fear.N
What is it that scares you so much?

(12) French

a. Qu'est-ce qui te fait peur?
Q:be.3SG-DEM REL 2SG.DAT make:3SG fear.N
'What is it that scares you?'
(Frantext: Labro (1982), *Des bâteaux dans la nuit*)

b'.* Qu'est-ce qui (tellement) peur te fait?
Q:be.3SG-DEM REL (INT) fear.N 2SG.DAT make:3SG
What is it that scares you (so much)?

The examples discussed in this section show that 'MAKE/GIVE FEAR'-type occur in different construction types, and the position of arguments is variable in Spanish. But they do not have full syntactic flexibility, neither in French nor in Spanish. As intransitive three place constructions, they exclude passivisation. In addition, pronominalisation and extraposition are restricted to the referential arguments (COR and EXP). The predicative noun can neither be pronominalised nor extraposed.[21] The link between the LV and the predicative noun is very tight in *dar miedo* (cf. ex. (13)). Other Spanish LVCs seem to be less restrictive, as Crego García (2000: 53) points out.

[20] Inversions of LV and NPRED are also attested with other MAKE/GIVE + FEAR-type LVCs in the CDE. In total, there are five cases.

[21] Other constructional preferences cannot be attributed to the properties of the LVC itself, such as the preference for a pronominal realisation of the EXP argument, which is roughly three times more frequent than its realisation as a lexical NP. This calculation is based on a sample of 560 examples from FRANTEXT (1980–2010) consisting of seven different *faire* + emotion-nominal LVCs. In the CDE, 360 *dar miedo* type LVCs were analysed. Lexical EXP are rare (1 x *dar miedo*, 1x *dar asco*, 2 x *dar rabia* in the CDE) and require clitic doubling as in ex. (4c) and (8) (cf. also Cuervo (2010: 141)). This approximates them to incorporated datives. For a discussion of datives and clitic doubling in general, and incorporated datives in particular in the framework of RRG cf. Belloro (2007: 202).

(13) a.* Miedo es lo que me da
fear be.3SG.PRS 3SG.IMPS that.REL me.CLI.DAT give:3SG.PRS

su actitud
POSS.3SG attitude

LIT: 'Fear is what his attitude gives me.' (His attitude scares me)

b. Lo que tengo es frío.
3SG.IMPS that.REL have-1SG.PRS be.3SG.PRS cold

LIT: Cold is what I am.

(Examples form Crego García (2000: 53); the translations are my own)

4 LVCs in RRG

LVCs have not received a lot of attention in RRG. Looking for reasons, we should consider both the "nature" of LVCs and the centre of interest in RRG. As to the nature of LVCs, Schultze-Berndt (2006: 372–373) observes with respect to their description in reference grammars:

> ...complex predicates formed with function verbs [i.e. light verbs, ES] have a double nature: they are syntactically complex and (more or less) productive, but at the same time, they also form (more or less) lexicalised collocations [...] As a result of their dual nature, complex predicates formed with function verbs do not easily fit into the standard outline of reference grammars with their division into "Morphology" [...] "Phrasal syntax" and "Clausal syntax". Occasionally, this seems to have the result that they are not discussed at all...

LVCs are syntactically structured, but at the same time, they can be subject to more or less specific collocational restrictions. Especially the less productive and strongly lexicalised ones cannot be described by making reference to general grammatical principles only. And this is why they are not an obvious subject for an RRG analysis. Traditionally, RRG is more concerned with what is systematic, both within individual languages and across typologically distinct languages.[22]

In the following, I will shortly discuss two different proposals for the analysis of different types of LVCs in RRG in order to examine whether existing RRG approaches are helpful for the description of the French and Spanish LVCs in question. When comparing these studies, Nolan (2014) on Irish and Romero Méndez (2007) on Spanish LVCs, it becomes again evident that "LVC" is a cover term for structurally distinct complex predicates that fulfil different functions in different languages.

[22] One of the central questions posed in RRG is: "how can the *interaction of syntax, semantics, and pragmatics* in different *grammatical systems* be best captured and explained?" (Van Valin 2005: 1, my emphasis).

The Irish LVCs analysed in Nolan (2014) are based on a combination of a LV and a verbal noun (VN). Nolan identifies 16 LVs, all of which have 'heavy' counterparts.[23] The semantic contributions of the LVs are manifold and include causation, phase information (cf. ex. (14b)) – i.e. the indication of the start, continuation or termination of an event – and sub-event modification. The LVs combine with a VN that denotes the event itself (cf. Nolan 2014: 144). Another function seems to be the formation of activity predicates as in ex. (14a).

(14) Irish LVCs

a. Ansin rinne na páistí gáire.
then do.make-PST DET-PL children laughing:VN
LIT: Then the children made laughing.
'Then the children laughed.'

b. Thosaigh siad go leír ag troid.
start-PST 3PL to all at fighting:VN
'They all started fighting.'
(Examples from Nolan 2014: 149–150)

Syntactically, these LVCs are treated as cases of nuclear cosubordination, which is the tightest possible juncture type (cf. Nolan 2014: 149). Two predicates are merged together to represent multiple facets of a situation as a single event. In nuclear junctures, all arguments are shared between the component parts, and nuclear operators (e.g. aspect, negation with narrow scope) always apply to the complex predicate as a whole. In textbook examples, the component parts of the complex nucleus are usually adjacent to one another.[24] The proximity of the component parts is taken to be a sign of their tight connection (cf. Van Valin 2005: 191, 202)

The LVCs discussed by Nolan seem to be compositional, i.e. the restrictions on the compatibility of the LVs and matrix verbs seem to be purely semantic. He does not mention any idiosyncratic restrictions typical of lexicalised combinations. The LVCs described by Nolan (2014) are thus of a different nature from the French and Spanish 'MAKE/GIVE FEAR'-type LVCs, where we cannot freely interchange the LV.

The Spanish LVCs analysed by Romero Méndez (2007) are semi-productive. The component parts are syntactically joined, but there are restrictions on the combinations of nouns and LVs. The noun *ducha* ('shower'), for example, can combine with *dar* ('give'), but not with

[23] The sources of the LVs on Nolan's list include generalised and basic action verbs (*déan* 'do/make', *tóg* 'take', *gabh* 'take', *tabhair* 'give', *chuir* 'put'), aspectual verbs (*thósaigh* 'start', *coinnigh* 'continue', *stad* 'stop'), and motion/posture and modalising verbs (*ceadaigh* 'let', *lig* 'allow', *éirigh* 'rise', *Iarraigh* 'try', *chaith* 'spend', *buail* 'hit', *téigh* 'go', *seas* 'stand') (cf. Nolan 2014: 144–145).

[24] Cf. for example *Je* [NUC [NUC *ferai*] [NUC *manger*]] *les gâteaux à Jean* ('I will make Jean eat the cake', Van Valin 2005: 191), where the causativiser *ferai* is adjacent to the matrix verb *manger* ('to eat').

hacer ('make') to yield the interpretation 'take a shower'. *Hacer una ducha* means 'make (build) a shower'.[25] The idiosyncracies are not completely arbitrary. LVCs pattern in families, and there is a lot of regularity within these families, i.e. once one understands the characteristics of a certain group of LVCs, it is possible to almost compositionally analyse the meaning of predicates (Romero Méndez 2007: 22–24).

Romero Méndez analyses LVCs as a special case of co-composition in the sense of Pustejovsky. The meaning of LVCs is related to the meaning of the component parts, but it cannot be derived completely by taking into account the meaning of the individual items. There are additional interpretive restrictions.[26] The meaning of the light verb changes as a function of the noun (cf. Romero Méndez 2007: 37–40).[27]

Given the semi-productive character of LVCs, Romero Méndez suggests a lexicon-based approach to the description of LVCs within RRG. He proposes that the LS of LVCs should be listed in the lexicon like that of any other predicate. In his approach, the semantics and LS of the light verbs is derived from the LS of the full verb. The LS of the full verb *dar* is given in ex. (15). It consists of two predicative parts – an activity predicate (**do'**) and a change of state operator plus state predicate (BECOME **have'**) – joined by a CAUSE operator.

(15) LS of the full verb *dar* ('to give'; Roméro Méndez 2007: 35)
[**do'**(x, ∅)] CAUSE [BECOME **have'** (y, z)]

According to Romero Méndez, different LVCs 'select' different parts of the LS of the full verb *dar* ('to give'). Whereas in his analysis *dar* + state noun LVCs are causative and preserve the bipartite structure of the LS of the full verb *dar* (cf. ex. (16a)), the LS of *dar* + activity NP only preserves the first part of the LS shown in ex. (16b), i.e. the activity predicate [**do'**(x,...)] (cf. Roméro Méndez 2007: 35, 38–39).

(16) a. El helado me dio frío.
DET.M.SG ice cream me.CLI.DAT give:3SG.PST cold
'The ice cream made me feel cold.'[28]
[**do'**(helado, ∅)] CAUSE [BECOME **feel'** (1SG, [**cold'**])]
(LVC: *dar* 'to give' + *frío* 'cold')

[25] The same applies to French. *Faire une douche* means 'to make (build) a shower', whereas 'take a shower' is *prendre une douche* (*prendre* = 'take').

[26] For a summary on the principle of co-compositionality cf. Pustejovsky (2012), where he explains that "co-compositonality is not the result of a failure of compositionality, and hence to be viewed as involving non-compositional processes. Rather, as the name would suggest, it entails at least conventional compositional mechanisms for the expressions involved, along with additional interpretive mechanisms not always exploited within a phrasal composition" (Pustejovsky 2012: 371).

[27] For further remarks on the semantics of Spanish as well as French LVCs cf. also section 3.1.

[28] The original translation provided by Romero Méndez (2007) is 'The ice cream got me cold'.

b. Jeon no se dio una ducha durante todo el semestre.
Jeon NEG REFL give:3SG.PST ART shower during all DET.M semester
'Jeon did not take a shower during the whole semester'
do'(x, [**shower'**(x, y)])
(LVC: *darse* 'to give onself' + *una ducha* 'a shower')
(Examples from Rómero Méndez 2007: 35, 39)

He directs the attention to the fact that the LS for *dar frío* as stated in (16a) contains the state predicate **feel'** and argues that this part of the predicate is contributed by the qualia structure of the noun *frío*. The state predicate component is absent in the LS for *darse una ducha* in ex. (16b). This is explained by the presence of *ducha*, which is analysed as an activity noun that highlights the activity component of the verb *dar*. In both examples, the LS of the LVC is made up of the LS of two lexical items. As a result of co-composition, certain facets of meaning of the component parts are highlighted to the detriment of others (cf. Roméro Méndez 2007: 30).

Romero Méndez shows that LSs are a useful means to model the results of LVC formation. Still, there are some remarks to be made about his analysis. The LS of *dar frío* and *darse una ducha* as represented in (16a) and (16b) adequately reflect that *dar frío* ('make cold') is a complex of two causatively related predications, whereas *darse una ducha* ('take a shower') is not. While the situation described by *dar frío* refers to an event/state (such as eating ice cream) that causes another event/state (feeling cold), *darse una ducha* ('take a shower') refers to a homogeneous event.

Although the representation given in (16b) represents that there are no subevents involved in the situation represented by *darse una ducha*, it is problematic from the point of view of predicate structure. It reduces the valency of *dar* to two slots in the first place by making it an activity and then introduces two further arguments by attributing two argument slots to the embedded predicate **shower'**. It is not explained, however, what makes *ducha* a two place predicate that can be represented by **shower'** (x, y) in the LS. The difficulty seems to be related to the fact that LVCs preserve the language-specific pattern of the underlying full verb *dar*, and that LSs require a representation based on universally valid predicates and operators which at the same time reflect structural properties from which macrorole (MR) assignment and argument realisation can be inferred. Finding a representation that adequately reflects all these aspects is challenging.

Romero Méndez (2007: 30) briefly addresses the issue of argument realisation and claims that it can be derived from the LS following general RRG linking principles. In the case of (16a), it is possible to derive argument realisation from the LS with some additional provisions concerning the status of the predicative noun. Contrary to this claim, however, it seems problematic to derive the syntactic realisation of *darse una ducha* from the LS given in (16b).

The LS proposed for *darse una ducha* is the same as for two place activity verbs such as *comer una pizza* ('eat a pizza'). The syntactic realisation of *darse una ducha*, however, requires a three-place reflexive construction related to *dar una ducha a alguien* ('to give a shower to somebody'). As a consequence, it is not possible to derive the syntactic construction given in (16b) by simply making reference to the LS.

With respect to the nexus juncture type, Romero Méndez (2007: 15) suggests that predicative NPs such as *frío* or *una ducha* have the status of an argument of the LV and that the link of the component parts does not correspond to a nuclear cosubordination juncture-nexus type. He justifies this by showing that predicative NPs behave like objects. They can be pronominalised and allow extraposition. This is, however, not the case for all LVCs (cf. ex. (13)), as was shown in section 3.2. *Dar miedo* type LVCs behave like nuclear junctures with respect to pronominalisation and extraction.[29]

Considering that LVC families have different syntactic behaviours and restrictions on the morphosyntactic properties of the NPRED (e.g. articles, pluralisation), a purely lexicon-based approach to their description is problematic. LSs only provide information about basic argument configurations that serve as a basis for macrorole assignment.

5 A constructional account of 'make/give fear'-type LVCs

5.1 Constructions in RRG

Although RRG is a theory that does not rely primarily on constructions as a tool for linguistic description, the wide range of linguistic phenomena that have been analysed using constructional accounts within RRG suggests that constructions are considered more and more important by linguists working within this framework.[30]

To mention but a few examples, Diedrichsen (2009) provides a constructional analysis of the different ways of passive formation in German and highlights the usefulness of constructions to explain mismatches between the arguments provided by the verb's LS and the argument structure in constructions.[31] Kailuweit (2011) proposes a constructional account for the analysis of anti-causative alternations in Romance languages. The paper by Nolan (2014) already discussed in section 4 uses constructional schemas for the description of LVCs, and Svorou (this volume) provides a constructional analysis of conjunction reduction with motion verb coordination.

[29] Their somewhat greater syntactic flexibility in Spanish as compared to French *faire peur* LVCs can be explained by a general flexibility of stress patterns and word order in Spanish.

[30] cf. Van Valin (1993: 110), where it is stated that RRG "falls between GB [Government and Binding] theory, on the one hand, which specifically denies the validity of the notion of grammatical construction (Chomsky 1988), and Fillmorean Construction Grammar (Fillmore 1988), on the other hand, in which only language-specific construction templates are posited."

[31] E.g. *It bores a hole in my head* (cf. Diedrichsen 2009: 3).

The status of constructions in RRG is described as follows by Van Valin (2005: 131–132):

> RRG recognizes the importance of grammatical constructions, and they are represented in terms of constructional schemas. Cross-constructional and cross-linguistic generalizations are captured in terms of the general principles and constraints that constitute the linking algorithms, e.g. the actor-undergoer hierarchy, the layered structure of the clause, the privileged syntactic argument selection hierarchy. Only the language-specific features are represented in constructional schemas. Hence constructional schemas, by virtue of their reference to general principles, permit the capturing of cross-linguistic generalizations, while at the same time expressing language-particular properties of grammars.

In RRG, it is assumed that linguistic description should rely on general principles and constraints where possible. It is admitted, however, that there are language-specific phenomena that can best be captured by constructional schemas. Nevertheless, it is emphasised that constructional schemas should make reference to general principles, so that cross-linguistic comparisons and generalisations are possible. The examples for constructional schemas given in Van Valin (2005: 132–133), such as WH-question formation and passive formation, show that the notion of construction in RRG is different from the notion of construction advised in Construction Grammar approaches. The examples for constructional schemas given by Van Valin (*ibid.*) contain information about the syntactic template to be used (e.g. English WH-question $\rightarrow$ PrCS-template), the morphological properties associated with the construction and pragmatic information concerning illocutionary force and focus structure. Constructional schemas also include some semantic information, which is, however, very general. In the case of the English passive construction, for example, it is specified that the "PSA is not the instigator of state of affairs, but is affected by it (default)" (*ibid.* 132). Constructional templates in RRG do not include lexical information. As Butler (2003: 132) points out, this is a major difference compared to Construction Grammar

> ...whereas Construction Grammar postulates that all information, including lexical, is integrated into the construction, RRG prefers to have two separate but interactive components: a syntactic inventory containing the syntactic templates and the lexicon containing all kinds of lexical entries.

Separating grammatical from lexical information is useful if one aims at getting to cross-linguistic generalisations.

5.2 The constructional schema for 'make/give fear'-type LVCs

The aim of the present section is to provide a description of the 'MAKE/GIVE FEAR'-type LVC family in French and Spanish using constructional schemas as a descriptive tool. In order to

develop the constructional schema, major characteristics of Spanish *dar miedo* and French *faire peur* type LVCs will be revised.

LVCs tend to be organised in 'families', as illustrated in (17). The curly brackets in the Spanish examples in 18 indicate that overt realisation of the argument is optional. The examples show utterances based on three different instances of the *faire mal* and *dar miedo* LVC type.[32]

(17) a. [Son regard / Ça] me fait peur.
POSS gaze / DEM.PRO me.CLI.DAT make:3SG fear.N
'[His gaze / this] scares me.'

b. [Ce succès / Ça] me fait plaisir.
DEM success / DEM.PRO me.CLI.DAT make:3SG pleasure.N
'[This success / this] makes me happy.'

c. [Cette défaite/ Ça] me fait mal.
DEM defeat / DEM.PRO make:3SG painN
'[This defeat / this] hurts me.'

(18) a. Me da miedo {la muerte}.
me.CLI.DAT give:3SG fear.N {ART.F death}
'[Death / it] scares me.'

b. Me da rabia {la ignorancia de la gente}.
me.CLI.DAT give:3SG rage.N {ART.F ignorance of ART.F people}
'[People's ignorance / it] makes me furious.'

c. Me da asco {la corrupción}.
me.CLI.DAT give:3SG disgust.N {ART.F corruption}
'[Corruption / it] disgusts me.'

The LVCs occur with the same syntactic patterns, and with the same argument structure. There is a first argument slot that allows for animate and state of affairs nouns (which are more frequent), and there is a third argument that is tightly linked with the predicative noun. The predicative noun is usually a bare noun and appears in the position of the second argument. It expresses an emotional state that is attributed to the referent of the third argument. [33]

[32] Note that for the definition of LVC families, LVCs should coincide in basic syntactic patterns. However, there will most certainly be different preferences for certain construction types between the members of the LVC family.

[33] Note that this is a major difference between LVCs of the 'MAKE/GIVE FEAR'-type and activity LVCs such as *faire signe* (lit. 'to do sign'), which at first sight seem to be based on a comparable pattern (cf. ex. (i)). In the case of *faire signe*, the predicative noun specifies the action carried out by the referent of the first argument.

i. Le client lui fait signe.
DET.M customer him/her.CLI.DAT make:3SG.PRS sign
'The customer beckons him/her over.'

CONSTRUCTION	Spanish 3-place DAT-EXP LVC construction
SYNTAX	Juncture: nuclear Nexus: cosubordination Construction type: 3-place LVC [CL [CORE RP_1 [[NUC_1 ...] [NUC_2 ...]] RP_2 ...] ...] ([NUC_1=V, [NUC_2=N) RP_2 is a non-macrorole core argument Linking: 1 MR-argument (derivable from LS and general linking principles) PSA: MR-argument
MORPHOLOGY	predicative noun is a bare noun, unless modified by an adjective (requires indefinite article) non-MR core argument receives dative case (according to general linking principles)
SEMANTICS	**pred'**$_{NUC2}$ denotes an emotional state that can be triggered by external influence LS: **pred'**$_{NUC1}$(x) CAUSE **feel'**(y, **pred'**$_{NUC2}$)
PRAGMATICS	Focus structure: PSA=topic (default) Illocutionary force: unspecified

Fig. 1: Constructional schema for Spanish *dar miedo* type LVC construction

In RRG, the number of arguments associated with a predicate is coded in the LS, which is specified in the lexicon. The LS provides a link between lexical and syntactic information. In the case of LVCs, the LS should not only figure in the lexicon, but also be part of the constructional schema associated with a particular LVC, as illustrated in figures 1 and 2.

LS are the basis for argument realisation, which is derived from them. This is also the case for 'MAKE/GIVE FEAR'-type LVCs. Although there are morphological and syntactic constraints and preferences, they reflect general principles of French and Spanish grammar. The LVCs in question form part of the grammatical system of the respective languages. This aspect is also reflected in the constructional description.

'MAKE/GIVE FEAR'-type LVCs are three-place constructions with two arguments. With respect to argument realisation, however, the LVCs in question behave very much like free combinations, except that the position of the second argument (that could potentially be assigned the second MR) is blocked by the predicative noun. On that condition, the LVCs can realise one MR according to general linking principles. This MR is assigned to the COR argument, which is the leftmost argument.

CONSTRUCTION	French 3-place DAT-EXP LVC construction
SYNTAX	Juncture: nuclear Nexus: cosubordination Construction type: 3-place LVC [CL [CORE RP_1 [[NUC_1 ...] [NUC_2 ...]] RP_2 ...] ...] ([NUC_1=V, [NUC_2=N) RP_2 is a non-macrorole core argument Linking: 1 MR-argument (derivable from LS and general linking principles) PSA: MR-argument
MORPHOLOGY	predicative noun is a bare noun, unless modified by an adjective (requires indefinite article) non-MR core argument receives dative case (according to general linking principles)
SEMANTICS	**pred'**$_{NUC2}$ denotes an emotional state that can be triggered by external influence LS: **do'**$_{NUC1}$(x, ∅) CAUSE **feel'**(y, **pred'**$_{NUC2}$) or LS: **pred'**$_{NUC1}$(x) CAUSE **feel'**(y, **pred'**$_{NUC2}$)
PRAGMATICS	Focus structure: PSA=topic (default) Illocutionary force: unspecified

Fig. 2: Constructional schema for French *faire peur* type LVC construction

The LS associated with the respective French and Spanish LVC only partially coincide. In French, animate subjects can be interpreted as agents, whereas Spanish clearly favors non-agentive readings. This is reflected by the presence of an LS containing **do'** for French *faire peur* LVCs.[34] In French, the COR argument can be an actor or an undergoer, depending on the context. In Spanish, the COR-argument is an argument of **pred'**, i.e. it is assigned the undergoer MR. This is reflected in the syntactic behaviour of the COR-argument, which follows the pattern of arguments of unaccusative verbs.

The status of the arguments is marked by case and word order (French).[35] The MR-argument receives nominative case and is the PSA, which agrees with the LV in person and number. The position of the second argument is not available due to the presence of the predicative noun.

[34] One could argue that the LS with **pred'** is not necessary in order to explain the behaviour of the French LVC, because non-agentive subjects and agentive subjects behave the same way.

[35] Case marking for nominative and accusative, however, is restricted to personal pronouns in French and Spanish. Lexical NPs and demonstrative pronouns are not overtly marked for nominative or accusative case. Dative case is marked by a preposition with lexical NPs or by the paradigm of indirect object pronouns.

The third argument is realised as a non-MR core argument and receives dative case. With respect to the form of argument realisation, the LVCs behave very much like simplex verbs despite the fact that there is a bare noun, which is part of the predicate nucleus, in the position where otherwise direct objects occur.

When specifying the properties of LVCs, we can take advantage of the fact that LVCs reflect the properties of the grammatical system they belong to. In the constructional schema we can both specify where the LVCs follow general syntactic principles and which constraints apply to them. We can specify the nexus-juncture type and the morphosyntactic properties of the NPRED.

The syntactic properties specified in the constructional schema account for the fact that LVCs are syntactically less flexible than free combinations. Extraction of the predicative noun, for example, is not possible because the LV and the NPRED are connected by a nuclear juncture. Syntactic inflexibility is also reflected by the property of MR intransitivity. Thus, syntactic operations that require two MRs, such as passive formation, are excluded.

In comparison to the constructional schemas proposed in Van Valin (2005) that handle phenomena such as WH-questions, causative constructions or reflexivisation, the constructional schema for the LVCs in question have to be much more specific. They merge lexical information such as the LS of the predicates they apply to and the semantic class of predicative elements admitted in this construction with morphological, syntactic and pragmatic information. They represent both the regular and the idiosyncratic characteristics of the LVC construction. A systematic description of LVCs by means of constructional schemas can reveal how LVCs are both shaped by general grammatical principles of the language they belong to and in which respects they are idiosyncratic constructions. This could also be a basis for a more systematic description of LVC types both within one language and cross-linguistically.

6 Conclusion

This paper has looked at the characteristics of lexicalised semi-compositional LVCs in French and has discussed possible ways to describe them in the framework of RRG. It was shown that the LVCs analysed are syntactically more restricted than free combinations, but that their syntax is nevertheless visibly shaped by the grammatical system they belong to. This was illustrated by making reference to the differences between the syntactic patterns found with French *faire mal* and Spanish *dar miedo* type LVCs, such as the preference of the postverbal realisation of the COR argument in Spanish.

It was argued that, although RRG aims at separating lexicon and grammar, this is not entirely possible with LVCs. There has to be a way to connect the lexical entries for LVCs with the

constructional schemas that specify their structural properties and constraints. The possibility to highlight that LVCs rely on general patterns of the grammatical systems they pertain to is seen as a major advantage of the description of LVCs in RRG constructional schemas. In order to substantiate this claim, further analyses of other LVC types are required.

An open issue is the representation of the predicate structure and meaning in the LS. Given that LVCs exploit language-specific patterns and language-specific semantic ranges of LVs, the advantages and disadvantages of using a universal semantic metalanguage for representing the LS of LVCs should be investigated in future work.

References

Alba-Salas, Josep. 2007. On the life and death of a collocation. A corpus-based diachronic study of *dar miedo/hacer miedo*-type structures in Spanish. *Diachronica* 24(2): 207–252.

Alonso Ramos, Margarita. 2004. *Las construcciones con verbo de apoyo.* Madrid: Visor.

ATILF, CNRS & Université de Lorraine. Base textuelle FRANTEXT.
Available online: http://www.frantext.fr

Belloro, Valeria. 2007. *Spanish Clitic Doubling: A Study of the Syntax-Pragmatics Interface.* Phd Thesis, State University of New York at Buffalo. [Available on RRG website]

Brinton, Laurel J. 2011. The grammaticalization of complex predicates. In *The Oxford Handbook of Grammaticalization*, B. Heine & H. Narrog (eds), 559–569. Oxford: OUP.

— & Traugott, Elizabeth. 2005. *Lexicalization and language change*. Cambridge: CUP.

Butler, Christopher S. 2003. *Structure and Function: A Guide to the Three Major Structural-Functional Theories, Part 1: Approaches to the simplex clause.* Amsterdam: John Benjamins.

Butt, Miriam. 2010. The light verb jungle: Still hacking away. In *Complex predicates: cross-linguistic perspectives on event structure*, M. Amberber, B. Baker & M. Harvey (eds), 48–78. Cambridge: CUP.

Crego García, María Victorina. 2000. En torno a las formas descompuestas. *Estudios de lingüística* 14: 45–74.

Cuervo, María Cristina. 2010. Two types of (apparently) ditransitive light verb constructions. In *Romance Linguistics 2008: Interactions in Romance*, K. Arregi, Z. Fagyal, S. A. Montrul & A. Tremblay (eds), 139–156. Amsterdam: John Benjamins.

Danlos, Laurence. 1992. Support verb constructions: linguistic properties, representation, translation. *French Language Studies* 2: 1–32.

Davies, Mark. 2002–. Corpus del Español: 1200s–1900s.
Available online: http://www.corpusdelespanol.org

Detges, Ulrich (1996). *Nominalprädikate: eine valenztheoretische Untersuchung*. Tübingen: Niemeyer.

Diedrichsen, Elke (2009). Towards a reconsideration of Constructional Schemas in RRG: Are all constructions driven by "constructions"? [Available on RRG website]

Family, Neiloufar. 2014. *Semantic Spaces of Persian Light Verbs: A Constructionist Account.* Leiden, NL: Brill.

Giry-Schneider, Jacqueline. 1987. *Les prédicats nominaux en français*. Geneva: Librairie Droz.

Gledhill, Christopher. 2008. Portée, Pivot, Paradigme: trois termes pour faire le point sur les expressions verbo-nominales. *Zeitschrift für Französische Sprache und Literatur, Beiheft* 35, P. Frath (ed.), 59–76. Stuttgart: Franz Steiner Verlag.

Gross, Maurice. 1981. Les bases empiriques de la notion de prédicat sémantique. *Langages* 63: 7–52.

Jouitteau, Mélanie & Rezac, Milan. 2007. The French ethical dative, 13 syntactic tests. In *Bucharest Working Papers in Linguistics*, IX(1): 97–108.

Kailuweit, Rolf. 2005. *Linking. Syntax und Semantik französischer und italienischer Gefühlsverben*. Tübingen: Niemeyer.

— 2011. Romance Anticausatives: A Constructionist RRG Approach, In *New Perspectives in Role and Reference Grammar*, W. Nakamura (ed.), 104–133. Newcastle upon Tyne: Cambridge Scholars Publishing.

Koch, Peter. 1981. *Verb, Valenz, Verfügung. Zur Satzsemantik und Valenz französischer Verben am Beispiel der Verfügungsverben.* Heidelberg: Winter.

Lauwers, Peter & Toback, Els. 2010. Les verbes attributifs: inventaire(s) et statut(s). *Langages* 179–180: 79–113.

Marchello-Nizia, Christiane. 1996. Les verbes supports en diachronie: le cas du français. *Langages* 30(121): 91–98.

Mayoral Hernández, Roberto. 2007. A variation study of verb types and subject position: verbs of light and sound emission. In *Romance linguistics 2006: selected papers from the 36th Linguistic Symposium on Romance Languages (*LSRL*), New Brunswick, March 31–April 2, 2006*, J. Camacho *et al.* (eds), 219–232. Amsterdam: John Benjamins.

Nolan, Brian. 2014. Complex predicates and light verb constructions in Modern Irish. *Revista Española de Lingüística Aplicada* 27(1): 140–167.

Pustejovsky, James. 2012. Co-compositionality in grammar. In *The Oxford Handbook of Compositionality*, W. Hinzen, E. Machery & M. Werning (eds), 371–382. Oxford: OUP.

Rohrer, Christian. 1967. Definition of locutions verbales. *The French Review* 41(3): 357–367.

Romero Méndez, Rodrigo. 2007. Spanish Light Verb Constructions: co-predication with syntactically formed complex predicates. [Available on RRG website]

Schøsler, Lene. 2008. Étude sur l'évolution des constructions à verbe support. In *Évolutions en français: études de linguistique diachronique*, B. Fagard (ed.), 345–362. Frankfurt am Main: Lang.

Schultze-Berndt, Eva. 2006. Taking a closer look at function verbs: lexicon, grammar, or both?. In *Catching Language: The Standing Challenge of Grammar Writing*, F. K. Ameka, A. Dench & N. Evans (eds), 359–391. Berlin: De Gruyter.

Staudinger, Eva (2013). *Syntax, Subjektivität, Subjektivierung. Schmerzprädikate im Französischen.* Freiburg: Rombach.

Ströbel, Liane. 2010. *Auf den Spuren der Entstehung einer neuen Kategorie: Leerverben als paralleler Kopulastrang.* Frankfurt am Main: Lang.

Vivès, Robert. 1993. La prédication nominale et l'analyse par verbes supports. *L'information grammaticale* 59(1): 8–15.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

—. 1993. A synopsis of Role and Reference Grammar. In *Advances in Role and Reference Grammar* R. D. Van Valin Jr. (ed.), 1–164. Amsterdam: John Benjamins.

—. 1990. Semantic Parameters of Split Intransitivity. *Language* 66: 221–260.

Valli, André. 2007. À propos de la notion de locution verbale: Examen de quelques constructions à verbe support en moyen français. *Langue française* 156: 45–60.

Winhart, Heike. 2005. *Funktionsverbgefüge im Deutschen. Zur Verbindung von Verben und Nominalisierungen.* PhD Thesis, University of Tübingen.

Wittenberg, Eva *et al.* 2014. The mental representation and processing of light verbs. In *Structuring the Argument. Multidisciplinary research on verb argument structure*, A. Bachrach, I. Roy & L. Stockall (eds), 61–80. Amsterdam: John Benjamins.

Yaqui Adverbial Clauses and the Interclausal Relations Hierarchy

Lilián Guerrero
IIFL – Universidad Nacional Autónoma de México

Abstract

In Yaqui, purpose, reason and temporal relations are expressed by multiple syntactic structures. The aim of this paper is to explore whether Yaqui syntactic linkages are compatible (or not) with the cross-linguistic predictions made by the Interclausal semantic relations hierarchy (Van Valin 2005: 209). This hierarchy predicts a scale where purpose outranks reason and reason outranks temporally ordered events. However, Yaqui does not conform to the expected rankings. In this language, purpose relations make use of the tightest syntactic linkages, while reason relations are expressed by the loosest syntactic structure. However, in terms of syntactic binding and semantic integration, simultaneous and sequential temporal expressions are placed right in the middle of purpose and reason.

1 Introduction[1]

The iconic correlation between the semantic and syntactic representations of complex sentences is very well known within the studies of clause union, especially in the domain of complementation (Silverstein 1976; Givón 1980; Haiman 1985; Noonan 2007 [1985]). The same syntactic devices that signal stronger interclausal dependency in complementation reflect also the stronger dependence of adverbial subordination but to a lesser degree (Givón 2001; Cristofaro 2003). As one step further, the theory of clause union in Role and Reference Grammar formalizes an implicational hierarchy which links the semantic and syntactic dimensions of complex constructions (Foley & Van Valin 1984; Bickel 1993; Van Valin & LaPolla 1997; Van Valin 2005, 2007). The Interclausal relations hierarchy predicts that the closer the semantic relationship between the two events is, the stronger the syntactic link

[1] This paper has been partially supported by the research project "Adverbial relations in Yaqui and other languages" (PAPIIT IN401816).

joining them will be. With respect to adverbials, the semantic relations can be ranked as in (1).

(1) Interclausal adverbial semantic hierarchy (adapted from Van Valin 2005: 207)
Modifying sub-event [manner, motion, position, means] > purposive > temporal or spatial circumstances > reason > conditional > concessive > simultaneous and sequential temporal > unordered temporal states of affairs

On the syntactic side of the Interclausal relations hierarchy, the linkage types are ranked in terms of their syntactic tightness. At the bottom of the scale are combinations of whole clauses constituting sentences (e.g. sentential or clausal coordination). As one goes up the hierarchy, the linked units lose more and more features of independent clauses until they are reduced to a bare nucleus or predicate (e.g. nuclear co-subordination). At the middle portion of the scale, in a non-subordinate core juncture, the two cores obligatorily share one core argument; in a subordinate core juncture, the linked core serves as a syntactic argument of the matrix core. These juncture-nexus types are abstract linkage relations, not grammatical construction types, meaning each linkage type may be realized by more than one grammatical construction in a language, and vice versa, the same grammatical construction may involve different linkage types. The crucial point here is that the looser the semantic relations, the more marked a tight syntactic linkage should be.

As part of a major project dealing with the syntax, semantics and pragmatics of clause union in Yaqui (Uto-Aztecan, Mexico), this paper focuses on adverbial constructions expressing purpose, reason and temporal relations. The Interclausal semantic hierarchy defines these relations as follows:

(2) Interclausal semantic relations [partial] (Van Valin 2005: 206–207):

e. Purposive: one action is done with the intent of realizing another state of affairs, e.g. *John went to the store to buy milk*, *Susan brought the book to read*

n. Circumstances: the spatial or temporal parameters of an event, e.g. *Kim saw Pat after she arrived at the party*

o. Reason: the motivation or cause for an action or event, e.g. *The baby cried because she was hungry*

r. Temporal

1. Simultaneous states of affairs: one state of affairs is temporally coterminous with another, e.g. *Max was dancing, and at the same time Susan played the piano.*

2. Sequential states of affairs: one state of affairs follows another temporally, with or without any temporal overlap, e.g. *Juan finished reading the newspaper, and then Carlos walked into the room.*

s. Temporally unordered states of affairs: the temporal relation between states of affairs is unexpressed, e.g. *John talked to Tom, and Sue chatted with Kim*

Yaqui displays considerable structural variation across these adverbial sentences. The aim of this paper is to explore whether Yaqui purpose, reason and temporal clauses are compatible with the cross-linguistic predictions made by the semantic relations hierarchy in (1). In this paper, I show that Yaqui purpose relations make use of the tightest syntactic linkages, while reason relations are expressed by the loosest structure. That is, the structures expressing purpose and reason relations reflect in an iconic way the semantic association between the two states of affairs. However, the units in temporal clauses are syntactically tighter and semantically closer than those involved in reason clauses. In fact, Yaqui makes a good case for re-examining the ranking of the Interclausal semantic hierarchy for a scale like purpose > simultaneous, sequential > reason.

2 Basic morpho-syntactic features of Yaqui

Yaqui is a synthetic/agglutinating, head-final, dependent marking, and primary object language (Lindenfeld 1973; Escalante 1990; Dedrick & Casad 1999; Guerrero & Van Valin 2004). In nominal arguments, the nominative case is unmarked (2a–b) and the accusative is marked by *-ta* (2b). Oblique core arguments are introduced by postpositions (2c). The accusative *-ta* and the plural *-(i)m* are mutually exclusive in nominals and determiners (2c).

(2) a. U-∅ o'ou-∅ batwe-u yeu=siika
DET-NOM man-NOM river-DIR out=go.SG.PFV
'The man went to the river.'

b. U-∅ o'ou-∅ u-ka toto'i-ta bicha-k
DET-NOM man-NOM DET-ACC hen-ACC see-PFV
'The man saw the woman.'

c. U-me toto'i-m=ne u-e jamut-ta-u nenki-ne
DET-PL hen-PL=1SG.NOM DET-OBL woman-ACC-DIR sell-POT
'I will sell the hens to the woman.'

Yaqui shows a rigid SOV word order, though other orders are possible for specific pragmatic functions. There are two sets of pronouns (Table 1). Full pronouns behave as lexical elements

in terms of their distribution, while reduced nominative pronouns behave as second position clitics, and reduced accusatives (available only for 3rd person) cliticize to the verb. There is some degree of morphological syncretism in the pronominal system; for instance, the reduced pronoun *ne* '1.SG' can serve as a nominative or accusative argument.

	nominative	accusative	oblique	genitive
1 SG	*ne*	*nee*, *ne*	*ne-*	*in*, *nim*
2 SG	*empo* = *'e*	*enchi*	*e-*	*em*
3 SG	*aapo* = ∅	*apo'ik*, *a'a*, *a*= ∅	*a-*	*apo'ik*, *a*=
1 PL	*itepo* = *te*	*itom*	*ito-*	*itom*
2 PL	*eme'e* = *'em*	*enchim*	*emo-*	*em*, *enchim*
3 PL	*bempo* = ∅	*apo'im*/*am*	*ame-*	*bem*, *bempo'im*

Table 1: Yaqui pronominal system

Except for a few suppletive forms, verbs do not express person or number. As typical in the family, there is little indication of suffixes that mark pure tense. Instead, the usual situation for the tense-aspect-mood affixes is to display a range of meanings (Table 2), including tense-aspectual values and distinct mood/epistemic states. The present tense is unmarked, meaning a simple clause (3a) in present tense and an unmarked verb in a complex predicate (3b) look the same. The difference is that the former can take any TAM affix.

V-∅	Present	RDP-V	Habitual
V-*k(a)*	Past perfective	RDP.RDP-V	Iterative, intense
V-*n*	Past continuative	V-*su*	Completive
V-*ne*	Future, potential	V-*'ean*	Subjunctive; hypothetical obligation
V-*na*	Future, potential passive	V-*maachi*	Obligation

Table 2: Tense-Aspect-Mood (TAM) verbal system

(3) a. Ketgo jo'a-wa-u=ne yepsa-∅ / yepsa-k
early house-GEN-DIR=1SG.NOM arrive.SG-PRS / arrive.SG-PFV
'I arrive / arrived to his place earlier.'

b. Ketgo jo'a-wa-u=ne yepsa-bae-k / *yepsa-ne-bae-k
early house-GEN-DIR=1SG.NOM arrive.SG-want-PFV / arrive.SG-POT-want-PFV
'I wanted to arrive to his place earlier.'

3 Adverbial constructions, some generalities

Previous work on Yaqui has described some basic features of adverbial subordination based on the selection of the subordinator or clause-linkage marker (henceforth CLM). There are two major CLMs in Yaqui showing a consistent, though not obligatory distribution: the marker *-kai* (*-ka* when the clause is non-final) is used when the main and linked subjects are the same, as in (4a), while *-o* tends to occur when they are different, as in (4b).

(4) a. Into=ne$_i$ ousi tomi-yo'o-k [_$_i$ yeewe-kai]
DM=1SG.NOM a.lot money-win-PFV play-CLM
'And then, I won a lot of money by playing.' (Buitimea; chapayeca: 83)

b. [Unison-po=nee estudiaroa-k-o] nim papa yo'owe-∅ muuku-k
Unison-LOC=1SG.ACC study-PFV-CLM 1SG.GEN father old-NOM die.SG-PFV
'When I studied at the Unison, my grandfather died.'

These adverbial markers are highly ambiguous. For instance, *-kai* can introduce a manner clause (4a), and a 'without' clause (5a); *-o* can mark a sequential temporal clause (4b), as well as conditional (5b) and counter-factive (5c) adverbial expressions.

(5) a. [kat]=e'e$_i$ [_$_i$ ji'i-bwa-ka] to'o-ne
NEG.IMP=2SG.NOM something-eat-CLM lie-POT
'You will go to bed without eating.' (Dedrick & Casad 1999: 394)

b. [jiba enchi nee$_i$ suaati-bicha-o] ne$_i$ ya'ura-ta tejwaa-ne
always 2SG.ACC 1SG.ACC death-see-CLM 1SG.NOM authority-ACC tell-POT
'If you keep bothering me, I'll talk to the authorities.'
(Silva *et al.*; turtle.coyote: 59)

c. [Kaa nee$_i$ 'u'utte bwite-o] nee$_i$ bwij-'ea-n
NEG 1SG.ACC strong run.SG-CLM 1SG.ACC grab-SHOULD-PSTC
'If I hadn't run fast, he would have grabbed me.' (Dedrick & Casad 1999: 395)

Although uncommon, some postpositions can introduce adverbial clauses too. For instance, the directional *-u*/*-wi* (6a) can mark an adverbial locative clause. Concessive, cause and reason clauses are exceptional since they are introduced by initial free particles. An example of *ella'apo* 'although' can be seen in (6b).

(6) a. 'aman jo'a-k [bwia-ta bweji-wa'a-wi]
there home-PFV dirt-ACC dig-PASS-DIR
'He lives there where they are digging dirt.' (Dedrick & Casad: 387)

b. Itepo aman kat-ne [ella'apo ili usi-∅ ko'okwe-∅]
1PL.NOM there go-POT CLM little child-NOM sick-PRS
'We will go there although the child is sick.' (Lindenfeld 1973: 85)

The study of adverbial sentences in Yaqui is based on a corpus of approximately 4,700 clauses comprised of several narratives including traditional and folk stories, life stories and short conversations. Table 3 shows the sample of adverbial clauses.

	adverbial-main	main-adverbial	total
Reason/causal	4	185	189
Conditional	5	3	8
Temporal	148	24	172
Purpose	35	185	220
Manner-like	33	21	54
Other relations	12	5	17
	237	423	660

Table 3: Yaqui adverbial clauses in the sample

Diessel (2005, 2008) has found some interesting facts regarding the ordering of the main and the adverbial units. He found that languages of the world use either, adverbial clauses both before and after the main clause, or adverbial clauses preceding the main clause; languages preferring final adverbial units are less common in his sample. When both orders are possible in the same language, the position of the adverbial unit can be motivated by several factors. One of these factors is word order. For instance, initial adverbial clauses are common in OV languages. Yaqui shows a rigid SOV order but, interestingly, 64% of the adverbial clauses in my sample are final (Table 3). Diessel also suggests that adverbial clauses with final markers tend to be sentence initial, while clauses taking initial subordinators may alternate their position. As for Yaqui, adverbial clauses marked by an initial subordinator prefer to be final (187 out of 191 clauses), but there is no clear preference for clauses taking final CLMs, as can be seen in Table 4. Indeed, Yaqui shows a *mixed order* pattern based on the semantic relation of the two units: temporal clauses tend to be sentence initial, while purpose and reason clauses tend to occur at the end of the sentence. This can be seen in Table 3 above.

	adverbial-main	main-adverbial	total
Final subordinator	233	236	469
Initial subordinator	4	187	191
	237 (36%)	423 (64%)	660 (100%)

Table 4: Distribution between main and adverbial units in Yaqui in corpus

4 Purpose, reason and temporal clause linkages in Yaqui

Purpose relations link two events, such as the main event is performed with the goal of obtaining the realization of another event (the purposive event). [2] Yaqui has four syntactic structures expressing the notion of purpose. The clause in (7a) consists of a V_1-V_2 structure, in which the V_2 is a motion verb. In (7b–c), the linked unit is marked by the adverbial CLM *-kai*; in addition to the main verb, there is a V_1-V_2 structure. In (7d), the postposition *-betchi'ibo* '(in order) to' serves as the linkage marker.

(7) a. Peo-$\emptyset_i$ a'abo $\__i$ ji'i-bwa-se-k
Peo-NOM here something-eat-move.PURP.SG-PFV
'Peter came here to eat.'

b. Peo-$\emptyset_i$ yepsa-k [$\__i$ ji'i-bwa-se-kai]
Peo-NOM arrive.SG-PFV something-eat-move.PURP.SG-CLM
'Peter arrived to eat (=moving to eat).'

c. Peo-$\emptyset_i$ yepsa-k [$\__i$ ji'i-bwa-bae-kai]
Peo-NOM arrive.SG-PFV something-eat-want-CLM
'Peter arrived to eat (=wanting to eat).'

d. Peo-$\emptyset$ enchi$_i$ a'abo bittua-k [enchi$_i$ ji'i-bwa-ne-betchi'ibo]
Peo-NOM 2SG.ACC here send-PFV 2SG.ACC something-eat-POT-CLM
'Peter sent you here in order for you to eat.'

Reason clauses provide the cause, reason or motivation for the proposition expressed in the main unit. There are three structures in Yaqui expressing causal and reason relations: a structure takes the linkage marker *-betchi'ibo* (8a); another structure takes the initial particle *bweituk* 'because' (8b), and another one starts with *kiali'ikun* 'so that' (8c). Out of 189 reason clauses in the sample, 20 clauses (expressing a cause-effect relation) take *-betchi'ibo* as the subordinator, meaning that *bweituk* and *kiali'ikun* clauses are by far the most common strategies to express reason relations. Because clauses like in (8a) are infrequent in the sample, the analysis focuses on the two most common structures.

(8) a. Inepo in jaboi-ta$_i$ uttia-$\emptyset$ [si jita a$_i$
1SG.NOM 1SG.GEN grandpa-ACC admire-PRS INT thing 3SG.GEN
ta'a-betchi'ibo]
know-CLM
'I admire my grandpa because he knows a lot of things.'

b. [Bweituk Peo-$\emptyset_i$ lauti yebij-pea-n] 'utte'a $\__{i/j}$ weye-$\emptyset$
CLM Peo-NOM early arrive.SG-desire-PASTC strong go.SG-PRS
'Since Pedro$_i$ desired to get earlier, he$_{i/j}$ goes very fast.'

[2] For an extended analysis on the way these adverbial relations are expressed in Yaqui see Guerrero (2017).

c. Peo-∅ jaibu yepsa-k [kiali'ikun empo aa yeu=siime-∅]
Peo-NOM already arrive.SG-PFV CLM 2SG.NOM can out=go.SG-PRS
'Pedro already arrived so that you can leave.'

Temporal clauses involve two states of affairs occurring in a sequence or overlapping. Typological studies usually distinguish between posterior (*before*-relations) and anterior (*after*-relations), and consider that *when*-clauses generally imply that the main and the linked events overlap in their realization (Cristofaro 2003: 159). Yaqui displays considerable structural variation across temporal clauses too (Guerrero 2015). In opposition to English where sequential clauses are marked by predicate prepositions (*before*/*after*), the most common and frequent clauses expressing simultaneous and sequential states of affairs are solely marked by the adverbial CLMs *-o*/*-kai*. These linkage markers merely indicate that the two units are within an adverbial relation, but they do not specify the temporal relationship of the clause. In (9a), the event described in the *-o* clause is located at a time point posterior to that of the main unit, e.g. *I will jump and then, he will arrive*. In (9b), the event in the *-kai* clause takes place before the main event, e.g. the fox first took the harp and then he started playing it.

(9) a. [junumpo bea aman eela a weye-o] bea nepo
over.there DM there almost 3SG.ACC walk.SG-CLM DM 1SG.NOM

yeu=tubukti-ne
out=jump-POT

'Before/when he [the coyote] almost arrives there, I will jump.'
(Silva *et al.*; turtle & coyote: 38)

b. Aapa-∅$_i$ jika-t cha'aka, [junak a$_i$ kom=wike-ka] bea
harp-NOM up-LOC begin.hang then 3SG-ACC down=take-CLM DM

jipon-taite-k
play-start-PFV

'As for the harp that was hanged up, after/when he [the fox] took it, he started playing it.' (Silva *et al.*; fox: 38)

The examples in (10), also marked by *-o* and *-kai*, favor simultaneous rather than sequentially ordered states of affairs:

(10) a. [Amureo-m ama ane-o] ama bea tukaa-po yeu am
hunter-PL there be.around-CLM there DM night-LOC out 3PL.ACC

nuk-sim-ne
take-go.SG-POT

'While the hunters are around, at the night he would keep them [the animals].'
(Silva *et al.*; Saint: 6)

b. [Junama'a bea=ne pu'akta-wa-ka] ejtacion-eu-bicha
later DM=1SG.ACC carry-PASS-CLM station-DIR-toward

nuk-saka-wa-k
take-go.PL-PASS-PFV

'Later, I was taken to the army station while I was being carried.'

(Silva; HVH: 183)

In addition to these general clauses, Yaqui has a way to specify the temporal relation between the two events. Thus, *before*-clauses may take the adverbial particle *(ketun) kee* 'not yet', as in (11a); *after*-clauses can take the completive suffix *-su* (historically related to *ansu* 'finish') in the linked verb, as shown in (11b).

(11) a. Enchi$_j$=ne$_i$ bicha-k [ketun_kee _$_i$ kaba'i-ta jinu-kai]
2SG.ACC=1SG.NOM see-PFV CLM horse-ACC buy-CLM
'I saw you before I bought the car.'

b. Naa=bea=te pakun-bicha yeu=saja-k
then=DM=SG.NOM out-toward out=go.PL-PFV

[a=puntaroa-su-k-o]
3SG.ACC=suture-CMP-PFV-CLM

'Then, we left [the hospital] after they sutured him.' (Guerrero; Lalo: 358)

However, out of 220 temporal clauses in the sample, only 18 correspond to specific temporal clauses. In other words, the most common and frequent sentences expressing temporal relations are introduced by the general markers *-o*/*-kai*. In what follows, the major syntactic features characterizing the most common and frequent purpose, reason and temporal clauses in Yaqui are described.

4.1 The selection of the subordinator

In Yaqui, complements, relatives and several adverbial clauses all take final clause linkage markers. Purpose clauses and temporal clauses both take final and bond subordinators, but reason clauses are exceptional because they can take a final subordinator, the postposition *-betchi'ibo* in (8a), as well as initial particles. However, reason clauses taking initial clause linkers are by far the most common strategy in Yaqui. This is the first major difference among purpose, reason and temporal clauses in Yaqui.

4.2 Identity and coding of the subject

The occurrence of non-nominative subjects is a strong feature for clausal dependency in Yaqui, i.e. complement clauses demand accusative subjects, and relative clauses require ac-

cusative subjects, when nominal, or genitive subjects, when pronominal. When the main and the dependent subjects are different, purpose and temporal clauses also require a non-nominative dependent subject. The purpose clause in (12a) takes the accusative pronoun *am* 'they', and the temporal clause in (12b) takes an accusative NP *juka'a go'ita* 'the coyote'.

(12) a. Itepo i-me'e yoeme-m$_i$ aman bittua-kan [am$_i$ yumjoe-ne-betchi'ibo]
1PL.NOM DEM-PL man-PL there send-PASTC 3PL.ACC rest-POT-CLM
'We sent these men there in order for them to rest.' (Silva; HVH: 69)

b. [junak ju-ka'a go'i-ta yeu=bwitek-o] ju'u mochik-∅ bea bwia-po
later DET-ACC coyote-ACC out=run.SG-CLM DET turtle-NOM DM soil-LOC

kibak-taite-k
enter.SG-begin-PFV

'After the coyote ran, the turtle begins to enter into the soil.'
(Silva *et al.*; turtle.coyote: 45)

In opposition, reason clauses must take a nominative subject inside the linked unit:

(13) Ili miisi-∅ pochi-ta bwasia-ka yeu=tomte-k [kiali'ikun miisi o'ou-∅
little cat-NOM short-ACC tail-ACC out=stomach-PFV CLM cat male-NOM

miisi jamut-ta-u o'omte-k]
cat woman-ACC-DIR RDP.angry-PFV

'A little cat with a short tail was born, so the male cat became angry with the female cat.' (Silva *et al.*; cat: 10)

Purpose, temporal and reason clauses also differ in terms of the lexical coding in same-subject clauses. On the one hand, purpose clauses demand the dependent subject to be implicit, as it is shown in (14a-b). Temporal clauses marked by *-kai* also take an implicit subject in co-reference to the main subject (14c).

(14) a. Bwite-k [_$_i$ au esso-se-kai] Waimam-mewi
run.SG-PFV 3SG.REF hide-move.PURP.SG-CLM Guaymas-DIR.PL
'He ran to Guaymas to hide himself (lit. moving to hide).' (Johnson; Cajeme: 34)

b. U o'ou-∅$_i$ juya-u siika [mas-ta _$_i$ me'e-ne-betchi'ibo]
DET man-NOM desert-DIR go.SG.PFV deer-ACC kill.SG-POT-CLM
'The man went to the desert to kill the deer.'

c. [bea sechupti _$_i$ pensasaroa-ka] ne$_i$ aman siika
DM suddenly think-CLM 1SG.NOM there go.SG.PFV
'And, as soon as I suddenly thought [about it], I went there.' (Guerrero; Lalo: 255)

Temporal clauses marked by *-o* allow both, implicit subjects as well as explicit subjects; the latter is the most common situation when the clause encodes sequential events. Notice that, when the subject is overtly expressed, as in (15b), it needs to take accusative case.

(15) a. Maria-$\emptyset_i$ Fermin-ta$_j$ bicha-k [kafe-ta _$_j$ saake-o]
Maria-NOM Fermin-ACC see-PFV coffee-ACC toast-CLM
'María saw Fermín when (he) was toasting coffee.'

b. [Ju-ka Sulumai-ta$_i$ omotria-u yepsa-k-o] jaibu
DET-ACC Sulumai-ACC brush-DIR arrive.SG-PFV-CLM already

_$_i kaa enchi tea-k
NEG 2SG.ACC find-PFV

'After Sulumai got back to the brushes, she couldn't find you.'
(Buitimea; toorokoyori: 102)

Again, in the very same context, reason clauses demand a nominative subject. In (16), the main and dependent subjects are the same, the nominative full pronoun is inside the adverbial unit, while there is a co-indexed missing argument in the main unit.

(16) Nee _$_i a=tetejwa-n [bwe'ituk aapo$_i$ ket bwibwika-n]
1SG.ACC 3SG.ACC=RDP.tell-PSTC CLM 3SG.NOM also RDP.sing-PSTC
'She [my mother] always told me that because she also sang.' (Silva; maejto: 34)

4.3 TAM operators

A purpose clause evokes a situation where the event predicated in the adverbial unit is posterior to the event expressed in the main unit. Because of this, it is expected that the linked verb be unmarked or be marked as a future-oriented verb. The structures coding purpose relations in Yaqui differ from each other with respect to operator dependency. For instance, the aspectual suffix *-ne* 'potential' is only allowed within the *-betchi'ibo* clause (17a), but it is completely prohibited in the other purposive linkages types. That is, the linked verb must be fully unmarked in the structures in (17b–c).

(17) a. Goyo-$\emptyset_i$ yeu=siika [enchi _$_i kaa bit-ne-betchi'ibo]
Goyo-NOM out=go.SG.PFV 2SG.ACC NEG see-POT-CLM
'Goyo left because he didn't want to see you.'

b. Ah, jewi, wam-bicha em saka'a-bo-$\emptyset$ (*saka'a-ne-bo)
ah yes there-toward 2PL.NOM go.PL-move.PURP.PL-PRS
'Ah, you all go and move away to another site.' (Silva; toad: 33)

c. Ili uusi-$\emptyset_i$ siika [demonio-ta jo'a-wi
little boy-NOM go.SG.PFV demon-ACC house-DIR

$_i$ am=to'i-se-kai] (*to'i-ne-se-kai /*to'i-se-ne-kai)
3PL-ACC=take-move.PURP.SG-CLM

'The boy left to take them inside the demon's house' (Johnson; boy: 11)

Temporal clauses show also some restrictions in terms of the TAM values. On one hand, in 'before' and 'after' relations, the linked events are sequential, meaning the TAM information is predetermined with respect to the event described in the main unit. On the other hand, when the main and linked events overlap to some extent, it would be expected that the two events also share the TAM operators (Cristofaro 2003: 164). In her sample, Cristofaro (2013) found that 84 out of 174 languages of the world require balanced (finite-like) adverbial 'when' clauses, 51 require deranked (non finite-like) adverbial units, while the other languages allow both, balanced and deranked 'when' clauses. It seems that Yaqui falls into the third group. In Yaqui, temporal clauses marked by *-kai* demand an unmarked verb form; this is true for simultaneous (18a) and sequential (18b) readings.

(18) a. [junaa'a bikam-po $_i$ bolo-toji-wa-ka nama] bea $_i$ ju-ka
over.there Vicam-LOC bolo-take-PAS-CLM there DM DET-ACC

in jaboi tata'a-k (*toji-ka-wa-kai)
1SG.GEN grandfather RDP.meet-PFV

'While she was engaged in Vicam, she met my grandfather.' (Félix; HVC)

b. [ajta bea $_i$ busam wasekte-ka] bea $_i$ primaria-u bea
until DM six become.old-CLM DM primary.school-DIR DM

kibake-k (*wasekte-k-kai)
enter.SG-PFV

'After he became six years old, he entered primary school.'(Guerrero; Lalo: 87)

c. [kee Sulumai-ta$_i$ bwij-wa-o] aapo$_i$ enchi juya-m nasuk
CLM Sulumai-ACC capture-PAS-CLM 3SG.NOM 2SG.ACC brush-PL middle

e'e-ria-k (*bwij-ne-wa-o)
hide-APPL-PFV

'Before Sulamai was captured, she was able to hide you inside the brush.'
(Buitimea; toorokoroyi: 78)

Although temporal clauses marked by *-o* may take the relevant TAM operators, as depicted in (19), the prototypical situation is for the linked verb to be unmarked for TAM affixes.

(19) Naa bea=te pakun-bichaa yeu=sajak
then DM=1SG.NOM outside-toward out=go.PL.PFV

[a puntaroa-su-k-o]
3SG.ACC suture-COMPL-PFV-CLM

'Then, we went out after they sutured him.' (Guerrero; Lalo: 358)

In contrast, the event described inside the adverbial unit in reason sentences is completely independent in terms of TAM operators, negation and illocutionary force. In the example in (20a), the adverbial unit is marked by the potential suffix *-ne*, and in (20b), by the past perfective *-k*.

(20) a. ¡Banse-ka em troopa-ta lijtaroa-ne ala!
hurry.up-CLM 2SG.GEN troop-ACC prepare-POT AFFIR

[bwe'ituk nee kompayaroa-ne]
CLM 1SG.ACC accompany-POT

'Hurry up and prepare your troop! because you are going to walk with me.' (Johnson; Cajeme: 38)

b. Aapo$_i$ siika [bweituk _$_i$ kaa im jo'a-pea-k]
3SG.NOM go.SG.PFV CLM NEG here live-desire-PFV
'He left because he didn't feel like living here.'

Notice that the two associated units inside reason clauses are even independent of each other in terms of illocutionary force which cannot be independently specified in a subordinate clause of any kind (Van Valin 2005: 214). That is, the event described in the main clause in (20a) is a command, and in (21a) a question; in both examples, the linked unit expresses an affirmative sentence. There are not examples in the sample of purposive or temporal clauses taking imperative or affirmative emphatic particles. If an imperative or question markers occur in a temporal or purpose clause, then it has scope over the two units. See the example in (21b).

(21) a. Jaksa into inim kaa mekka? [bwe'ituk ne jakko juni keet
where DM there NEG far CLM 1SG.NOM never even before

enchim bicha-n-ti] ne ame-u jiia-k
2PL.ACC see-PASTC-CLM 1SG.NOM 3PL-OBL say-PFV

"But where is not far? because I have never seen you two before", I said to them.'
(Buitimea; grandfather: 37)

b. Aman _$_i$ Edgar-ta jiu-se, Jorge-∅$_i$!
there Edgar-ACC look-move.PURP.SG Jorge-NOM
'Jorge, go look for Edgar over there!' (Guerrero; ConvG & A: 46)

4.4 Discourse markers

There is another feature distinguishing reason from purpose and temporal clauses. The occurrence of discourse markers such as *bea* 'so, then', *into* 'and so', at some point of the left-edge, is very common in independent clauses. For reason cases (22a), these markers may freely occur inside the adverbial unit but they can only occur inside the main unit in a purpose clause, except when they are preposed (22b).

(22) a. Em animaalim-tia-'u-m$_i$ juna-me-san tua wa-e kobanamyi-ta
2SG.GEN animal-PL-say-CLM-PL DET-PL-DM INT DET-OBL knowledge-ACC

jippue [bwei'ituk juna-me$_i$ bea che'a wa-ka bwia-ta tomtam-ta-mak
have CLM DET-PL DM more DET-ACC land-ACC give.birth-ACC-INS

emo tu'ure] [bwe'ituk into juna-me$_i$ bea moream ket jippue]
REFL like CLM DM DET-PL DM wisdom also have

'Those that you called animals are creatures with a lot of knowledge, because they live and enjoy the natural environment, and they also have wisdom.' (Buitimea; grandfather: 115–118)

b. [Kaa nee$_j$ ta'a-na-betchi'ibo] bea nee$_j$ _$_i$ ama kobatua-k.
NEG 1SG.ACC know-POT.PASS-CLM DM 1SG.ACC there get.dressed-PFV
'In order for nobody to recognize me, they dressed me up over there.' (Buitimea; grandfather: 53)

The distribution (and function) of adverbial particles inside temporal clauses is still unclear. For instance, *bea* occurs twice in (23), once inside the linked unit marked by *-kai*, and then right after the adverbial clause.

(23) [Num bea jita-su-mak nau ya'a-ka] bea aapa-reo-tu-ka'u
there DM thing-COMPL-COM together make-CLM DM harp-AGT-VBLZ-NMLZ

na'a bea jupa-taka yeu=bwite-k
DEM DM skunk-being out=run.SG-PFV

'When they finished making love, the harper dressed like the skunk ran out.'
(Silva *et al.*; skunk: 23)

To summarize, purpose clauses in Yaqui show a higher degree of syntactic integration. They demand a missing syntactic argument when the main and dependent subjects are the same, and an accusative dependent subject when they are different; the dependent verb must be unmarked or be marked by the potential suffix *-ne* only within the *-betchi'ibo* purposive clause; the linked unit depends on the main unit for the other TAM information including illocutionary force. Now, simultaneous temporal clauses marked by *-kai* share most of the structural

properties of purpose linkages except that the linked verb must always be unmarked. Because they describe anterior or posterior events with respect to the main clause, sequential temporal clauses marked by *-o* may take the relevant tense-aspect operators, though the typical situation is to take an unmarked verb form. Also, the dependent subject must be accusative regardless if it is coreferential or not to the main subject. In opposition, reason clauses do not show any of these dependence features: the linked subject must take nominative case, and it can be overly expressed even when coreferential to the main subject; the two events are completely independent of each other in terms of TAM operators and illocutionary force.

Semantically, purpose and temporal relations predetermine the temporal associations between the two events. Purpose clauses involve two sequential events (the realization of the intended event is presented as possible at a future point in time). For temporal clauses, the two events can overlap in time or one event can follow the other event, but the two states of affairs must hold a temporal relationship. In opposition, reason clauses do not place any of these temporal constraints.

5 The Interclausal semantic relation hierarchy, revisited

As mentioned before, Role and Reference Grammar's approach to clause-internal relational structures predicts that the closer the semantic relationship between the two events is, the stronger the syntactic link joining them will be. In the semantic side of the hierarchy, as depicted in (1) and repeated below, purpose relations outrank reason relations and reason outranks simultaneous and sequential temporal relations. It would be expected then that purpose and reason clauses would make use of a tighter syntactic linkage in comparison to temporal clauses. We have seen that the most common and frequent syntactic structures coding purpose, reason and temporal relations in Yaqui do not follow these predictions.

(1) Interclausal adverbial semantic hierarchy (adapted from Van Valin 2005: 207)
Modifying sub-event [manner, motion, position, means] > purposive > temporal or spatial circumstances > reason > conditional > concessive > simultaneous and sequential temporal > unordered temporal states of affairs

It is worth noting, however, that this semantic hierarchy places two positions for clauses expressing temporal relations: temporal circumstances outranking reason, conditionals and concessive relations, and then simultaneous and sequential temporal relations just above temporally unordered states of affairs.

Accordingly, there are two semantic relations relevant to temporally ordered events. First, circumstances describing the locative or temporal parameters of an event have the logical

structure of a core modifier **be-loc/temp'**([LS$_1$], [LS$_2$]). Van Valin (2005: 194) argues that some prepositions in English can introduce both nominal (24a) and clausal (24b) objects. That is, the relationship of the adverbial clause to the core it modifies is the same as that of a peripheral PP modifying the core, meaning both adjuncts occur at the periphery$_{\text{CORE}}$, e.g. ad-core subordination.

(24) a. Kim saw Pat after the concert
a.' **be-after'**(concert, [**see'**(Kim, Pat)])

b. Kim saw Pat after she arrived at the party
b.' **be-after'**([BECOME **be-at'**(party, 3sgF)], **see'**(Kim, Pat)])

In Yaqui, locative phrases (25a) and locative clauses (25b) are marked by the temporal/locative predicative postposition *-po*.

(25) a. Lunes-po Kantoora-m=ne tea-k plasa-po
Monday-LOC singer-PL=1SG.NOM find-PFV plaza-LOC
'Last Monday, I found the church singers in the plaza.'
a.' **be-at'**(lunes, [**be-at'**(plasa, [**find'**(1sg, kantooram)])])
b. Kantoora-m=ne tea-k [am bwiika-po]
singer-PL=1SG.NOM find-PFV 3PL.ACC sing-LOC
'Last Monday, I found the church singers where they were singing.'
b.' **be-at'**(lunes, [**sing'**(3pl)], [**find'**(1sg, kantooram)])

Thus, locative clauses resemble English temporal clauses modifying the core of the main unit. In other words, they fit perfectly in the description of locative circumstances. However, temporal clauses in Yaqui are not marked by predicative adpositions, but they take adverbial CLMs introducing clauses, not objects. Lower in the semantic continuum, there is another semantic relation that has to do with temporality. This temporal relation distinguishes between (i) one state of affairs being co-temporal with another (simultaneous), which can be decomposed as [LS$_1$] $\wedge$ [LS$_2$], and (ii) one state of affairs following another event temporally, with or without any temporal overlap (sequential), which can be decomposed as in [LS$_1$] & [LS$_2$]. In contrast to temporal circumstances, the English clauses illustrating these relations do not involve the 'after', 'before', or 'when' temporal clause linkage markers:

(26) a. Max was dancing, and at the same time Susan played the piano
[**do'**(Max, [**dance'**(Max)])] $\wedge$ [**do'**(Susan, [**play'**(Susan, piano)])]
b. Juan finished reading the newspaper, and then Carlos walked into the room
[**do'**(John, [**read'**(John, paper)])] & [**do'**(Carlos, [**walk'**(Carlos) & INGR **be-at'**(room, Carlos)])]

Since there is not a specific temporal clause linkage marker (nor adverbs) specifying the association between the two units in the English examples in (26), one may assume that the temporal semantic relations are inferred from the construction as a whole. Although uncommon, juxtaposed clauses as in (27) also favor a sequential relationship in Yaqui. This sentence looks similar to the English temporally ordered constructions.

(27) Junak bea senu jamut-∅ ne-u-bichaa sikaa, wajiwa-bichaa
then DM one woman-NOM 1SG-DIR-toward go.SG.PFV inside-toward

nee wike-k
1SG.ACC drag-PFV

'Then, one woman followed me (and then) dragged me inside.'

(Buitimea; kamam: 51)

In consequence, Yaqui provides good examples for re-examining the ranking of the Interclausal semantic hierarchy. One possibility is to propose a scale like purpose > simultaneous, sequential > reason > inferred temporally ordered events > temporally unordered events. That is, that the so-called circumstance relations specify the temporal relations hold between the two events; lower in the hierarchy are those linkages expressing a lesser degree of semantic cohesion between the propositional units, such as temporality can be inferred (or not) among the units linked in the complex structure.

6 Final comments

This paper describes the major features defining purpose, reason and temporal adverbial constructions in Yaqui. Based on the selection of the CLM, the identity and coding of the dependent subject, and the TAM operator information, it has been shown that simultaneous and sequential clauses show a higher degree of syntactic bonding in comparison to reason clauses, but a less tight linkage when compared with purpose structures. In order to reflect the Yaqui adverbial clause facts, a revision of the Interclausal semantic hierarchy is then suggested, such as purpose outranks simultaneous and sequential relations, and then reason outranks inferred temporally ordered events. It is expected that further studies will verify the cross-linguistic relevance of the proposed semantic scale.

References

Bickel, Balthasar. 1993. Belhare subordination and the theory of topic. In *Studies in Clause Linkage*, K. H. Ebert (ed.), 23–55. Zürich: ASAS.

Cristofaro, Sonia. 2003. *Subordination*. New York: OUP.

—. 2013. 'When' Clauses. In *The World Atlas of Language Structures Online*, M. Dryer & M. Haspelmath (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/126

Dedrick, John & Eugene Casad. 1999. *Sonora Yaqui Language Structure*. Tucson: The University of Arizona Press.

Diessel, Holger. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 45: 449–470.

—. 2008. Iconicity of sequence: A corpus-based analysis on the positioning of temporal adverbial clauses in English. *Cognitive Linguistics* 19(3): 465–490.

Dixon, R.M.W. & Aikhenvald, Alexandra. 2009. *The Semantics of Clause Linking*. Oxford: OUP.

Escalante, Fernando. 1990. *Voice and argument structure* in Yaqui. Ph.D. dissertation, University of Arizona.

Foley, William & Van Valin, Robert D., Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: CUP.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. In *Studies in Language* 4: 333–377.

—. 2001. *Syntax. An Introduction*. Amsterdam/Philadelphia: John Benjamins.

Guerrero, Lilián. 2006. *The Structure and Function of Yaqui Complementation*. Munich: Lincom.

—. 2015. Clause Linkage and temporal relations in Yaqui. *Role and Reference International Conference*. Heinrich-Heine-University Düsseldorf.

—. 2017. On purpose and causal adverbial clauses in Yaqui. *International Journal of American Linguistics* 83(4): 679–718.

Guerrero, Lilián & Van Valin, Robert D., Jr. 2004. Yaqui and the analysis of primary object languages. *International Journal of American Linguistics* 70(3): 290–319.

Haiman, John. 1985. *Natural Syntax: Iconicity and Erosion*. Cambridge: CUP.

Lindenfeld, Jacqueline. 1973. *Yaqui Syntax*. Berkeley: University of California Press.

Noonan, Michel. 2007. Complementation. In *Language Typology and Syntactic Description, Volume 2: Complex Constructions*, T. Shopen (ed.), 42–140. Cambridge: CUP.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical Categories in Australian Languages*, R.M.W. Dixon (ed.), 112–171. Australian Institute of Aboriginal Studies.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Motion Verb Integration and Core Cosubordination in Modern Greek

Soteria Svorou
San José State University

Abstract

This paper explores a subtype of the Modern Greek conjunction reduction construction within the Role and Reference Grammar framework. Focusing on the string *V1 ke V2*, where V_1 is a motion verb, several arguments are offered in support of an analysis for a construction that resembles verb serialization in which syntactically V_1 is cosubordinate to V_2 at the core level and semantically V_1 integrates with V_2 by being interpreted as a frame for V_2. The data and analysis have implications for the role of lexical templates and their interaction with syntactic constructions within sentences but also the interaction of constructions with larger discourse structures.

Introduction

In this paper, I explore the expression of sequential states of affairs in Modern Greek *V1 ke V2* construction ('V and V') with specific emphasis on an instantiation of this construction where V_1 is the motion verb *piγeno/pao* 'I go'. Under certain conditions, the string $V_{piγeno/pao}$ *ke* V_2 is interpreted as a single event involving primarily the event of V_2. I present arguments for a construction that involves core cosubordination at the syntactic level and offer a proposal for handling the conceptual integration of the events. The proposed analysis has implications for the Interclausal Relations Hierarchy, specifically, the temporally related sequential states of affairs, and the semantic representation [LS_1] & [LS_2], as well as the manipulation of the lexical template associated with the verb *piγeno/pao* 'I go'.

1 Conjunction Reduction and the Interclausal Relations Hierarchy

The Interclausal Relations Hierarchy (IRH) developed within Role and Reference Grammar (RRG) represents the interaction of a syntactic relations hierarchy and a semantic relations hierarchy with the goal of capturing cross-linguistic regularities of form-meaning correspondences (Van Valin 2007: 85). It predicts that the closer the semantic relation between two propositions is, the tighter the syntactic bond between them will be. Temporally related sequential states of affairs appear close to the bottom of the hierarchy among the loosest semantic relations representing clearly delineated distinct events or actions with the semantic representation [LS_1] & [LS_2]. Their predicted syntactic juncture level is the sentence and the nexus type is subordination, as in '*Before Juan had finished talking, Carlos entered the room.*' (Van Valin 2007: 82) However, sequential states of affairs with a shared participant may be expressed without an additional explicit mention of that participant in the second clause leading to a syntactic construction known as 'conjunction reduction', as in 'Carlos went to the store and ____ $_i$ bought bread.' Although the events of going to the store and buying bread are sequential and, consequently, identifiable as separate conforming to the [LS_1] & [LS_2] representation, the shared participant contributes to an understanding of a tighter connection between the two events. Langacker (2003, 2012), who associates a clause with an attentional frame, attributes the closeness of the events in a conjunction reduction construction to their occurrence within the same attentional frame.[1] Syntactically, conjunction reduction is captured in RRG by a constructional schema representing clausal cosubordination (Van Valin 2005, 2007). The interpretation resulting from filtering the two logical structures (LS) through the conjunction reduction construction brings the two events conceptually closer, more so than what is predicted by the IRH.

In further exploring the interaction between sequential states of affairs and conjunction reduction, I discuss data and propose an analysis for one subtype of the Modern Greek conjuction reduction construction that integrates the events expressed by the conjoined verbs even further, demonstrating closer syntactic bonds at the core level and tighter semantic relations.

2 Modern Greek Conjunction Reduction

Modern Greek is a pro-drop language and, similar to other pro-drop languages like Serbo-Croatian (Van Valin 2005: 231), allows for the "reduction" of a topical participant in clausal conjunction. In (1), the experiencer of *iðe* 'saw' is understood to be coreferential with the

[1] Attentional framing represents the packaging of expressive and conceptual content into chunks of manageable size (Langacker 2012: 562).

agent of *piγe* 'went', Vasilis, who is the topical participant.[2] A pronoun indicating the topical participant (*aftos* 'he') is normally not expected in this position, unless contrastive emphasis is desired.

(1) o vasili-s_i piγ-e s-ti krakovia-ø ke ____$_i$ ið-e
ART Vasilis-NOM go.PST.PFV-3SG to.ART Krakow-ACC and see.PST.PFV-3SG

tus roliŋg stounz
ART.ACC Rolling Stones

'Vasilisi went to Krakow and ____$_i$ saw the Rolling stones.'

Syntactically, the juncture-nexus type of this construction is clausal cosubordination, similar to what was proposed for English. The main argument for establishing this type relies on illocutionary force. Illocutionary force must be shared across all conjuncts. In (2a), the interrogative particle *mipos* has scope over the entire sentence, as it is not possible to question only the first conjunct, (2b), or only the second conjunct, (2c).[3]

(2) a. mipos o vasili-s_i piγ-e s-ti krakovia-ø ke ____$_i$
textscq ART Vasilis-NOM go.PST.PFV-3SG to.ART Krakow-ACC and

ið-e tus roliŋg stounz?
see.PST.PFV-3SG ART.ACC Rolling Stones

'Did Vasilis go to Krakow and see the Rolling Stones?'

b. *[mipos o vasili-s_i piγ-e s-ti krakovia-ø?] ke [____$_i$
Q ART Vasilis-NOM go.PST.PFV-3SG to.ART Krakow-ACC and

ið-e tus roliŋg stounz]
see.PST.PFV-3SG ART.ACC Rolling Stones

'*Did Vasilis go to Krakow? and he saw the Rolling Stones.'

c. *[o vasili-s_i piγ-e s-ti krakovia-ø] ke [mipos ____$_i$
ART Vasilis-NOM go.PST.PFV-3SG to.ART Krakow-ACC and Q

ið-e tus roliŋg stounz?]
see.PST.PFV-3SG ART.ACC Rolling Stones

'*Vasilis went to Krakow and did he see the Rolling Stones?'

Additionally, a look at operator sharing, which is a criterion for core cosubordination, yields negative results. In (3), the tense/aspect operator need not be the same in the two conjuncts.

[2] The glosses of the examples are simplified in certain ways. Although articles, nouns, and adjectives all carry affixes marking gender, number, and case, such information is generally not included here, except in the first few examples.

[3] (2c) could be acceptable only if the first conjunct is separated from the second by a long pause, which would indicate a different fit into the discourse.

(3) o vasili-si piγ-e s-ti krakovia-ø] ke [____$_i$ θa
ART Vasilis-NOM go.PST.PFV-3SG to.ART Krakow-ACC and FUT

ð-i tus roliŋg stounz]
see.PFV-3SG ART.ACC Rolling Stones

'Vasilis went to Krakow and he will see the Rolling stones'

Also, the negative operator *ðen* can have scope over only one of the conjuncts, as in (4).

(4) [o vasili-si ðen piγ-e s-ti krakovia-ø] ke [____$_i$
ART Vasilis-NOM NEG go.PST.PFV-3SG to.ART Krakow-ACC and

ið-e tus roliŋg stounz]
see.PST.PFV-3SG ART.ACC Rolling Stones

'Vasilis did not go to Krakow and he saw the Rolling stones'

Based on these arguments, I conclude that the juncture-nexus type of the Modern Greek conjunction reduction construction is clausal cosubordination. Figure 1 presents a simplified representation of the constituent and operator projection of the conjunction reduction construction.

In terms of meaning, the verb of the first conjunct in (1), *piγe* 'went', is understood as an active accomplishment expressing a motion event carried out by Vasilis which resulted in the relocation of Vasilis to Krakow; the second verb, *iðe* 'saw', is understood as an event also carried out by Vasilis and the clause linkage marker (CLM) *ke* 'and' contributes to the interpretation that the second event follows the completion of the first event. Following Van Valin (2009) based on Ohori (2001, 2005), the semantic representation of a given conjunction reduction sentence would be an instantiation of [LS_1] & [LS_2] with coreferencing for the common topic participant. Hence, the semantic representation of (1) would be as follows:

[**do'**(Vasilis$_i$, [**move.away.from.reference.point'**(Vasilis$_i$)]) & INGR **be-at'**(Krakow, Vasilis$_i$)] & [**see'**(x_$_i$, roliŋg stounz)]

In terms of linking syntactic and semantic representations, the linking algorithm proposed in Van Valin (2005: 279–281) is assumed to be at work, with the provisions for a dependent-marking accusative language with a voice opposition. The lack of lexical instantiation of the privileged syntactic argument of the second clause can be accounted for in terms of a constructional schema. Given the similarities between the English and the Modern Greek conjunction reduction schemas, the reader is directed to Van Valin (2005: 231) for an explication of the schema.

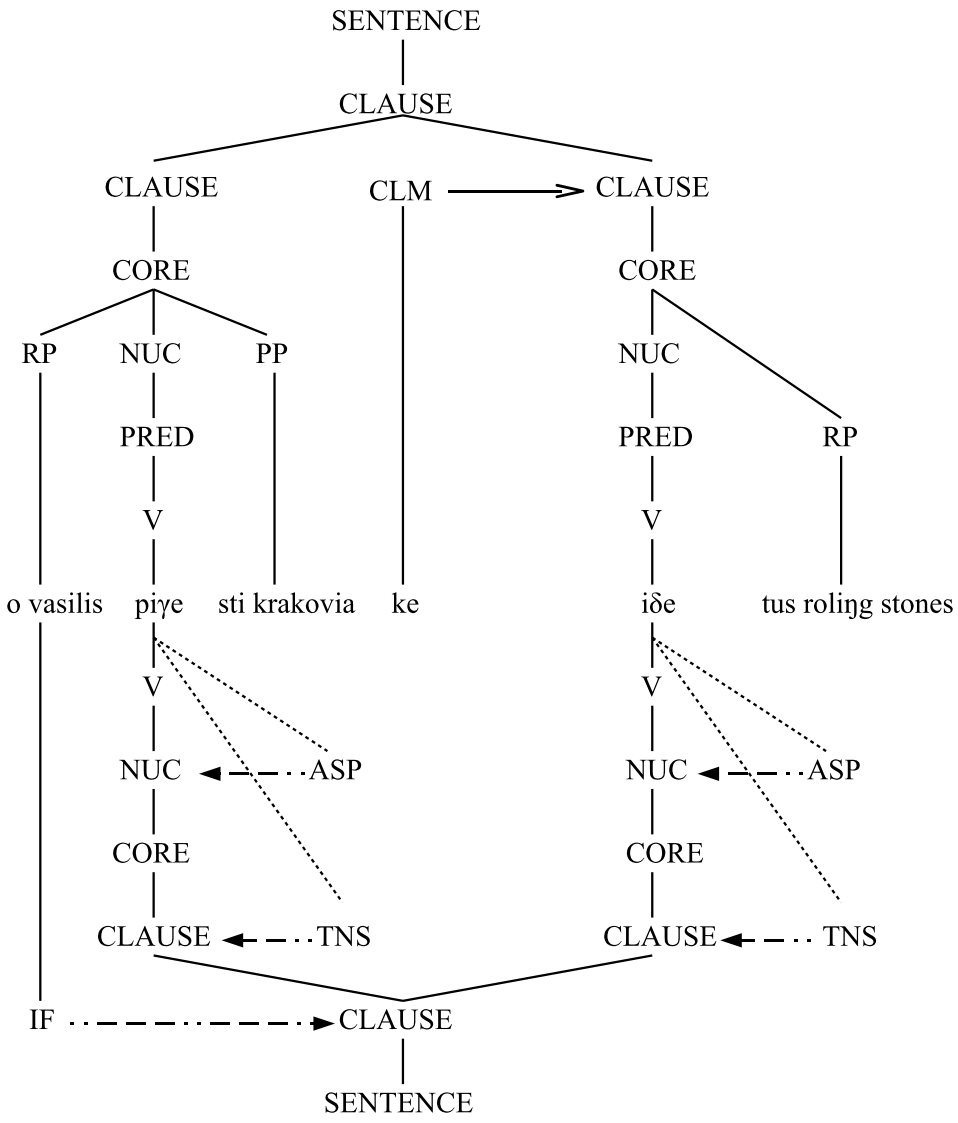


Fig. 1: Constituent and operator projection for (1)

Providing a general schema for this construction entails a possibly similar behavior across types of verb instantiating the nuclei of the conjoined clauses. An investigation, however, of some representatives of the class of motion verbs reveals that, when such verbs occur in the first conjunct clause, within certain discourse environments, they create ecologies that contribute to semantic and pragmatic interpretations that deviate from the ones specified in the schema and allow for further tightening of the syntactic relations among the conjunct verbs. Specifically, $V_{pi\gamma eno/pao}$ *ke* V_2 'go and V_2', where V_2 is an accomplishment, and V_{erxome} *ke* V_2 'come and V_2', where V_2 is an achievement, show cosubordination at the core level and interface with a semantic interpretation that involves a conceptually more integrated macro-event (Bohnemeyer & Van Valin 2009) than what is expected based on the conjunction reduction constructional schema. I refer to this construction as the Integrated V_1 *ke* V_2.

3 Integrated V_1 ke V_2

As part of a larger study that involves the semantics of Modern Greek motion verbs *piγeno/pao* 'I go' and *erxome* 'I come' and posture verbs *kaθome* 'I sit' and *sikonome* 'I stand up' in structures conjoined with other verbs, I conducted the query [lemma MOTION VERB] – [lemma *ke*] - [POS verb] using the Hellenic National Corpus (HNC), allowing for one word to intervene between *ke* and V_2 in the string.[4] I manually organized the resulting examples in three categories according to whether V_1 is interpreted as clearly involving motion, as clearly not involving motion, and as resulting in an ambiguous interpretation. Examples (5) and (6) illustrate the no-motion interpretation, which was clear in 40 of the 361 tokens retrieved for the lemma *piγeno/pao*. This interpretation is the focus of the current discussion.

(5) o kostas gatsjuðis bik-e ston aγona me stoxo na riks-i
ART K. G. enter.PST-3SG in.ART race with goal CLM throw.PFV-3SG

pano apo 83 metra, aftos omos **piγ-e** **ke** **kseperas**-e
over from meters he however go.PST.PFV-3SG and surpass.PST.PFV-3SG

ta 87!
ART

'Kostas Gatsioudis entered the competition with the goal to throw more than 83 meters, but he (went and) surpassed 87 meters!'

[4] The Hellenic National Corpus (HNC) is maintained by the Institute of Language and Speech Processing and is available at http://hnc.ilsp.gr/en/default.asp. It currently contains 47 million words of written text from a variety of contemporary sources.

Kostas Gatsioudis, as a javelin thrower, is in control of the act of throwing but not necessarily of the act of surpassing a certain threshold. In (5), the expression *piγe ke kseperase* 'went and surpassed' is not understood as involving Gatsioudis's movement followed by his surpassing 87 meters in throwing; rather, in using this expression, the speaker is attributing intentionality to the actor of the event of surpassing a threshold, even though this event is not necessarily under complete control of the actor. In so doing, a contrast is created between the statement of the clause containing *piγe ke kseperase* 'went and surpassed' and the initial clause, reinforced by the particle *omos* 'but'; the contrast results in the understanding of surprise (Note also the use of the exclamation mark.)

Similarly in (6), there is no understanding of motion towards the deictic center in the expression *irθe ke γalinepse*. Even though Stavros, as a human, is capable of motion, there is no motion understood in this case; rather, the understanding is that Stavros underwent a change of state in calming down and the speaker emphasizes the culmination of that process, with an added implication of speaker's satisfaction at the outcome.

(6) ala otan se liγo γiris-e o stavros ki iδ-e ekina ta omorfa
but when in little turn.PST-3SG ART S. and see.PST.3SG those ART beautiful

mavra matia tis etima na klaps-un, ksexas-e ke θimus ke
black eyes POSS ready CLM cry.PERF-3SG forget.PST-3SG and anger and

katakrisis ki **irθ-e** **ke** **γalineps-e**, metaniomenos tora
condemnations CLM come.PST-3SG CLM calm.down.PST-3SG remorseful now

γia ti skliri simberifora tu.
for ART harsh behavior POSS

'But when Stavros returned in a while and saw her beautiful black eyes ready to cry, he forgot his anger and his condemnations and he (came and) calmed down, remorseful for his harsh behavior.'

The examples retrieved from the HNC showed some important tendencies. Verbs that appear as V_2 in the $V_{piγeno/pao}$ *ke* V_2 string are generally accomplishments. The shared arguments in the examples are mostly animate (in 39 of the 40 examples) and actors, although in 11 examples the shared arguments are undergoers. A representative example is given in (7), in which the first person plural is the Undergoer of the act of losing but additionally it is attributed with a certain degree of intentionality or resoluteness in undergoing this event. The understanding of intentionality is contributed by *piγe ke* ... 'went and ...'

(7) paniγiris-ame kapos perisotero me ton olimbiako ke **piγ-ame**
celebrate.PST-1PL somewhat more with ART Olympiakos CLM go.PST.PFV-1PL

ke xas-ame sto proto pexniði sto OAKA
CLM lose.PST.PFV-1PL at.ART first game in.ART OAKA

'We celebrated somewhat more with Olympiakos and (went and) lost at the first game in OAKA.'

Verbs that appear as V_2 in the *V_{erxome} ke V_2* string are generally inchoative, change of state verbs, or indicate a process that results in a new state. The shared argument tends to be an inanimate or a non-volitional entity (such as chocolate, sauce, policy, state, elements, forces, experience, money, political party, models, conscience, revolution, mythology, etc.), with twice as many inanimate than animate subjects (32 out of 47 tokens of the construction). Additionally, in 27 of the 47 tokens, V_2 is an intransitive verb with an undergoer argument. Example (8) illustrates a non-volitional shared participant, *to spiri* 'the pimple', which stands metaphorically for an explosive situation. Even though it is the (only) argument of *irθe* 'came', *to spiri* 'the pimple' is clearly an entity incapable of either initiating or undergoing motion; as the only argument of *eskase* 'popped', it is clearly an undergoer.

(8) to spiri **irθ-e** **ke eskas-e**, me θimata tus θites oso
ART pimple come.PST-3SG CLM pop.PST-3SG with victims ART bullies as.much.as

ke ta θimata tus
and ART victims POSS

'The pimple (came and) popped, with victims the bullies as much as their victims.'

As a subtype of the conjunction reduction construction, the integrated *V_1 ke V_2* construction inherits the sharing of illocutionary force of the conjoined clauses. But beyond that, the data below indicate that the integrated *V_1 ke V_2* construction involves a tighter syntactic relation exhibiting characteristics of core cosubordination, such as sharing of core-level operators (internal negation, modality, and event quantification) and modification by temporal and spatial expressions (adverbials, clauses, tenses). In this paper, I discuss primarily the *$V_{piγeno/pao}$ ke V_2* instantiation of the construction and only make reference to *V_{erxome} ke V_2* for the sake of contrast.

3.1 Shared Negation Operator

In (9) the negative operator *ðen* has scope not only over the first predicate, *pas*, but also over the entire core of the following conjunct, *pernis proaγoγi se mja vð omað a*.

(9) Se opjaðipote ðulja ine ðiskolo na anaðixt-is, ðen pas ke
in any job is hard CLM move.up-2SG NEG go.NPST-2SG and

pern-is proaɣoɣi se mja vðomaða
get.NPST-2SG promotion in ART week
'In any job it's hard to move up, you don't go and get a promotion in a week.'

Similarly, the negative operator *min*, which generally appears in the imperative and in embedded structures (10), has scope over both cores. The sense, however, that arises is that the negative operator does not negate two events but rather one.

(10) Pola eɣraps-a ... afto pu iθel-a na p-o ine
many write.PST.PFV-1SG this REL want.PST-1SG CLM say.PFV-1SG is

na min pa-s ke ðos-is tosa lefta ɣia ena bufan,
CLM NEG go.PFV-2SG and give.PFV-2SG so.much money for one jacket

ena kranos ke ena zevɣari ɣantja.
one helmet and one pair gloves
'I wrote a lot ... what I wanted to say is to not go and give so much money for a jacket, a helmet and a pair of gloves.'

3.2 Shared Modal Operator

Deontic modality operators in Modern Greek, such as the impersonal verb of obligation *prepi* 'must' (11a), which appears with a calcified 3SG morphology, and the personal verb of ability *boro* 'can, may' (11b) both require that the modified verb be preceded by the CLM *na* joined in a cosubordinate structure.

(11) a. prepi na me stiriks-is
must CLM 1SG.ACC support.PFV-2SG
'You must support me.'

b. bor-o na se stiriks-o
can-1SG CLM 2SG.ACC support.PFV-1SG
'I can support you.'

These modal operators have scope over both cores in the $V_{pi\gamma eno}$ *ke* V_2 construction. So, in (12a), the obligation on the part of the 2PL participant is understood to include the visiting and, as such, the operator is shared among the two cores. Repeating the operator with each verb, (12b), creates juncture at the clausal level. But in the case where V_2 is the verb *episkeptome* 'I visit', repetition of the modal creates a redundancy that, with other verbs, would only be interpreted as emphatic obligation to carry out the second event, except, here, it creates some

dissonance as the events of going and visiting are not distinguishable, because visiting entails going.

(12) a. afto to meros ine ena apo ta meri pu prepi na pate ke na
this ART place is one of ART places REL must CLM go.2PL and CLM

episkefθite
visit.2PL

'This place is one of the places that you must go and visit.'

b. ?Afto to meros ine ena apo ta meri pu prepi na pat-e ke
this ART place is one of ART places REL must CLM go.PFV-2SG and

prepi na episkefθ-ite
must CLM visit.PFV-2SG

'This place is one of the places that you must go and must visit.'

3.3 Shared Event Quantification Operator

Modern Greek has a verbal prefix *ksana-* 'again', that comes from a similar adverb, which has the effect of indicating the repetition of the event expressed by the host verb. Prefixation of *ksana* on *piγeno* in the $V_{piγeno}$ *ke* V_2 construction leads to ambiguity in (13) as to whether the second verb is to be interpreted as being repeated as well. Ambiguity can be avoided by repeating *ksana*, as in (14).

(13) Ksana-piγ-e ke epsaks-e kati pu
EVQ-go.PST.PFV-3SG and look.for.PST.PFV-3SG something REL

ix-e γraps-i
have.PST.3SG write.PFV-3SG

'He went again and looked (again) for something he had written.'

(14) Oxi, ðen mu apandis-an, ala ksana-piγ-a
no NEG 1SG.PRO.GEN reply.PST.3PL but EVQ-go.PST.PFV-1SG

ke ksana-rotis-a
and EVQ-ask.PST.PFV-3SG

'No they did not reply, but I went again and asked again.'

In (15), however, which instantiates the no-motion interpretation of the verb *piγeno/pao*, the sense of repetition applies more to the second event. The exhortation here, rather than stressing not going again, emphasizes that the addressee should not say such things again intentionally. This interpretation can only be achieved if the juncture happens at the core level.

(15) Min ksana-pas ke pis tetja praɣmata!
NEG EVQ-go.IMP and say.IMP such things
'Don't go and say such things again!'

3.4 Shared tense and adverbial modifiers

According to the Macro-Event Property (MEP) proposed by Bohnemeyer & Van Valin (2009), cosubordinate events are equally modified by the same temporal and spatial expressions (adverbials, clauses, tenses).[5] The verbs in the *$V_{piɣeno/pao}$ ke V_2* construction must agree in tense/aspect, obligatory components of the verbal morphology. As (16b) indicates, if the sequence of tenses is not observed the resulting utterance is non-sensical.

(16) a. o olimbiakos piɣ-e ke proselav-e ton texniko
ART Olympiakos go.PST.PFV-3SG and hire.PST.PFV-3SG ART technical
simvulo tis eθnikis
advisor ART Ethniki
'Olympiakos went and hired the technical advisor of Ethniki.

b. ??o olimbiakos piɣ-e ke θa proslav-i ton texniko
ART Olympiakos go.PST.PFV-3SG and FUT hire.PFV.3SG ART technical
simvulo tis eθnik-is
advisor ART.GEN Ethniki-GEN
??'Olympiakos went and will hire the technical advisor of Ethniki.'

As the tense operator, however, is entangled with the aspect operator in morphological and semantic terms, a valid question may be raised as to whether the aspect operator is also shared, which would then make an argument for a nuclear juncture. All 383 examples of *$V_{piɣeno/pao}$ ke V_2* retrieved from the corpus, regardless of the event integration category they represent, show agreement of V_1 and V_2 in terms of tense and aspect. If the two verbs carry different aspectual settings, the two events must be conceived as completely separate (17). Note also that the inclusion of an argument adjunct PP specifying the destination is an additional indicator of a separate motion event.

(17) o ɣiatros piɣ-e sto nosokomio ke xiruɾɣus-e
ART doctor go.PST.PFV-3SG to.ART hospital CLM operate.PST.IPFV-3SG
oli ti nixta
all ART night
'The doctor went to the hospital and was operating the whole night.'

[5] 'The Macro-Event Property (MEP): An event-denoting construction has the MEP iff it combines only with those time-positional and durational modifiers that have scope over all subevents it entails.' (Bohnemeyer & Van Valin 2009)

The requirements for agreement of V_1 and V_2 in aspect in this construction, in addition to all other shared operators, is a serious argument for nuclear cosubordination. It is, however, challenged by a counterargument presented in 3.6.

Temporal and spatial adverbial modifiers also apply to the whole conjunction; there cannot be separate modifiers of the two predicates. In (18b), the addition of separate temporal modifiers *tin triti* 'on Tuesday' and *tin tetarti* 'on Wednesday' for the two verbs leads to the interpretation of sequential but temporally distant separate events; in contrast, in (18a) there's no time lapse as *piγe ke sakateftike* '(went and) became crippled' only allows for the conceptualization of one event, that of becoming crippled. Similarly in (18c), the addition of a locative modifier *stin ispania* 'to Spain' right after *piγe* 'went' dissociates the two predicates and leads to the interpretation as two separate events with temporal lapse between them. Examples (18b, c) instantiate clausal cosubordination, whereas (18a) involves cosubordination at the core level.

(18) a. ala o arian piγ-e ke sakateftik-e
but ART Arian go.PST.PFV-3SG and become.crippled.PST.PFV-3SG

sta prokrimatika tu evrobasket '99 kj afis-e
in-ART qualifiers ART Eurobasket '99 and leave.PST.PFV-3SG

ton olimbiako orfano
ART Olympiakos orphan

'But Arian went and became crippled in the Eurobasket '99 qualifiers and left Olympiakos orphaned'

b. ala o arian piγ-e tin triti ke
but ART Arian go.PST.PFV-3SG ART Tuesday and

sakateftik-e tin tetarti sta prokrimatika
become.crippled.PST.PFV-3SG ART Wednesday in-ART qualifiers

tu evrobasket '99 kj afise ton olimbiako orfano
ART Eurobasket '99 and leave.PST.PFV-3SG ART Olympiakos orphan

'But Arian went on Tuesday and became crippled on Wednesday in the Eurobasket '99 qualifiers and left Olympiakos orphaned'

c. ala o arian piγ-e s-tin ispania ke
but ART Arian go.PST.PFV-3SG to.ART Spain and

sakateftik-e s-ta prokrimatika tu evrobasket '99 kj
become.crippled.PST.PFV-3SG in-ART qualifiers ART Eurobasket '99 and

afis-e ton olimbiako orfano
leave.PST-3SG ART Olympiakos orphan

'But Arian went to Spain and became crippled in the Eurobasket '99 qualifiers and left Olympiakos orphaned'

Note also that in the case of clausal cosubordination where the verb *piγeno* may be accompanied by prepositional complements or adjuncts, it is possible to have different tense/aspect operators in the two conjuncts, as illustrated by example (3).

3.5 Fronting of V_2 argument RP in PrCS

Narrow focus on an RP constituent can be indicated with intonation but also by fronting it to PrCS. In (19), the second argument RP of V_2, *afta* 'these', is fronted preceding not only *ðjavazis* 'you read' but also *pas ke* 'go and'. In (20), the fronting involves the interrogative word *ti* 'what', which is a core argument of the verb *les* 'you say'. These examples show that the second argument of V_2 can appear in PrCS not of that verb's core, but rather that it is directly attached to the CLAUSE node, preceding both of the linked cores.

(19) Mja zoi ðiamartirese ja ta englimata ke prota prota afta pas
one life protest.PRES.2SG for ART crimes and first first these go.PRES.2SG

ke ðjavazis.
and read.PRES.2SG

'You always protest for the crimes but you go and read these first.'

(20) Ti pas ke les?
What go.PRES.2SG and say.PRES.2SG
'What do you go and say?'

In summary, I have shown above that sentences involving the integrated $V_{piγeno/pao}$ *ke* V_2 construction should be analyzed syntactically as involving core cosubordination because of a) being treated as a unit in fronting of the second V_2 argument to PrCS, b) shared adverbial modifiers and c) negation, modality, event quantification, and tense operator sharing.

3.6 Is the juncture at the nuclear level?

Sharing of the aspectual operator in the integrated $V_{piγeno/pao}$ *ke* V_2 construction, as discussed in 3.4, presents a hypothesis for nuclear juncture of the two verbs. Nuclear juncture requires the combination of two (or more) nuclei into "a single, complex nucleus with a single set of core arguments." (Van Valin 2005: 191) In the integrated $V_{piγeno/pao}$ *ke* V_2 construction, a subtype of conjunction reduction, one core argument, the subject, is shared between the two verbs. A potential locative argument of the motion verb V_1 is suppressed and not even presumed to exist in this construction. In the case of an intransitive V_2 (as in 18a), the construction may appear to be an example of nuclear cosubordination. However, in also looking

at examples with a transitive V_2 (as in 16a), an argument against the nuclear juncture hypothesis can be offered. The argument stems from the phenomenon of clitic doubling. Clitic doubling involves the appearance, right before the matrix verb, of a clitic pronoun that agrees in gender, number, and case with its "double" RP, which may be an object of the matrix verb, among other roles (Papangeli 2000; Alexiadou & Anagnostopoulou 2000; Kechagias 2011.) In (21), the clitic *to* cliticizes to the nucleus *ðiavasa*.

(21) to ðiavas-a to piima-ø
CLI.N.SG.ACC read.PST.PFV-1SG ART poem-ACC
'I read the poem.'

The only item that can interfere between the verb and the clitic is the perfect auxiliary *exo* 'I have', a nucleus operator, as in *to exo diavasi to piima* 'I have read the poem'. In the case of the construction $V_{pi\gamma eno/pao}$ *ke* V_2, if we hypothesize a juncture as the nuclear level, we would expect the clitic to appear before the complex nucleus. This, however, is not an option, as the ungrammaticality of (22a) attests; rather, the preferred position for the clitic would be right before V_2 (21b).

(22) a. *i simpatriotes tu ton piγ-an ke
ART compatriots POSS CLI.M.SG.ACC go.PST.PFV-3PL CLM
θimiθik-an to γalo
remember.PST.PFV-3PL ART French
'His compatriots (went and) remembered the French guy.'

b. i simpatriotes tu piγ-an ke
ART compatriots POSS go.PST.PFV-3PL CLM
ton θimiθik-an to γalo
CLI.M.SG.ACC remember.PST.PFV-3PL ART French
'His compatriots (went and) remembered the French guy.'

This possibility shows that the clitic is within the core of V_2 and that the juncture is at the level of the core, not the nucleus, at least for transitive V_2. It could very well be the case that transitive and intransitive V_2 present different levels of juncture.

3.7 Construction type: Serial verb?

Unlike conjunction reduction, the integrated $V_{pi\gamma eno/pao}$ *ke* V_2 construction exhibits many of the typological characteristics of verb serialization, as specified in Aikhenvald (2006); namely, it involves a monoclausal sequence of verbs that have the same tense, aspect, and polarity

value, share a core argument, and are conceptualized as a single event. In verb serialization, the two verbs constitute a single predicate, without any markers of subordination or coordination. This last characteristic poses a certain difficulty to a clear-cut analysis of the construction at hand as a serial verb. The intervening element *ke* instantiates a coordinating conjunction in the conjunction reduction construction and has the function of conjoining the two clauses. It could be argued, however, that, in the integrated $V_{pi\gamma eno/pao}$ *ke* V_2 construction, *ke* does not have the function of conjoining clauses, but cores, and, as such, is not a typical conjunction. Holton, Mackridge, & Philippaki-Warburton (1997), commenting on *ke*, list a number of functions, in addition to clear coordination, that resemble those of subordinating conjunctions, including temporal sequence, result, and reason, and hint at idiosyncratic meanings after perception verbs and verbs such as *arxizo* 'I begin'. As already hinted at, with further details to be explained in 3.8, the motion verb instantiating V_1 in the construction does not express an event, but aspects of its semantic structure frame the event expressed by V_2. In such a structure, the semantic contribution of *ke* is rather minimal, and in fact, it is possible to leave out *ke* altogether, as (23) shows.

(23) enas γnostos mu fititis piγe pire to paso apo
a known POSS student go.PST.PFV-3SG take.PST.PFV-3SG ART pass from

to katastima
ART store

'A student I know went and got the pass from the store.'
(retrieved from: http://www.myphone.gr/forum/showthread.php?t=345377&page=4)

For these reasons, I propose that the integrated $V_{pi\gamma eno/pao}$ *ke* V_2 construction is a type of serial verb construction, even though it does not have the grammatical generality of the prototypical serial verb constructions in prototypical serial verb languages, such as Thai.

3.8 Semantic representation

If the construction presented above is assumed to be a particular instantiation of conjunction reduction, then the semantic representation would be [LS_1] & [LS_2]. The logical structure of the motion verb, which occupies the LS_1 position, can be lifted from the lexicon, where it presumably has the following specification:

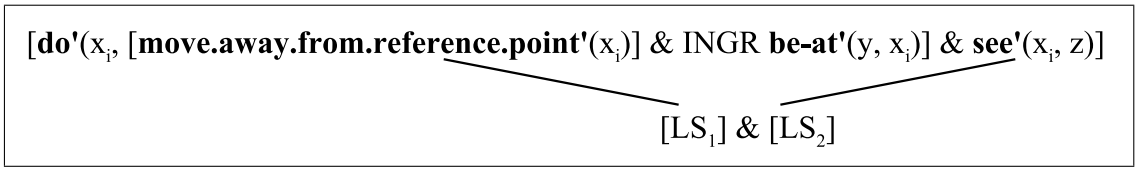
piγeno/pao 'I go': **do'**(x, [**move.away.from.reference.point'**(x)] & INGR **be-at'**(y, x)

This lexical template may be employed in actual sentences in a way that all the variables in the semantic representation also appear in the syntactic representation, fulfilling the Completeness Constraint (Van Valin 2005: 129); alternatively, it may appear with the accomplishment part of the representation suppressed, i.e., with no specification of destination. In such case, there is still an underlying understanding that a participant who moves will end up in a different location but the specification of the location is not considered important to mention in the current discourse. It appears, then, that only one maximal lexical template needs be specified in the lexicon and that constructional and discourse environments contribute to manipulations of elements of the template. This approach has also been advocated in Cortés Rodrigues & Torres Medina (2002) in discussing Old English motion verbs and their activity, active accomplishment, and causative active accomplishment uses. In the Lexical Construction Model (Ruiz de Mendoza & Mairal 2008; Mairal & Ruiz de Mendoza 2009), lexical templates are considered low level constructions that can be absorbed within higher level constructions adapting to the specifications of the construction.

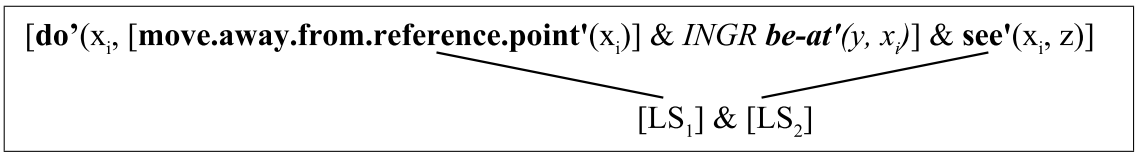
Taking an approach similar to the Lexical Construction Model, the following is a sketch of the linking of lexicon, construction, and semantic representation. Lexical templates from the lexicon are selected and interfaced with a construction. In the case of conjunction reduction, as exemplified in (1), the templates for the two predicates are incorporated into the semantic representation appropriate to the construction. The construction calls for a shared topical participant which is represented by x, coreferenced in the two templates.

Lexicon
piγeno/pao: **do'**(x, [**move.away.from.reference.point'**(x)] & INGR **be-at'**(y, x)
vlepo/δo: **see'**(x, y)

Conjunction Reduction Construction



In cases where the end-point location is not expressed, (24), the inchoative predicate is suppressed in the lexical template (indicated by the *italicized* part), leading to the following representation:



(24) o vasilis piγ-e ke ið-e tus roliŋg stounz
ART Vasilis go.PST.PFV-3SG and see.PST.PFV-3SG ART Rolling Stones
'Vasilis went and saw the Rolling stones'

In the case of examples of the integrated $V_{piγeno/pao}$ *ke* V_2, as I pointed out in the discussion above, the motion verb in V_1 does not express a motion event.[6] Then, the lexical template **do'**(x, [**move.away.from.reference.point'**(x)]) for *piγeno* 'I go' does not reflect the conceptualization that ensues in the above specified context. If filtering a lexical template through a construction is an effective way of accounting for the resulting semantics, then in order to capture the resulting semantics, the construction must specify that V_1 loses its predicate character, is de-predicativized. If it is de-predicativized, most of the representation would be suppressed. What would remain in the understanding of (5) with V *piγeno/pao*, for example, is only a sense of intentionality attributed by the speaker to the shared participant in his involvement in the V_2 event.

The sense of intentionality in the integrated $V_{piγeno/pao}$ *ke* V_2 can be best seen in testing for compatibility with an adverbial of chance. While the event of surpassing the threshold of 87 meters itself can be attributed to chance, as in (25a) where the verb is modified by *tixea* 'accidentally', the expression *piγe ke kseperase* 'went and surpassed' could not be interpreted as having happened accidentally, as in (25b).

(25) o kostas gatsjuðis **tixea** **kseperas-e** ta 87 metra.
ART K. G. accidentally surpass.PST.PFV-3SG ART meters
'Kostas Gatsioudis accidentally surpassed 87 meters.'

a. ??o kostas gatsjuðis tixea **piγ-e** **ke**
ART Kostas Gatsioudis accidentally go.PST.PFV-3SG and

kseperas-e ta 87.
surpass.PST.PFV-3SG ART 87

'Kostas Gatsioudis accidentally (went and) surpassed 87 meters.'

The examples in (25) are my renditions, based on my intuition, of corpus example (5) to highlight the incompatibility of *piγe ke kseperase* 'went and surpassed' with an adverbial of

[6] Similar English constructions with similar effects are discussed in Newman & Li (2007) and Newman & Rice (2008).

chance. Example (26), however, is from the corpus and nicely addresses the issue at hand, as the speaker of the utterance actually rejects an attempt (presumably by the addresses) to claim that an event described by *piγe ke ðolofonise* 'went and murdered' can be modified by *tixea* 'accidentally'.

(26) Ke veveos afto to opio prospaθ-ite na p-ite oti tixea kapios
and of.course this ART REL try.PRES-2PL CLM say-2PL CLM accidentally someone

piγ-e ke ðolofonis-e eksinda zoa, ðen stek-i.
go.PST.PFV-3SG CLM murder.PST.PFV-3SG sixty animals NEG stand.PRES-3SG

'And of course that which you're trying to say, that someone accidentally went and murdered sixty animals, does not hold.'

The intentionality that $V_{piγeno/pao}$ contributes in the integrated $V_{piγeno/pao}$ *ke* V_2 can be viewed as a conceptual frame imposed on the meaning of V_2. Such a frame reflects the original semantics of the motion verb from which it is distilled. The morphophonological form *piγeno* 'I go' gives conceptual access to the initial stage of moving which is fueled by intention – but without a subsequent motion entailment; (by contrast, the morphophonological form *erxome* 'I come' gives access to the final stage of moving – but without previous motion entailment — resulting in modifying V_2 with a culmination frame). The construction facilitates the distillation process. In a decompositional semantics, the result could be captured by an operator, such as AIM. So, the representation of (27), a simplified version of (5), would be:

AIM **surpass'**(gatsjuðis, 87 metra) ∧ BECOME **surpass'**(gatsjuðis, 87 metra)

(27) o gatsjuðis **piγe ke kseperase** ta 87 metra
ART Gatsioudis go.PST.PFV-3SG and surpass.PST.PFV-3SG ART meters
'Gatsioudis (went and) surpassed 87 meters.'

AIM is proposed as an operator that imposes an intentional conceptual frame on a completed event. Purposives ('He went to buy bread'), represented by the predicate **want'**, also involve an intentionality frame. The intentionality in the latter is assumed to be the driving force of the second verb but there are no entailments as to the fulfillment of the second verb.

One difficulty with capturing the semantics of the construction with an operator is showing the relation of the operator to the original motion verb semantics. A potential way to circumvent such difficulty is to adopt an imagistic representation of the semantics of the motion. Langacker's (2012) Cognitive Grammar imagistic representations are amenable to highlighting relevant portions of the meaning potential of an expression while allowing other parts to remain in the background. As an imagistic representation has not been pursued so far within RRG, this option requires further exploration.

3.9 Constituent and operator projection

The arguments presented above support an analysis of the Modern Greek integrated $V_{pi\gamma eno/pao}$ *ke* V_2 construction as a core cosubordination construction. A simplified constituent and operator projection of (28) is represented in Figure 2 below (page 300).

(28) i simpatriotes tu piγ-an ke θimiθik-an to γalo
ART compatriots POSS go.PST.PFV-3PL CLM remember.PST.PFV-3PL ART French
'His compatriots (went and) remembered the French guy.'

The de-predicativization of *piγeno* 'I go' could be handled in the constituent projection by removing the PRED node from the projection yet retaining the verbal status, as the verb is still fully inflected for person/number and tense/aspect.

3.10 Linking

When the aforementioned lexical template for *piγeno/pao* is linked to constituent structure in actual sentences, by default, participant *x* is selected as actor, following the actor-undergoer hierarchy (Van Valin 2005: 61). However, when it is employed as V_1 in sentences instantiating the integrated *V1 ke V2* construction, the macrorole of the argument of the motion verb depends on the macrorole of the argument of V_2 in the construction and, consequently, it can be either actor or undergoer. So, in (29), *i simpatriotes tu* 'his compatriots' has the macrorole of actor, as the first argument of *θimiθikan* 'they remembered' is an experiencer and gets selected as the actor; but in (30) (same as 18a) *o arian* 'Arian' is in fact an undergoer, as the only argument of *sakateftike* 'was crippled', a passive form, is a patient and gets selected as an undergoer. In (31), the only argument of *emine engios* 'got pregnant', an active form, is similarly an undergoer.

(29) i simpatriotes tu piγ-an ke θimiθik-an to γalo
ART compatriots POSS go.PST.PFV-3PL and remember.PST.PFV-3PL ART French
'His compatriots (went and) remembered the French guy.'

(30) ala o arian piγ-e ke sakateftik-e sta prokrimatika
but ART Arian go.PST.PFV-3SG and cripple.PASS.PST.PFV-3SG in-ART qualifiers

tu evrobasket '99 kj afis-e ton olimbiako orfano
ART Eurobasket '99 and leave.PST.PFV-3SG ART Olympiakos orphan

'But Arian (went and) was crippled in the Eurobasket '99 qualifiers and left Olympiakos orphaned'

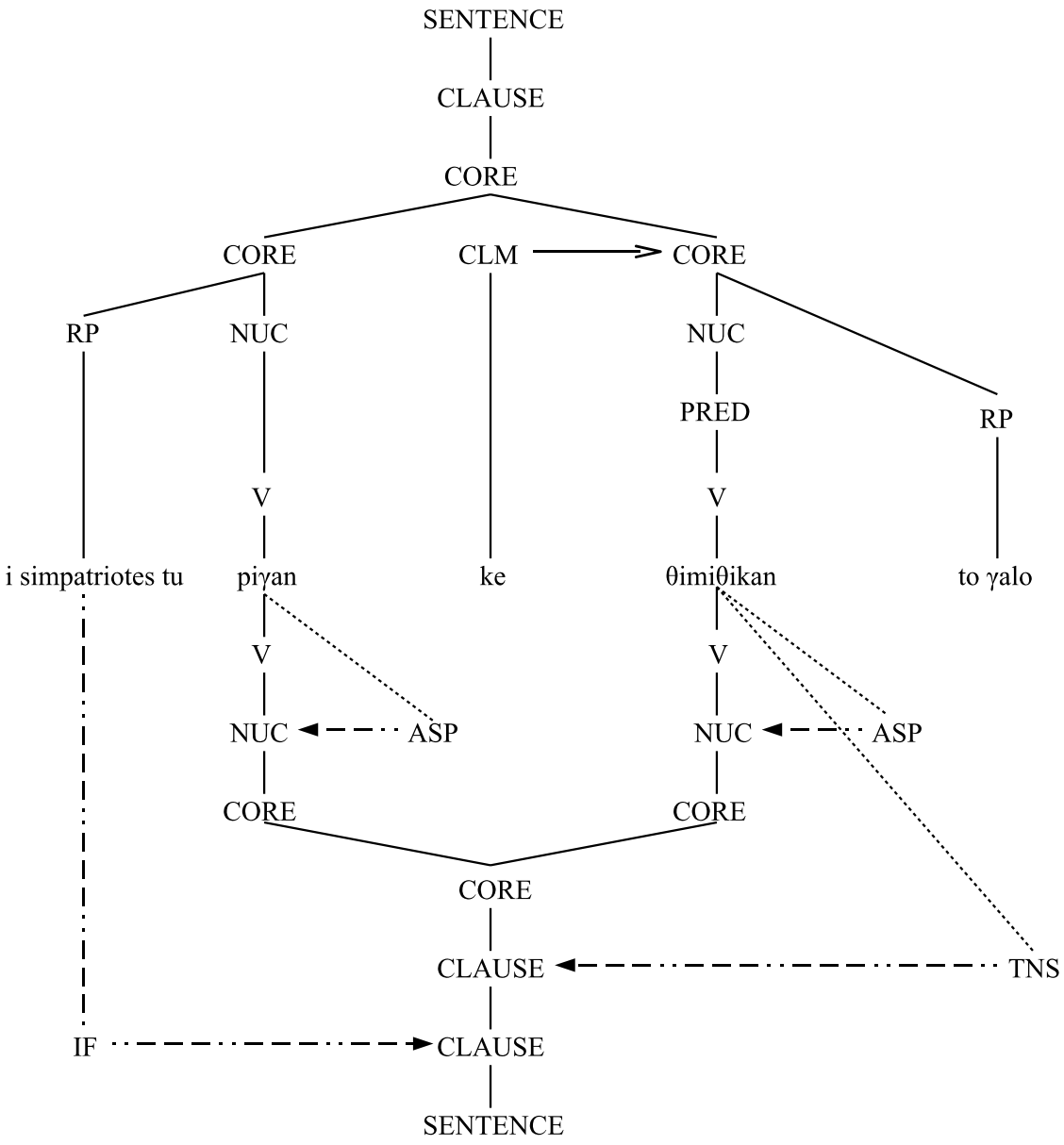


Fig. 2: Constituent and operator projection of core cosubordination structure

(31) ke me ton tropo mu tin agapus-a ti maria ke tora afti
and with ART way POSS CLI.F.ACC love.PST.IMP-1SG ART Maria and now she

piγ-e ke emine engios me kapjon alo…
go.PST.PFV-3SG and stay.PST-3SG pregnant with someone else

'And in my own way I loved Mary and now she (went and) got pregnant with someone else…'

Macrorole assignment then is in line with the integration analysis presented here. The shared argument receives its thematic role from V_2, making V_2 the matrix verb and leaving the motion verb in a framing role as an operator.

3.11 The proposed constructional schema

To summarize the analysis of the Modern Greek integrated *V_1 ke V_2* construction, where V_1 is instantiated by the verb *piγeno* 'I go', I present Table 1. The constructional specifications with regards to syntax and morphology are the same across the class of verbs that show integration, such as motion and posture verbs. The semantics and pragmatics are adjusted to the source V_1 verb.

CONSTRUCTION	Modern Greek integrated conjunction reduction
SYNTAX	Juncture: core
	Nexus: cosubordination
	Construction type: serial verb
	[CL [CORE **NP** [NUC …]] CLM [CORE [NUC…]…]…]
	Unit Template(s): (to be specified later)
	PSA: Core 1: controller = syntactic controller
	Core 2: pivot=variable syntactic pivot = pragmatic pivot
	Linking: controlled argument in core 2 = pragmatic pivot (accusative construction)
MORPHOLOGY	Optional CLM: conjunction **ke / kj**
SEMANTICS	If V_1 = *piγeno/pao* → intentional framing of linked core
PRAGMATICS	Illocutionary force: shared across conjuncts
	Focus structure: unspecified
	Topic: shared topical participant
	Speaker attitude: surprise, indignation

Table 1: Constructional schema for Modern Greek 'integrated V_1 ke V_2'

4 Conclusion

In concluding this discussion, I hope to have shown that Modern Greek *V_1 ke V_2* constructions, under certain conditions, exhibit juncture-nexus relations of core cosubordination and present a closely knit semantic integration of the lexical templates of the two verbs. Similar English constructions have been explored by Newman & Rice (2008) and the analysis proposed here could be extended to the English data as well. Constructions of this kind appear to follow the iconicity principle, which is the basis of the Interclausal Relations Hierarchy, in that core cosubordination is interfaced with semantically integrated conjunction of verbs. What needs to be stressed, however, is that constructions play an important role vis-à-vis the Interclausal Relations Hierarchy, which is a default characterization of clausal syntactic and semantic relations.

Several other points can be stressed but perhaps what's important is the articulation of what more needs to be done in exploring the construction under discussion. The inherent semantics of the verb in V_1 of the integrated *V_1 ke V_2* construction plays a role in the semantic interpretation of the construction, as I pointed out. Investigating the classes of verbs (e.g., motion, posture, phase) that can appear in such a string will provide a firmer basis for articulating what aspects of a V_1 can be constructionalized. In considering the representation of the semantics of verbs, lexical templates that are based on predicates from a metalanguage may be too rigid to account for conceptual modifications due to constructional and contextual influences. An exploration of proposals for imagistic representations (e.g., Cognitive Grammar) may provide a potential solution but would require adaptation to the RRG formalism. Finally, corpus data provide a wealth of information regarding factors that facilitate a particular understanding of a construction, including the discourse environment of such constructions. For example, the particular understanding of surprise on the part of the speaker that arises in the integrated *$V_{pi\gamma eno/pao}$ ke V_2* construction, as discussed briefly above, appears in discourse settings that invite contrast, with a preceding adverbial clause (e.g. 'although', 'instead'), a contrastive particle (*omos*, *ala* 'but'), in a rhetorical question (*ti pas ke kanis?* "What do you go and do?") or exclamatory statement (*pu piγes ke eblekses!* "Where did you go and get caught up!"). The intricate interaction of discourse structures with constructions and the lexicon is an important next step towards a fuller understanding of the contribution of each structure.

References

Aikhenvald, Alexandra. 2006. Serial verb constructions in a typological perspective. In *Serial Verb Constructions: A Cross-Linguistic Typology*, A. Y. Aikhenvald & R. M.W. Dixon (eds) 1–87. Oxford: OUP.

Alexiadou, Artemis & Anagnostopoulou, Elena. 2000. Clitic doubling and (non-) configurationality. *Proceedings of NELS* 30. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.513.3527&rep=rep1&type=pdf

Bohnemeyer, Jürgen & Van Valin, Robert D., Jr. 2009. The Macro-Event property and the LSC. Presentation at the *2009 International Conference on Role and Reference Grammar*. Berkeley, CA. [Available on RRG website]

Butler, Christopher S. & Martin Arista, Javier (eds). 2009. *Deconstructing Constructions*. Amsterdam: John Benjamins.

Cortés Rodrigues, Francisco & Torres Medina, Dolores. 2002. Lexical templates and linking rules for some Old English verbs of motion. [Available on RRG website]

Holton, David, Mackridge, Peter & Philippaki-Warburton, Irene. 1997. *Greek Grammar: A Comprehensive Grammar of the Modern Language*. London: Routledge.

Institute of Language and Speech Processing. *Hellenic National Corpus*. http://hnc.ilsp.gr/en/default.asp

Kechagias, Axiotis. 2011. A new look at clitic doubling in Standard Modern Greek. *UCL Working Papers in Linguistics* 2011(1): 1–29.

Langacker, Ronald W. 2003. Constructional integration, grammaticization, and serial verb constructions. *Language and Linguistics* 4(2): 251–278.

—. 2012. Elliptic coordination. *Cognitive Linguistics* 23(3): 555–599.

Newman, John & Lin, Jingxia. 2007. The purposefulness of going: A corpus-linguistic study. In *Corpora and ICT in Language Studies* [Lodz Studies in Language 13], J. Walinski, K. Kredens & S. Gozdz-Roszkowski (eds) 293–308. Frankfurt am Main: Peter Lang.

Newman, John & Rice, Sally. 2008. Asymmetry in English multi-verb sequences: A corpus-based approach. In *Asymmetric Events: An interpretation*, B. Lewandowska-Tomaszczyk (ed.) 3–22. Amsterdam: John Benjamins.

Mairal, Ricardo & Ruiz de Mendoza, Francisco. 2009. Levels of description and explanation in meaning construction. In *Deconstructing Constructions* [Studies in Language Companion Series 107], C. S. Butler & J. M. Arista (eds) 153–198. Amsterdam: John Benjamins.

Ohori, Toshio. 2001. Some thoughts on a new systematization of Interclausal Semantic Relations. Paper presented at the 2001 Role and Reference Grammar Conference, University of California, Santa Barbara. [Available on RRG website]

—. 2005. More thoughts on the semantic representation in RRG: Event types and the semantics of clause linkage. Paper presented at the 2005 International Role and Reference Grammar Conference, Taiwan. [Available on RRG website]

Papangeli, Dimitra. 2000. Clitic doubling in Modern Greek: A head-complement relation. *UCL Working Papers in Linguistics* 12: 473–499.

Roberts, John R. 2010. How to find serial verbs in English: an RRG analysis of phase verb constructions. In *Proceedings of the Tenth International Conference on Role and Reference Grammar*, W. Nakamura (ed.) 301–314. [Available on RRG website]

Ruiz de Mendoza Ibañez, Francisco Jose & Mairal Usón, Ricardo. 2008. Levels of description and constraining factors in meaning construction: an introduction to the Lexical Construction Model. *Folia Linguistica* 42(3–4): 355–400.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

—. 2007. Recent developments in the Role and Reference Grammar theory of clause linkage. *Language and Linguistics* 8(1): 71–93.

—. 2009. An overview of Role and Reference Grammar. [Available on RRG website]

—. 2012. Lexical Representation, Co-composition, and Linking Syntax and Semantics. In *Advances in Generative Lexicon Theory*, J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki & C. Lee (eds) 67–107. Dordrecht: Springer.

The Syntax of the Directed Motion Subtype of the Spanish Gerund Construction

Luis París
CONICET
UNCuyo

Abstract

The directed motion subtype (DM) of the Spanish Gerund Construction (SGC) consists of a directed motion matrix verb adverbially modified by a gerund phrase (GP) without the mediation of any explicit linking marker. I drew evidence from different distribution patterns to conclude that DM comprises an ad-nuclear subordination linkage. The analysis of DM hints towards two broader hypotheses that would need further explorations. First, it supports a non-categorical division between arguments and adjuncts since GP is a non-lexically required constituent that behaves like an argument. Second, DM shows that the interclausal relations hierarchy is also sensitive to the meaning type of the matrix verb in addition to the semantic relation between events.

1 Introduction

The Spanish Gerund Construction (SGC) consists of a matrix verb predicate adverbially modified by a gerund phrase (GP) without the mediation of any explicit linking marker. The key issue in the syntax of SGC is about establishing the specific linkage that joins the structures that contain as nuclei the matrix verb and the gerund form, respectively. On earlier papers, I have shown that SGC encompasses a rather complex set of subtypes that is best represented as an *isa* network of (sub)constructions. The upper section of the network is schematically represented in Figure 1.

They all have in common a GP that modifies adverbially a matrix verb. The GP is not lexically required; therefore it is an adjunct that should be represented as a modifier in some periphery. From a purely semantic stand point, both the matrix verb and the verb in GP denote each an eventuality and, hence, SGC encompasses necessarily a relation between two eventualities.

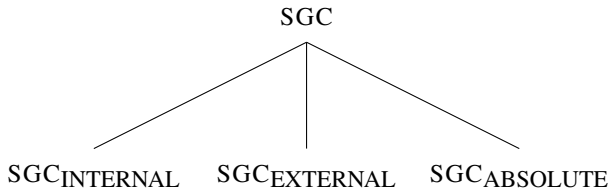


Fig. 1: SGC and its subtypes

Depending upon the subtype of SGC involved the relation might be cause, consequence, condition, temporal overlap or sequence, or different types of manner (either means or incidental). The subtypes directly dominated by SGC in the construction network are absolute, internal and external or SGC$_{ABSOLUTE}$, SGC$_{INTERNAL}$ and SGC$_{EXTERNAL}$, respectively. SGC$_{ABSOLUTE}$ posits GP as a clause in the left detached position (it is, thus, separated from the main clause by a pause), allows it to have independent PSA and aspectual markers and the relation between the eventualities can be condition, consequence or just mere temporal sequence. In stark contrast, SGC$_{INTERNAL}$ has GP in the post-verbal position (typical of manner adverbs), GP cannot have a PSA and the eventualities hold a part-whole relation that is consistent with manner modification. The subtype SGC$_{EXTERNAL}$ represents a type restricted to GPs with individual state verbs. The semantic relation between the eventualities is that of temporal overlap; the GP offers the temporal location of the event denoted by the matrix verb. Sentence (1) is an example of SGC$_{ABSOLUTE}$; sentence (2) is an instance of SGC$_{INTERNAL}$, and sentence (3) represents the SGC$_{EXTERNAL}$ subtype.

(1) Habiendo terminado el libro, Pedro se tomó vacaciones.
having finished the book, Pedro REFL take-PRET.3SG vacations
'Having finished the book, Pedro went on vacation.'

(2) Mi hermana se gana la vida escribiendo.
my sister REFL earns her living writing
'My sister earns her living as a writer.'

(3) María llegó a Madrid siendo abogada.
María arrive-PRET.3SG to Madrid being lawyer
'María was already a lawyer when she came to Madrid.'

In this paper I will concentrate on SGC$_{INTERNAL}$. In the network that represents SGC, SGC$_{INTERNAL}$ subsumes two subtypes, namely, Directed Motion (DM) and non-Directed Motion (non-DM). The former consists of a matrix verb whose logical structure denotes bounded motion events typically through non-durative telic representations (i.e. Achievements) like *entrar* ('enter'), *llegar* ('arrive'), *salir* ('exit') but also durative ones (i.e. Accomplishments) like *venir* ('come')

and *ir* ('go') while the nucleus of GP can be any kind of verb but individual states (París 2003). Sentence (4) illustrates an instance of this type.

(4) El professor entró al aula rengueando.
the teacher enter-PRET.3SG the classroom limping
'The teacher limped into the classroom.'

In contrast, non-DM has a matrix verb with any type of logical structure (but individual states) like in, for example, sentence (2) above. The same open variability holds for the nucleus of GP. The main thesis of this paper is that DM comprises a syntax-semantics interface pairing that differentiates it from the rest of SGC types and, in particular, from its closest node in the construction network, namely, the one occupied by the non-DM type. Specifically, the clause linkage that corresponds to DM is a non-daughter nuclear subordination juncture-nexus; in contrast, non-DM involves a non-daughter core subordination nexus. The evidence that supports this analysis is drawn from *wh*-questions, cleft sentences and secondary predication. The two predicates cannot be separated in a cleft structure in the DM type whereas they can in non-DM instances. In addition, the DM type allows a core macrorole argument of GP to show up in the pre-core slot of the matrix clause whereas this is not possible in the non-DM type. Finally, GP might be combined with a modifying secondary predicate in the DM type but it cannot in the non-DM one.

The behavior of GP in the DM type under my analysis calls for a gradient view of the argument-adjunct distinction. This contrast has been central to most syntactic theories for decades and RRG has been indeed a strong advocate for it. RRG makes a strong distinction between lexically required and non-required constituents (i.e. arguments and adjuncts), that is, constituents that show up in the LS of a verb and constituents that do not (Van Valin & LaPolla 1997: 159ff., 326ff.). In particular, the linking algorithm connects the LS of a verb (or verb like) predicate with a specific syntactic template . In the context of clause linkage, complement clauses satisfy lexically required elements and the same is true for infinitive forms that follow matrix verbs in Spanish. GP in the DM type is not a required constituent but it will be shown throughout the paper that it behaves syntactically closer to complement forms than to adverbial ones. This analysis supports the need of refinements to the categorical distinction between arguments and adjuncts by acknowledging different classes of adjuncts organized on a scale that set its highest member on an equal foot with arguments from the distribution standpoint.

The paper is organized as follows. In the next section I will present an overall description of the construction. In the third section I will concentrate on characterizing formally both the DM and the non-DM subtypes. This is the central part of the paper since it is devoted to justify the

difference among those major subtypes on the base of different clause linkages. Finally, the conclusion should summarize the points that have been substantiated in the previous section and it also includes some final comments framing the proposal in the large discussion about the formal properties of adjuncts.

2 An overview of the construction

SGC constitutes an array of grammatical structures involving a variety of syntactic templates coupled with semantic relations that connect the matrix event e_M with the gerund event e_G in different ways.

The subtype of SGC 'Internal/Directed Motion/Means' or $SGC_{INTERNAL/DM/MEANS}$ contains a matrix verb that denotes a directed motion event and the event denoted by GP represents a Means to perform or achieve the matrix event. Specifically, GP describes the kind of motion carried out in order to reach the goal denoted by the matrix verb like in example (5). GP is a non-required constituent (i.e. an adjunct); hence, it can be dropped without any effect on the grammaticality of the sentence as seen in (6).

(5) Pedro entró al aula corriendo.
Pedro enter-PRET.3SG to-the classroom running
'Pedro ran into the classroom'

(6) Pedro entró al aula.
Pedro enter-PRET.3SG to-the classroom
'Pedro entered the classroom'

The matrix verb *entrar* (enter) is an achievement and a directed motion verb that describes an instantaneous change of location. It does not codify semantically a Path as can be seen by the fact that a measure phrase cannot be predicated of it (see Kita (1999) for a similar point in Japanese and Kawachi (2005) for the opposite view).

(7) El cartero entró (#dos metros) a la casa (#dos metros).
the mailman enter-PRET.3SG (seven feet) to the house (seven feet)'
'The mailman entered (seven feet) the house (seven feet).' (Literal)

Even if the Path is syntactically invisible, it is conceptually present since it is impossible to refer with *entrar* ('enter') to an event that does not involve the motion of a participant that traverses the intermediate points of a Path (a sort of 'star trek' motion). GP contributes indeed this missing conceptual information. It presents the activity that renders a description of the motion involved in traversing that Path (Talmy 2000). *Correr* ('running') is an activity that denotes a particular means of motion. Without this information – that is, if GP were absent like

in (6) – the means of motion should be pragmatically inferred. The addressee is expected to assume that the actor walked into the classroom through a Q2 generalized implicature (París 2006a; 2008). Namely, the typical way for a human being to traverse the Path leading to a classroom is by walking. Since this is predictable, the Q2 maxim of conversation – say the least you can regarding the requirements of the communicative situation – relieves Spanish speakers from the need of expressing means of motion. They are indeed prevented to do it. If means of motion is expressed and predictable, the utterance becomes pragmatically odd as shown in (8). This sentence would only be appropriate if the shared background knowledge contains a proposition that has the teacher unable to walk at that time. Only unexpected means of motion can be expressed by GP since it is the unmarked potential narrow focus (París 2006a). In other words, GP is necessarily part of the potential focus domain of SGC and, hence, it needs to carry new information.

(8) #El profesor entró al aula caminando.
the professor enter-PRET.3SG to-the classroom walking
'The professor walked into the classroom.'

The two eventualities – the matrix event e_M and the running event e_G – constitute a single complex one where e_G is a proper part of the e_M through the means relation (París 2003, 2006b; París and Koenig 2003; Krifka 1998).

Alternatively, GP might introduce an eventuality that is incidental to e_M, namely, e_M and e_G occur in different causal chains while performed simultaneously by the same participant. Still, the form conveys a single event representation made out of two events. This is the subtype $\text{SGC}_{\text{INTERNAL/DM/INCIDENTAL}}$ illustrated in (8).

(9) El niño entró cantando.
the child enter-PRET.3SG singing
'The child entered singing.' (Literal)

In addition, the subtype $\text{SGC}_{\text{INTERNAL/CAUSE}}$ illustrated in (9) conveys events in the same causal chain. Its particular feature is that the matrix verb LS conveys a causative event (*romper* 'break'). Typically, causative verbs contain an unspecified causing event in their LS and GP describes the causing eventuality in that LS.

(10) El comando ruso acusó a los independentistas de haber roto seis
the Russian command accused to the independentistas from having broken six
semanas de tregua lanzando una ofensiva contra las posiciones rusas...
weeks of truce launching an attack against the Russian positions...
'The Russian commander in Chechenia blamed the independentist from having broken six months of peace talk by launching an attack against Russian positions...'
[Associated Press]

Furthermore, the subtypes $\text{SGC}_{\text{INTERNAL/DM/MEANS}}$ and $\text{SGC}_{\text{INTERNAL/}\neg\text{DM/INCIDENTAL}}$ hold the same semantic relation than the related subtypes in (5) and (8). However, the matrix verb in non-DM is not a directed motion (DM) verb. Both non-DM subtypes are represented in (10) and (11), respectively.

(11) Corría cojeando con tanto ímpetu que nadie era más ágil [...]
run-IMP.3SG limping with so.much energy that nobody was more agile
'S/he used to run limping with so much energy that nobody was more agile.' [CREA]

(12) María corrió por la calle gritando.
María run-PRET.3SG by the street screaming
'María ran through the streets screaming.'

All the subtypes presented so far share the same abstract schema labeled $\text{SGC}_{\text{INTERNAL}}$. They all denote a single complex event constituted by a mereological relation: e_G is a proper part of e_M. The imperfectivity associated with the aspectual meaning of the Spanish gerund makes e_G unbounded and facilitates it to be contained as a part by another event. Formally, GP is an adjunct that functions like a manner adverb modifying the meaning of the matrix verb. Every instance of the subtypes included within $\text{SGC}_{\text{INTERNAL}}$ above might have been an appropriate answer to a manner question (in relation to (1) 'how did Pedro enter?') and they all can be paraphrased in (2), (4) or (5) above by 'y is a manner of x' where 'y' is the whole complex event description that adds the meaning of the matrix verb plus the meaning of the gerund and 'x' involves only the meaning of the matrix verb (i.e. *entrar corriendo es una manera de entrar*; literally 'enter running is a way of entering').
An important restriction on $\text{SGC}_{\text{INTERNAL}}$ is that GP cannot denote an 'individual-state' as shown in (13).

(13) #Pedro corrió por la calle siendo abogado.
Pedro run-PRET.3SG through the street being lawyer
'*Pedro ran through the street being a lawyer.' (Literal)

A permanent condition of a participant cannot hold a manner relation to another event. Grammatical sentences containing GPs that introduce an individual state invoke an event relation different from manner. Unlike $\text{SGC}_{\text{INTERNAL}}$, in (14) GP introduces an eventuality that temporally frames the matrix event. In this sense GP is 'external' to the matrix event and hence this subtype is labeled $\text{SGC}_{\text{EXTERNAL}}$.

(14) María llegó a Madrid siendo abogada.
María arrive-PRET.3SG to Madrid being lawyer
'María was already a lawyer when she came to Madrid.'

The adverbial function of GP in SGC makes it sensitive to the opposition between pre-posed and post-posed adverbs. Just like pre-posed adverbs, GPs can be placed in the LDP as in (14), separated by a pause from the matrix clause. GP describes an event that is externally connected to the matrix event by relations like 'consequence' as in (14), 'condition', 'concession', 'temporal sequence/overlap'. In the Latin tradition, this type of left detached adjunct clauses headed by non-finite forms have been called 'Absolute' (Stump 1985) and this is the term I use for it.

(15) Habiendo vendido Pedro la casa, sus sobrinas se quedaron
having sell-PSTP Pedro the house, his nieces REF stay-PRET.3PL

sin vacaciones.
without vacations

'Pedro having sold his house, his nieces were left without vacations.'

Figure 2 makes certain assumptions by representing SGC as an inheritance *isa* network of related subconstructions. First, it takes SGC to be a grammatical unitary object and, hence, it attributes properties that are shared by every subtype. Second, the dominating nodes consist of partial grammatical representations that are further specified down the network. Third, each subtype differentiates itself from the rest by possessing particular properties. The basic characterization of the clause linkage of every major subtype I will offer in the next section should help to make these assumptions realistic. Every node embodies a specific meaning and a formal structure making each of them a recognizable unit. The only controversial node in this sense is $SGC_{INTERNAL/\neg DM}$ since it is identified by the negation of a particular content (DM) rather than by a positive one. The exposure of the properties of this node should motivate this otherwise unattractive labeling.

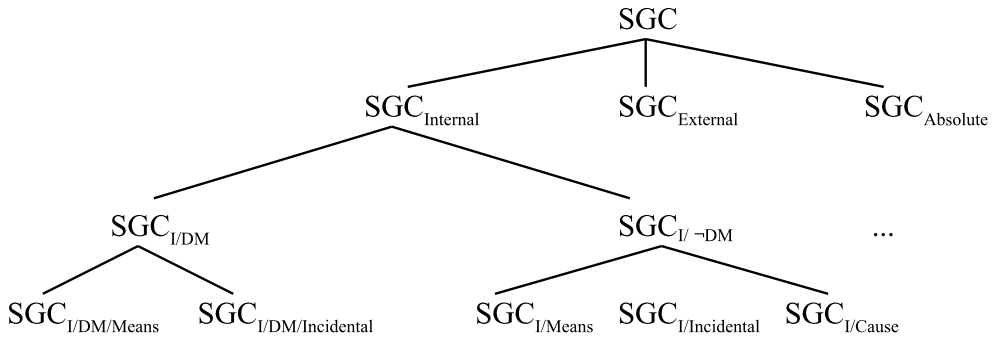


Fig. 2: *isa*-network of SGC

3 The clause linkage that corresponds to each subtype of SGC

The first divergence in the SGC network can be easily accounted for. GP in the Absolute subtype is pre-posed, is allowed to have an independent PSA as well as independent markers of grammatical aspect. None of these properties are licensed by $SGC_{INTERNAL}$ as shown in (15a), (15b) and (15c), respectively.

(16) a. *Corriendo, Pedro entró al aula.
running, Pedro enter-PRET.3SG to-the classroom
'*Running, Pedro entered the classroom' (Literal)

b. Pedro entró al aula (*María) corriendo (*María).
Pedro enter-PRET.3SG to-the classroom (María) running (María)
'*Pedro entered the classroom (María) running (María).' (Literal)

c. *Pedro entró al aula habiendo corrido.
Pedro enter-PRET.3SG to-the classroom having run-PSTP
'*Pedro entered the classroom having run.' (Literal)

For reasons of space, I cannot offer an appropriate discussion of the full scope of the relevant data but the evidence given so far seems provisionally consistent with a sentential subordination structure for $SGC_{ABSOLUTE}$ (Van Valin 2005: 192). GP is in the LDP since separated necessarily by a pause. Furthermore, it always carries presupposed information and, hence, it cannot be under the scope of the potential focus domain of the sentence and, hence, both GP and the matrix clause can fulfill different Speech Acts.

The subtype $SGC_{EXTERNAL}$ represents a very restricted set that has mixed properties. The postposed GP as in (13) above presents the circumstance that surrounds and temporally frames the matrix event. Just like $SGC_{ABSOLUTE}$, GP can have an independent PSA as shown in (16).

(17) María llegó a Madrid siendo Juan abogado.
María arrive-PRET.3SG to Madrid being Juan lawyer
'María came to Madrid when Juan was already a lawyer.'

However, GP is not the presupposition of the sentence – unlike in $SGC_{ABSOLUTE}$ – but the unmarked narrow focus and, hence, it is necessarily part of the potential focus domain. For example, GP – but not the matrix verb – is within the scope of a sentential negation operator in (17). She was not a lawyer but she did arrive to Madrid.

(18) María no llegó a Madrid siendo abogada.
María not arrive-PRET.3SG to Madrid being lawyer
'María was not a lawyer when she came to Madrid.'

Furthermore, if the GPs in (16) or (17) are fronted to LDP, the interpretation of the sentence changes dramatically; the GP does not present merely the temporal circumstance of the matrix event but now both events are connected by a consequence or reason relation (i.e. *She did not reach Madrid because she was a lawyer*). Assigning to $\text{SGC}_{\text{EXTERNAL}}$ an ad-clausal subordination juncture-nexus allows us to derive all the properties mentioned above.

$\text{SGC}_{\text{INTERNAL}}$ encompasses different subtypes and each of them deserves a particular consideration. Since they encode semantic relations that are higher up in the interclausal hierarchy, the possibility of a nuclear co-subordination juncture-nexus should be considered. This is the tightest juncture and corresponds to verbs that form a lexical unit through syntactic rather than morphological means. Nuclear co-subordination junctures in Spanish display the attachment to the matrix verb of clitic pronouns that express arguments of the non-finite verb. This is illustrated by an instance of the causative construction in (18) (París 2012).

(19) Pedro lo hizo lavar.
Pedro it-ACC make-PRET.3SG wash-INF
'Pedro had it washed.'

The pronoun *lo* expresses the undergoer argument of *lavar* ('wash') but it is attached to the matrix verb *hacer* ('make'/'have'). This placement of the clitic signals that the two nuclei form a complex nucleus domain where all their arguments can be expressed. None of the subtypes of $\text{SGC}_{\text{INTERNAL}}$ permit this clitic placement as shown in (19), (20) and (21) for DM/MEANS, DM/INCIDENTAL and non-DM, respectively.

(20) *Pedro lo salió arrastrando.
Pedro it-ACC exit-PRET.3SG dragging
'Pedro exited while dragging it.'

(21) *Pedro lo salió cantando.
Pedro it-ACC.M exit-PRET.3SG singing
'Pedro exited (somewhere) singing it.'

(22) *Pedro la rompió la caja apretando.
Pedro it-ACC.F break-PRET.3SG the box squeezing
'Pedro broke the box by squeezing it.'

$\text{SGC}_{\text{INTERNAL}}$ is an obligatory but non-fixed control structure thereby the PSA of GP cannot be expressed (i.e. it is the Pivot of the construction) and the potential argument associated with it should be co-referential with either the PSA or the non-PSA macro-role argument of the matrix verb. In sentence (22), either the coach or the child could be the runners.

(23) El entrenador trajo al niño corriendo.
the coach bring-PRET.3SG to-the child running
'The coach brought the child by running.'

Since an argument of the matrix verb determines the reference of an argument of the second verb, this type of structures is named 'control' and in RRG they all involve core junctures. In English there is clear evidence that the typical (fixed) control structure (i.e. 'John wants to buy a cake') involves co-subordination nexus. The same does not need to hold for $\text{SGC}_{\text{INTERNAL}}$, though. First, the control is not fixed; second, GP is an adjunct; third, the status as complement of the infinitive phrase in Spanish that satisfies lexical requirements cannot be ruled out. The evidence is inconsistent. Namely, they cannot occupy PSA positions – as seen in (23) – but they can be promoted to PSA in passive structures, like in (24).

(24) *Salir a pescar de noche quiere los niños.
go-out-INF to fish-INF at night wants the kids
'*Going out to fish at night wants the kids.' (Literal)

(25) Comprar un auto usado fue desestimado de plano por Pedro.
buy-INF a car used was dismissed of plan by Pedro
'Buying a used car was entirely dismissed by Pedro.'

Sentence (23) is related to an obligatory and fixed control structure. The actor of the infinitive form *salir* ('go out') is necessarily co-referential with an argument of the matrix verb. The infinitive phrase cannot be interpreted as the PSA of the sentence. In (24) the infinitive phrase *comprar un auto* ('buy a car') occupies the PSA position of the passive form *fue desestimado* ('was dismissed'). Since the characteristic feature of co-subordination is operator sharing at the level of juncture, a piece of argument that suggests that the infinitive phrase is not co-subordinated in Spanish is represented by examples like (26). This sentence shows that the infinitive can be independently modified by an aspectual operator. *Poder* ('can') has scope over *vender* ('sell'); it does not affect the meaning of *querer* ('want').

(26) Pedro quiere poder vender el auto.
Pedro wants can-INF sell-INF the car
'Pedro wants to be able to sell the car.'

In contrast, sentence (27) shows that the modal *puede* ('can') – in its internal possibility/ability interpretation – affects the matrix verb only. The speaker infers that somebody might want to do something. It does not mean that Juan might be eating fish.

(27) Juan puede querer comer pescado.
Juan can want-INF eat-INF fish
'John might want to eat fish.'

It is quite likely the case that this is not possible with every infinitive control template but I cannot devote more space to this discussion. Nevertheless, the point can be made, though; namely that the generalization that prevents control from taken place in subordinate (i.e. complement) instances should be weakened in Spanish by making it sensitive to lexical properties (i.e. some specific verbs). Therefore, it is possible for control to occur in templates with subordinate cores and, provisionally, we can claim that $\text{SGC}_{\text{INTERNAL}}$ involves a subordinate nexus; in fact, it involves an ad-core subordinate nexus rather than daughter subordination.

There are major differences between the two direct subtypes of $\text{SGC}_{\text{INTERNAL}}$. First, $\text{SGC}_{\text{INTERNAL/DM}}$ permits the controlled argument of the non-matrix verb to be expressed as a *wh*-word in the pre-core slot of the matrix verb, as shown in (28). In contrast, example (29) illustrates that the same structure is not possible with the non-DM subtype.

(28) ¿Qué llegó diciendo?
what arrive-PRET.3SG saying
'What did s/he come here saying?'

(29) *¿Qué asustó Pedro a los niños gritando?
what scare-PRET.3SG Pedro to the children screaming
*What did Pedro scare the children by screaming? (Literal)

This data points unquestionably to a tighter juncture-nexus in DM than in non-DM. The grammaticality of (28) can be derived if GP and the matrix verb are joined at the matrix verb nucleus so that there is a single shared pre-core slot where a *wh*- question word denoting an argument of any of the two predicates can be placed. That is, DM needs to have GP in a core internal position within the matrix clause assuming that only core internal arguments are allowed to appear in the pre-core slot of the relevant core. This requirement is met by a nuclear juncture (which has been shown not to be the case) or by a core co-subordination nexus or by an ad-nuclear juncture. This last one is the right choice given the upcoming evidence. In contrast, (29) shows that non-DM involves a different structure; namely, a linkage where the arguments of GP are not core internal to the matrix clause but core peripheral.

Cleft sentences give us additional evidence that further support the assignment of different constituent structures to the DM and non-DM subtypes. In (30) the undergoer of GP can show up in the matrix clause in a DM instance whereas example (31) shows that it cannot in non-DM examples.

(30) Fue un tango lo que Juan entró cantando.
was a tango it.NEUT that Juan enter-PRET.3SG singing
'?It was a tango what Juan came home singing.' (Literal)

(31) *Fue un tango lo que Juan asustó a los niños cantando.
was a tango it.NEUT that Juan scare-PRET.3SG to the children singing
'*It was a tango what Juan scared the children singing' (Literal)

Spanish cleft sentences permit an argument of the cleft clause to be expressed as a *wh*-pronominal neutral form (*lo que* 'what') that is co-referential with a full NP in the matrix clause (i.e. *un tango* 'a tango'). According to the analysis of '*it*-clefts' in Pavey (2004: 206), the cleft clause occupies the periphery of the matrix core. The *wh*-pronoun is set in the pre-core slot of the matrix verb. Hence, as with *wh*-words, DM allows an argument of GP to be place in the matrix pre-core slot whereas non-DM outlaws this placement.

The two structures that I have used so far to differentiate DM and non-DM instances – i.e. *wh*-placement and cleft sentences – involve the same kind of syntagmatic relation; namely, a core argument of GP occupies the pre-core slot of a matrix clause. What are syntactic templates of DM and non-DM that both allow and prevent this position of a core internal argument of GP, respectively? The difference cannot lie on the nexus since both structures are peripheral and, hence, they are supposed to involve subordination. The issue should be handled at the juncture. My proposal is that DM contains an ad-nucleus juncture and a subordination nexus. In contrast, non-DM involves an ad-core juncture and a subordination nexus. If this is right, the *wh*-placement in (28) and the pronominal relation in (30) can be derived from the fact that the arguments of GP in DM are all core internal of the matrix core. This core internal argument restriction prevents the arguments of GP in non-DM from showing up in the matrix pre-core slot since they are indeed core peripheral. The assignment of these templates also predicts that any core internal constituent should be able to be associated with a template that contains an argument of the embedded verb in the matrix pre-core slot. This is the case with complement clauses and infinitives in Spanish. Both structures contain the arguments of the embedded verb within the matrix core and, hence, should permit their placement in the matrix pre-core slot as it is corroborated by sentences (31) and (32) in relation to *wh*-placement and sentences (33) and (34) regarding cleft sentences.

(32) ¿Qué dijo Juan que María iba a traer?
what say-PRET.3SG Juan that María was to bring-INF
'What did Juan say that María was going to bring?' (Literal)

(33) ¿Qué quiere comer Pedro?
what wants eat-INF Pedro
'What does Pedro want to eat?' (Literal)

(34) Era un vino lo que Pedro dijo que María iba a traer.
was a wine it-NEUT that Pedro say-PRET.3SG that María was to bring-INF
'It was a wine what Pedro said that María was going to bring.'

(35) Era un asado lo que Pedro quería comer.
was a barbecue it.NEUT that Pedro wanted eat-INF
'It was a barbecue what Pedro wanted to eat.' (Literal)

Furthermore, the ad-nuclear juncture of DM – and, hence, the core internal status of GP – prevents it from splitting from the matrix verb as shown in (36). In contrast, since non-DM contains GP in the periphery of the matrix verb, it can be placed in the matrix pre-core slot and, hence, separated from the matrix verb, which is shown in (37).

(36) *Fue corriendo que Pedro entró.
was running that Pedro enter-PRET.3SG
'*It was (by) running that Pedro entered (the room).' (Literal)

(37) Fue ahorrando que Juan pudo construir la casa.
was saving that Juan can-PRET.3SG build-INF the house
'It was by saving that Juan could build the house.'

In addition, the matrix verb cannot be placed in the matrix clause of a pseudo-cleft structure in DM but it can in non-DM examples, as illustrated by (38) and (39), respectively.

(38) *Fue entrar lo que Pedro hizo corriendo.
was enter-INF it.NEUT that Pedro did running
'*It was entering what Pedro did running.' (Literal)

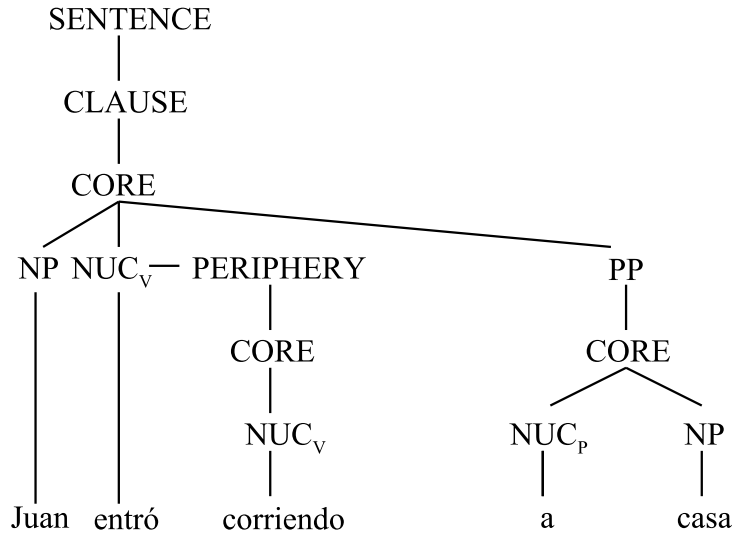
(39) Fue construir la casa lo que Pedro hizo ahorrando.
was build-INF the house it.NEUT that Pedro did saving
'*It was building the house what Pedro did saving.' (Literal)

The last two set of examples unquestionably indicates that the two verb structures cannot split in DM but they can in non-DM. This behavior is expected from the properties of their respective linkages. The ad-nuclear juncture of DM prevents the two verbs from separating; in contrast, the ad-core juncture of non-DM does permit the separation of the verbs. In fact, the core internal position of GP in DM suffices to derive the ungrammaticality of (38) as seen by the ungrammaticality of infinitive and complement clauses in (40) and (41), respectively, in comparable structures. The relation of the matrix verb to the embedded structure cannot withstand splitting and this is also the case with an ad-nuclear structure.

(40) *Era prometer lo que Pedro hizo lavar el auto.
was promise-INF it-ACC that Pedro made wash-INF the car
'*It was promise what Pedro did wash the car.' (Literal)

(41) *Era que dijo lo que María hizo traer.
was that say-PRET.3SG it-ACC that María did bring-INF'
'*It was that said what María did bring.' (Literal)

Figure 3 represents the ad-nuclear subordination juncture-nexus of the DM type whereas Figure 4 displays the ad-core subordination embodied by the non-DM type.



'Juan ran into the house'

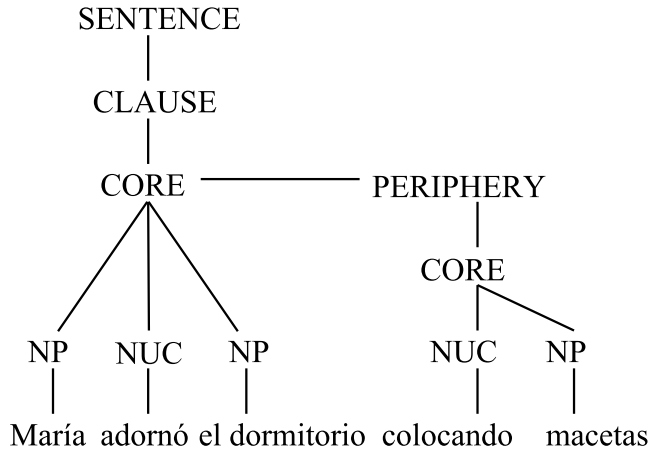
Fig. 3: DM type – ad-nuclear subordination

The contrast between DM and non-DM when combined with depictive secondary predicates should give us further evidence of the structural differences. Depictive predicates in Spanish are – just like the English ones (Rothstein 2000) – obligatory non-fixed control structures, as exemplified in (42).

(42) Pedro trajo al niño cansado.
Pedro bring-PRET.3SG to-the child tired
'Pedro brought the child tired.'

The secondary predicate *cansado* ('tired') might be predicated of either the subject *Pedro* or the undergoer argument *niño* ('child'). Since a control relation is involved, we need to assume that the predicates hold a core or a nuclear juncture connection. The obligatory agreement relation controlled by the shared argument shown in (43) reinforces the nuclear/core juncture analysis.

(43) Pedro presentó su propuesta nervioso/*nerviosa.
Pedro introduce-PRET.3SG his proposal nervous-M/*nervous-F
'Pedro presented his proposal restless.'



'María decorated the bedroom by placing flowerpots'

Fig. 4: non-DM type – ad-core subordination

In addition, there is no semantic relation between the event denoted by the matrix verb and the stage-state denoted by the adjective other than simultaneity and participant sharing. Therefore, the secondary predicate should be interpreted as an adjunct and this means that a subordinate nexus is involved. This is consistent with the independent behavior of operators. As seen in (44) the negation operator has only scope over GP but not on the matrix verb.

(44) Pedro no llegó cansado.
Pedro not arrive-PRET.3SG tired-M
'Pedro did not arrived tired.'

The unmarked interpretation of (44) is that Pedro arrived but he was not tired; the negative adverb introduces an operator that has scope only over the predicative AP. Taking together the evidence presented so far on control, agreement and scope of negation into consideration the assignment of ad-core subordination seems to be a safe assumption. In addition, the semantics of the structure is consistent with this juncture rather than with a nuclear one since the AP really modifies primarily a participant (*Pedro* in (44)) and only derivatively it can be inferred that the event itself is modified. However, this should count as a pragmatic enrichment rather than as the meaning of the sentence.

Sentence (45) shows the combination of $\text{SGC}_{\text{INTERNAL/DM}}$ with a depictive secondary predicate. The grammaticality of this seems to depend on the condition that the AP of the secondary predicate specifies the gerund event and, in this sense, it is highly restricted.

(45) Pedro llegó a casa gritando furioso.
Pedro arrive-PRET.3SG to house screaming furious
'*Pedro came home screaming furious.' (Literal)

If GP and AP interchange positions, the result is unacceptable – or, at least, marginal – as shown in (46).

(46) *Pedro llegó a casa furioso gritando.
Pedro arrive-PRET.3SG to house furious screaming
'*Pedro got home furious screaming.' (Literal)

I propose an analysis of (45) where the AP adds up information to the one contributed by GP and, hence, it needs to be consistent with it, namely, it denotes a participant state that typically co-occurs with the event described by GP. Being modifiers of different layers, they can co-exist in the same sentence if the information provided by the more external modifier (i.e. AP) is consistent with the one offered by the more internal one (i.e. GP). This analysis would prevent the lawful combination of GPs in the non-DM type with APs as secondary predicates. In contrast, the combination of non-DM and the AP is ruled out because they modify the same layer and, hence, this makes difficult to establish a modification relation between them. These are two different forms (i.e. a verb form and an adjective form) and, hence, cannot simply coordinate. In fact, the combination of an adverb – or a form with an adverbial function like GP in Spanish – with an adjective phrase results typically in the adverb modifying the adjective. This is the opposite of what we see here in (45) and (46). Thus, the only factor we can claim makes (45) possible is juncture, more precisely, two junctures at different layers. This predicts the ungrammaticality of the co-occurrence of non-DM and secondary predication as seen in (47).

(47) *Pedro construyó la casa ahorrando apurado/cansado.
Pedro build-PRET.3SG the house saving rushed/tired
'*Pedro built the house by saving rushed/tired.' (Literal)

Is there any semantic difference between DM and non-DM that would motivate their syntactic divergence? In this paper I have not devoted any section to a semantic analysis but I have assumed that the relations between the events are basically the same in both types; namely, both contains either Means or Incidental relations; both are different kinds of manner relations (París 2003, 2006b; París & Koenig 2003). There is, though, a visible semantic difference between them and it comes from the logical structure of the matrix predicates. DM contains a directed motion matrix verb while non-DM involves a matrix verb of any other semantic class. I assume that ultimately this is the differentiating factor that calls for linking into different syntactic templates. If this claim is right, the interclausal relations hierarchy has to be thought to be not only sensitive to semantic relations between events but also to the logical structure type of the matrix verb. Directed motion verbs as matrix verbs motivate a specific type of

linking even if the semantic relation between the events is of the same kind than the one invoke by other matrix verbs.

Is this a Spanish peculiarity? It turns out that different languages single out achievement and directed motion verbs to give them a special treatment. According to Foley & Olsen (1985), DM verbs have a privilege status to form serial verb construction. There is an implicative hierarchy for serial verb formation: if a language has serial verb constructions, they are formed necessarily by telic directional motion verbs ('come' and 'go'); only then, they might have posture verbs ('sit', 'lay', 'stand') and, then, they may also have state and process verbs. In fact, English and Norwegian seems to have counterparts of $\text{SGC}_{\text{INTERNAL/DM}}$. In fact, data of *wh*-placement of an argument of the gerund form verbs as shown in (48) and (49), respectively, should lead us to think of an ad-nuclear subordination linkage (examples taken from Truswell 2007). Both sentences have the undergoer argument of the gerund or gerundive form in the pre-core slot of the matrix verb.

(48) What did John arrive whistling?

(49) Hvilken sang kom han plystrende på?
Which song came he whistling on
'Which song did he arrive whistling?'

There is a second consequence for the syntax-semantic interface. The relation means and incidental are quite different in one important respect; namely, means involves two events in the same causal course whereas incidental relates two events in different causal courses. This difference is neutralized in both types by the fact that the related events constitute a complex macroevent in which e_G is a proper part of e_M. Thus, the part-whole relation does not require a unique causal course while demanding basically participant sharing and temporal overlap. In sum, the part-whole relation is more significant to the interclausal relations hierarchy than the specificities of type of manner relation involved.

4 Conclusions

The salient issue analyzed here has been the contrast between the major subtypes under $\text{SGC}_{\text{INTERNAL}}$, namely DM and non-DM. The evidence presented is compelling in signaling that each subtype involves a different kind of linkage. Both are peripheral constituents and, hence, both are deemed to involve a subordination nexus. This seems like the natural conclusion because it is unexpected of an adverbial adjunct to be on an equal syntactic status with its modified constituents. This would be exactly the situation for any non-subordinate nexus, namely, coordination and co-subordination. More important, I have provided throughout the paper significant syntactic evidence that supports a subordination nexus. In addition,

it is apparent that the juncture is tighter in DM than in non-DM. The two predicates cannot be separated in a cleft structure in DM whereas they can in non-DM; DM allows a core macrorole argument of GP to show up in the pre-core slot of the matrix clause whereas this is not possible in non-DM.

This analysis of DM has consequences for our understanding of the interclausal relations hierarchy. The first conclusion is that an adverbial modifier that holds a manner relation to the matrix verb might behave in many respects like a complement. The semantic distinction between arguments (lexically required participants) and optional semantic information is not equally reflected by the syntactic realization of these meaning components. In particular, there are some kinds of adjuncts that are closer to arguments regarding their syntactic realization than others. DM is precisely the kind of adjunct that behaves syntactically tightly attached to the matrix verb just like arguments do. This fact seems central for RRG since the linking algorithm projects lexically required information into the core whereas non-lexically required occupies the periphery of the core. The point is that DM teaches us that this crucial distinction should not be bluntly projected to the entire syntactic distribution of adjuncts.

An important corollary of this thesis refers to the variability of the syntactic realization of adjuncts; that is, adjuncts are characterized by variable linking. In SGC the same form – i.e. GP – holding basically the same structural meaning relation – i.e. GP adverbially modifies the matrix verb – manages to show up in various and quite different syntactic templates while holding also different semantic relations between the referred events.

A second corollary is that the syntactic realization of adjuncts might be highly sensitive to the semantic class of the matrix verb. In particular, the difference in the syntactic realization of DM and non-DM meanings – which is quite significant – can be motivated only on the semantic class of the matrix verb. The difference is not about the relation between the denoted events, which is manner and it is the same for the two subtypes. It is neither about the meaning of the verb within GP. It is just the fact that DM contains a directed motion matrix verb and this represents a very small set of verbs and, hence, a minor semantic class. The fact that linking is sensitive to this nuance tells us about the nature of linking and this should have an impact on our understanding of it.

References

Foley, William & Olson, Mike. 1985. Clausehood and verb serialization. In *Grammar inside and outside the clause*, J. Nichols & A. Woodbury (ed.), 17–60. London: CUP.

Kawachi, Kazuhiro. 2005. Japanese Enter/Exit verbs are motion verbs. In *Proceedings of the 41st Regional Meeting of the Chicago Linguistic Society*, R. Edwards, P. Midtlyng, C. Sprague & K. Stensrud (eds), 223–239. Chicago: CLS.

Kita, Sotaro. 1999. Japanese Enter/Exit verbs without motion. *Studies in Language* 23(2): 307–330.

Krifka, Manfred. 1998. The Origins of Telicity. In *Events and Grammar*, S. Rothstein (ed.), 197–235. London: Kluwer Academic Publishers.

París, Luis. 2012. Sintaxis y semántica de la construcción causativa. In *Avances en la Gramática del Rol y la Referencia*, R. Mairal, L. Guerrero & C. Vergara (eds), 287–307. Madrid: Editorial Akal.

—. 2008. On Expanditures. In *Proceedings of the 40th Regional Meeting of the Chicago Linguistic Society.* N. Adams, A. Cooper, F. Parrill & T. Wier (eds), 291–304. Chicago: CLS.

—. 2006a. Implicating and Focusing on Underspecified Lexical Information. In *When Semantics Meets Pragmatics.* K. von Heusinger & K. Turner (eds), 421–444. Oxford: Elsevier.

—. 2006b. La codificación gramatical de eventos: Medio_{E}. *Signos* 39: 259–262.

—. 2003. The grammatical encoding of event relations: Gerund Phrases in Spanish. PhD dissertation, State University of New York at Buffalo.

París, Luis & Koenig, Jean Pierre (2003) 'What does it mean to be a dependent?' In *Proceedings of the 10th International Conference on Head Driven Phrase Structure Grammar*, S. Müller (ed.), 298–317. Stanford: CSLI. http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2003/

Pavey, Emma. 2004. The English It-cleft construction: a Role and Reference Grammar analysis. PhD dissertation, University of Sussex. [Available on RRG website]

Rothstein, Susan. 2000. Secondary Predication and Aspectual Structure. In *Approaching the grammar of adjuncts: Proceedings of the Oslo Conference* [ZAS Papers in Linguistics 17], C. Fabricius-Hansen, E. Lang & Claudia Maienborn (eds). 241–264. Berlin: ZAS. http://www.zas.gwz-berlin.de/fileadmin/material/ZASPiL_Volltexte/zp17/zaspil17-rothstein.pdf

Stump, Gregory. 1985. *The Semantic Variability of Absolute Constructions.* Dordrecht: Kluwer Academic Publishers.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics.* Cambridge, MA: MIT Press.

Truswell, Robert. 2007. Extraction from Adjuncts and the Structure of Events. *Lingua* 117: 1355–1377.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface.* Cambridge: CUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax. Structure, meaning and function.* Cambridge: CUP.

On the Application of Conceptual Graphs in RRG – First Steps towards a Functional Computational Processing Model

Judith Gottschalk[α] & Nicolai Winther-Nielsen[β][1]
Aalborg University[α]
Fjellhaug International University College Denmark[β]

Abstract

How can computer-supported vocabulary learning be improved by using a computational adequate model of Role and Reference Grammar [RRG] (cf. Van Valin 2005)? This paper suggests conceptual graphs (cf. Sowa 2000) as a new approach semantics in RRG to develop a computationally tractable version of RRG, which can be implemented in a semantic parser as an extension of existing learning software for Biblical Hebrew. In this approach the linking algorithm from syntax to semantics is reduced to a set of lexical rules matching attribute value matrices defining the layered structure of the clause against an ontology (cf. Gottschalk 2010; Gottschalk 2012a; Gottschalk 2014); an algorithm for the automatic determination of *Aktionsarten* for Biblical Hebrew developed by Winther-Nielsen (2017) is employed for this task.

1 Introduction

When learning Biblical Hebrew, what causes most learners big problems is to remember and to correctly apply the huge amount of vocabulary. Many lexemes have several different meanings and when translating from the Hebrew Bible it is nearly impossible for the students to choose the meaning which correctly fits in the context. Students have various strategies to learn their vocabulary: Some write hundreds of record cards by hand with the Hebrew lexeme on the front side and the translations in their native language on the backside so that they can use them to recall the foreign words, others use technical means. Computer programs like Anki (http://ankisrs.net), where one can generate digital record cards, called flash cards, are

[1] We would like to thank Hagen Langer, Brian Nolan, Peter Øhrstrøm, Ulrik Sandborg-Petersen and Claus Tøndering for discussions on earlier drafts of this paper.

commonly used or a webpage called Memrise (http://www.memrise.com) is consulted. Here the students can learn their vocabulary in a playful way.

In the particular approach to vocabulary learning developed by Thompson (2011), vocabulary-learning strategies are used as a means to enable vocabulary learning in a learner. For Thompson the leading question is what vocabulary learning strategies are and which vocabulary learning strategies would be most helpful for learning an amount of Biblical Hebrew vocabulary which is suitable for reading the Old testament? He asks the question of what vocabulary learning strategies are, because they are not clearly defined in the literature. Thompson explains that there is no consensus in the applied linguistics literature as to an adequate definition of learning strategies. Nation (2001: 217) for example states that: "It is not easy to arrive at a definition of what a strategy is . . . ", also Schmitt (1997), who did some research on vocabulary acquisition and vocabulary learning strategies does not propose an explicit definition although he discusses possibilities for vocabulary learning strategies in great detail (cf. Thompson 2011: 24). Thompson himself lists the following broad categories for strategies of vocabulary learning:

- strategies involving authentic language use
- strategies involving creative activities
- strategies used for self-motivation
- strategies used to create mental linkages
- memory strategies
- visual/auditory strategies
  (cf. Thompson 2011: 65)

For Thompson, all these strategies are means to enable a learner to acquire vocabulary. A strategy in this context is something like a technique, either mental, physical or audio/visual to make learning vocabulary easier for a learner. Several studies have shown that using computers for vocabulary learning has the advantage of learning vocabulary in context. The problem with learning single words, like many students do, when they use word lists or flash cards, is, as pointed out above, that it is not successful because the learners miss the authentic contexts in which the vocabulary occurs (cf. Thompson 2011, Nagy 1997).

Thompson cites in particular the study by Kang (1995), which suggests that computer-supported vocabulary learning is a promising approach. Kang tested the effects of using a computer program to introduce words in genuine contexts and found out that presenting words along with both written and visual contexts was better for definition recall, listening comprehension, and knowledge transfer than paired-associate learning with paper and pencil, with a

computer having words but no picture, or with a computer having only a picture (1995: 43) (cf. Thompson 2011).

Like Kang (1995) we present an approach to learn Biblical Hebrew vocabulary, which uses a corpus-driven tool called Bible Online Learner (http://bibleol.3bmoodle.dk). Bible Online Learner uses the corpus of Biblical Hebrew, which is provided by the Eep Talstra Center for Bible and Computer [ETCBC] in Amsterdam. Claus Tøndering has programmed Bible Online Learner in continuation of the EuroPLOT project (http://www.eplot.eu) in close collaboration with Nicolai Winther-Nielsen, who developed Bible Online Learner's persuasive design (cf. also Gottschalk & Winter-Nielsen 2013).

In the course of the extensive testing of Bible Online Learner, learners found on the one hand that corpus-driven language learning can be very persuasive but that on the other hand effective learning requires the knowledge of lexical data that is not available in the corpus (cf. Winther-Nielsen 2017). The problem laid out in the beginning made using Bible Online Learner not as successful for the learners as one might expect in the first place: Learning the ancient Semitic language of Biblical Hebrew in a corpus-driven technology is not only a matter of being able to read and parse the Hebrew Bible but it also means translating from the Bible, and translating means to have an active vocabulary and not only being able to recall morphological and syntactic information in order to process the language (cf. Winther-Nielsen 2017).

While Thompson (2011) suggests approaches from applied linguistics to improve vocabulary learning, which do not use a computer, like Winther-Nielsen (2017) I come to the conclusion that a stronger linguistic framework is needed to enhance the learning technology, which can be computationally implemented to better support computer-supported vocabulary learning. Thompson (2011) thinks that computers in general do not enhance vocabulary learning given their limited availability. Nevertheless Behringer *et al.* (2013) has shown that computer-supported language learning can even be established in the countries of the Global South, e.g. Madagascar.

Winther-Nielsen explains that using a linguistic framework as an engine for computer-supported language learning will improve corpus-driven language learning as it enables learners to derive lexical rules from the corpus. This way the learners can choose the correct interpretation of a word in their target language of translation. Winther-Nielsen shows that deriving meaning from structural valency patterns, as it is done at the ETCBC (cf. Dyk 2016), is not a sufficient solution as it is not reliable in all contexts. Therefore he suggests a solution within the framework of Role and Reference Grammar (RRG), which does not treat valency patterns as a purely syntactic phenomenon, but rather relies on semantic roles and pragmatic reference (cf. Winther-Nielsen 2017).

In a recent paper, Winther-Nielsen explores what RRG can offer in its current status and expands the theory with a solution using a WordNet ontology (http://wordnet.princeton.edu/), following a proposal made by Petersen (2007) to enrich parser output to generate conceptual graphs [CG] from it (cf. Winther-Nielsen 2017). In the course of his research, Winther-Nielsen develops a pedagogical framework, which enables learners of Biblical Hebrew to derive the meaning of a word in a specific syntactic context with the help of a simple algorithm, which uses the theoretical basis of RRG.

In this paper, building on previous work by Winther-Nielsen, I develop the theoretical basis for adding a semantic parser to Bible Online Learner, which will improve the automatic generation of glosses in Bible Online Learner and supports learners in their desire to derive the correct meaning of verbs from the corpus. Section 2 shortly introduces Bible Online Learner and a possible design for the integration of the semantic parser into the tool. The questions the paper deals with are: 1) What does a computationally adequate model of RRG which enables the derivation of meaning from a corpus of Biblical Hebrew, have to look like, and does RRG need to be revised in order to be used in a computational implementation of a semantic parser which can derive meaning from the corpus? 2) How can a lexicalist functional approach to RRG using conceptual graphs support the development of a computationally adequate model of RRG, which can be implemented in a semantic parser?

The paper is organized as follows: Section 2 introduces the background of this study and explains the motivation for using RRG to automatically derive meaning from a corpus. It makes reference to earlier and current computational approaches to semantics in RRG. Section 3 shows how vocabulary learning with Bible Online learner, the tool I seek to improve, works at the moment, and how I seek to improve it with the model developed in this paper. In Section 4 I introduce my new approach to RRG semantics, which is based on work done by Winther-Nielsen (2017), using conceptual graphs, and I develop a new approach to the linking of syntax to semantics in RRG, which enables the derivation of meaning in terms of semantic structures in the ETCBC corpus.

2 Background

RRG uses the term 'linking algorithm'. This use naturally results in associations of implementability on a computational device in terms of a procedurally executable processing model (cf. Gottschalk 2014). The motivation for specifically using RRG as the linguistic engine for the semantic parser lies in its specific architecture. The linking algorithm is bidirectional and it links the semantic representation of a clause with the syntactic representation and vice versa.

This means that it is possible to derive the semantics of a syntactic structure via the linking algorithm. Van Valin (2005: 129) writes about the linking algorithm:

> Viewed in terms of a processing model, the semantics-to-syntax linking is an aspect of the production process, while the syntax-to-semantics linking is an aspect of the comprehension process. In the comprehension process, the parser would take the input and produce a structured syntactic representation of it, identifying the elements of the layered structure of the clause and the cases, adpositions and other grammatically relevant elements. It is then the task of the grammar to map this structure into a semantic representation, as the first step in interpreting it, and this is where the syntax-to-semantics linking algorithm is required. The same syntactic and semantic representations are used in both algorithms. (Van Valin 2005: 129)

It is this strong claim by Van Valin that the linking algorithm in RRG is naturally tied to being a linguistic processing model. Van Valin uses a perspective from the early times of computational linguistics in his architecture of RRG and the linking algorithm, which perceives syntax, semantics, morphology and phonology as different and separated levels. In this system the steps are sequentially executed as in a procedural computer program. This view led to a first attempt in the development of a formal framework of RRG, which can be implemented into a semantic parsing system (cf. Gottschalk 2014).

Parts of RRG have already been implemented computationally. When it comes to syntax, Nolan (2004) presents an approach to a computational lexicon in RRG, which he used in the first implementation of an RRG-based parser. In another account, Guest (2008) developed an RRG-based parser. It uses extensions of a chart parser to parse languages with different degrees of word order flexibility. Guest executes parsing via syntactic templates as suggested in Van Valin (2005) instead of rules. Winther-Nielsen (2009) and Wilson (2009) have developed a corpus-driven tool called Role-Lexicon Module to support linguistic analysis of Biblical Hebrew. This system uses an EMDROS database and implements an active chart parser to generate the layered structure of the clause. RRG has also been successfully used by Nolan and Salem (2009) and Salem (2009) to develop a system for machine translation called UniArab. Here the linking algorithm is employed via an interlingua to bridge the translation from Arab to English. Murtagh (2011) uses RRG for translation purposes. In this case a conversational agent using Irish Sign Language is developed which uses RRG as its linguistic engine. Nolan (2013) presents an approach which used speech acts within an agent-framework to implement RRG. Gottschalk (2012b) presents the design study of an intelligent teaching agent using RRG

as a linguistic engine while Gottschalk (2014) discussed theoretical aspects of the computational implementation of RRG from a broad perspective. In Diedrichsen (2014) an approach to an RRG-based parser for German which can handle free word order is presented.

Besides several approaches to the implementation of syntax in RRG there has been extensive research on computational approaches to semantics in RRG. Winther-Nielsen (2017) gives a rather extensive overview, which I will briefly paraphrase: Mairal-Usón (2003) developed a semantic meta-language, which relies on semantic primitives. The problem with this approach, however, is that it creates endless *ad hoc* lists of logical terms. The meta-language does not have a means to clearly determine what semantic primitives are (cf. Winther-Nielsen 2017). Therefore, in another approach to semantics in RRG, Guest and Mairal-Usón (2005) developed a solution using the mathematical notion of intervals. They tie it to a double ontology of predicates and objects into open-ended sets and use distinctions from fuzzy logic. The result is yet another complex meta-language. Winther-Nielsen (2009) implemented a lexicon for verbs, which employs the Functional-Lexical Module developed by Faber and Mairal (1999). Here the problem is, however, that this lexicon has proved not to be ideal for an ancient language like Biblical Hebrew (cf. Winther-Nielsen 2017).

Periñán-Pascual and Mairal-Usón (2009) have proposed the Lexical Constructional Model. This model uses a multi-purpose lexico-conceptual knowledge base suitable for natural language processing called FunGramKB (cf. Winther-Nielsen 2017). In the FunGramKB aspects of RRG, especially the semantic representations are used to create a knowledge base for natural language processing (cf. Periñán-Pascual and Arcas-Túnez 2010). In the ARTEMIS project (cf. Periñán-Pascual and Arcas-Túnez 2014), various stages in the development of a computational system employing a lexico-semantical knowledgebase are defined. These steps generate the logical structure of sentences and demonstrate that for a computerized model of the RRG syntax-semantics interface constructional schemas should be used (cf. Gottschalk 2014).

The FunGramKB approach, however, does not work for the Hebrew Bible (cf. Winther-Nielsen 2017). A crucial assumption within FunGramKB is that all knowledge can be formalized. The problem with respect to Biblical Hebrew is, however, that for ancient Israel many cultural situations are barely known. The other main problem is that only access to the literary portrayal of figures and events in the ancient documents exist, so that an ontology as necessary for the FunGramKB cannot be developed based on language data from Biblical Hebrew (cf. Winther-Nielsen 2017).

Winther-Nielsen's suggestion in order to overcome this problem is to step back and use a much more widely known and adaptable ontology for the lexicon. WordNet is a lexical database of English, which is quite large. It was developed in the Cognitive Science Labo-

ratory at Princeton University (cf. Fellbaum 1998). Later WordNet was developed into the successful www.wordreference.com from the development of EuroNet (Vossen 2002) (cf. Winther-Nielsen 2017). WordNets are relational databases for linguistic applications; they are freely available and under constant development (cf. Winther-Nielsen 2017).

In a WordNet, lexical meaning is defined by relation to other lexical units in a network of meaningfully related words and concepts; this way semantic relationships between synsets are formed, which form clusters of synonymous words (cf. Winther-Nielsen 2017). Through the lexical systems, which are created by the relation of entities in the ontology, WordNets overcome the challenge of constructing a conceptual knowledge base as in the FunGramKB. Periñán-Pascual's (2013) objection against the WordNet ontology is that most natural language systems fail because they do not have a sufficiently developed deep semantic knowledge base; therefore he rejects the shallow semantics in WordNet (cf. Winther-Nielsen 2017).

Winther-Nielsen (2017) argues against Periñán-Pascual and explains that because WordNet was influenced by Sowa's (2000) development of Conceptual Graphs in many ways, it offers a user-friendly visual representation, and it can still be implemented in computational applications. Petersen (2003, 2004, 2007) has already implemented an approach to derive conceptual graphs from the ETCBC corpus using shallow parsing and a Hebrew-English dictionary. He matches English glosses with the WordNet categories and constructs conceptual graphs, which represent an ontology-guided, syntax-driven and rule-based joining and refinement of the graphs (cf. Winther-Nielsen 2017). Wilson (2009) has implemented this work in the LEX project where he used it for mapping Hebrew syntax into the logical structures of RRG (Winther-Nielsen 2008, Winther-Nielsen 2017).

In this paper, I will, in turn, develop a computational model focusing on lexical rules, which match attribute value matrices defining the layered structure of the clause against an ontology, in order to realize the theoretical model of a semantic parser which generates conceptual graphs as its output. This, as a result, will enable the correction of glosses in the ETCBC corpus used in Bible Online Learner and will support contextualized vocabulary learning with the tool. The work in this paper is an extension of approaches developed in Gottschalk (2012a, 2012b, 2014). It is based on Nolan (2004) and develops a new computational approach to semantics in RRG as well as to the syntax to semantics linking in RRG by employing Winther-Nielsen's (2017) new approach to semantics in RRG. Instead of the Dowty-based semantic representation used in RRG (cf. Van Valin and LaPolla 1997; Van Valin 2005), I am using a semantic representation using conceptual graphs as developed in Winther-Nielsen (2017; cf. Sowa 2000), which will be generated via a revision of Van Valin's approach.

3 Vocabulary Learning with Bible Online Learner

Before approaches to computer-supported vocabulary learning such as those with Bible Online Learner had been established, learners of Biblical Hebrew used comprehensive dictionaries for grammar and translation exercises. Two things were necessary for successful translation with such a dictionary: a) knowledge of paradigms needed to be activated because the lexemes in the dictionary are not inflected b) it was necessary to choose the appropriate meaning of a verse in the Hebrew Bible (cf. Winther-Nielsen 2017).

With Bible Online Learner this task can be automated as the software can display the appropriate meaning of a word in its syntactic context. The downside of this approach, however, is that the students are not forced to use their knowledge of grammar to derive the appropriate meaning of a word in the syntactic context and hence they rarely memorize the correct translation of a word (cf. Winther-Nielsen 2017).

Extensive evaluation by Winther-Nielsen during a course in Copenhagen has shown that better means for learning vocabulary with Bible Online learner are necessary. The current solution in Bible Online Learner is that Hebrew glosses are displayed with their frequency of occurrence in the Hebrew Bible. How this is done for the Hebrew *nātan* (נתן) is shown in Figure 1.

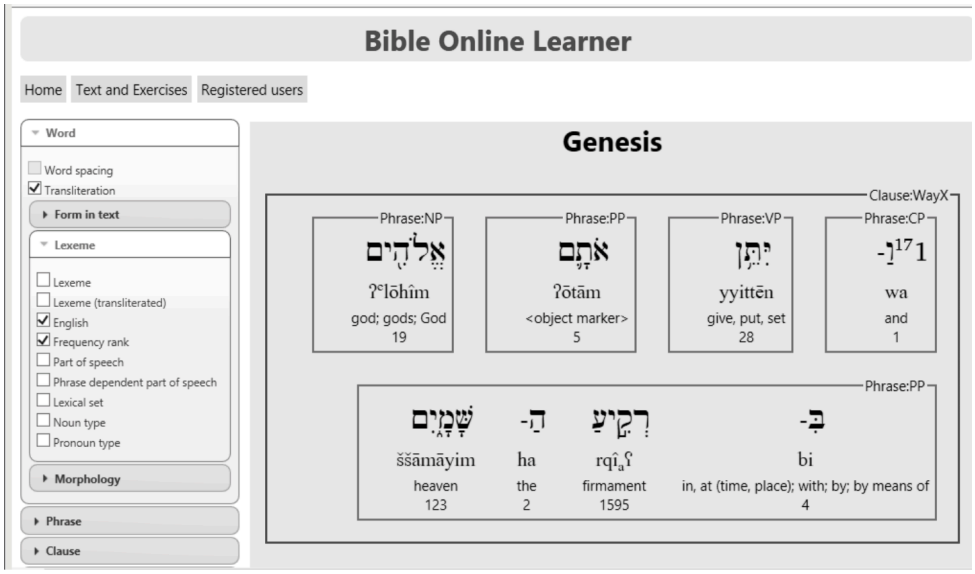


Fig. 1: Display of glosses and frequency in Bible Online Learner

The idea was, that the learners could explore the glosses in the corpus and practice memorization of the meaning of a word in the context of a clause. In the case of *nātan* (נתן) three glosses are displayed: *give*, *put*, *set*. The verb is ranked at position 28 in the list of relative frequency for all words in the Hebrew Bible. In Bible Online Learner, in turn, the learners are

provided with exercises which are based on the frequency of the words. The various lexemes are listed and the learners can input a free translation of the word in a textbox.

The problem with the current form of exercises in Bible Online Learner is, however, that they do not support the learners in choosing between one of the three meanings listed in the glosses in Figure 1 (cf. Winther-Nielsen 2017). Winther-Nielsen illustrates the challenge for the learners as follows:

> To illustrate the challenge for the learner, the verb נתן *nātan* is used 5 times in Genesis 1–3. For the example from Gen 1:17 in Fig. 1, the learner has to deduce from the context that the luminaries are the referents for the masculine plural clitic pronoun attached as suffix to the object marker אֵת *ʔet*. To 'give' them 'in the firmament' clearly does not involve an animate receiver, and therefore the two synonymous glosses 'put' and 'set' are more appropriate for placing the luminaries in visible view in the sky. In Genesis 1:29–30 נתן *nātan* has the meaning 'give' and is followed by a receiver 'to you' in a prepositional pronom followed by the gift itself, 'all herbs and trees'. (Winther-Nielsen 2017: 3)

Obviously, the learners can search for the translations in a dictionary and choose the one which might fit the rest of their translation. But this way of hunting for dictionary interpretations and their attempts to match a preferred Bible translation from their memory prevents them from learning lexical rules they can memorize and automate for language acquisition. The current approach does not support the learners in constructing rules as part of the process of language acquisition. This is where the RRG-based semantic parser, for which I am developing the theoretical basis in this paper, comes into play: Instead of providing the learners with several translations in the glosses in Bible Online Learner, the task will be to automatically derive the translation which exactly fits the syntactic context a verb *nātan* (נתן) occurs in, and to identify that in the case of the gloss in Gen 1:17 *put* is a better translation than *give*.

In addition, the exercises in Bible Online Learner will be improved: Instead of providing the learners with multiple lexemes outside of their syntactic context and asking them to enter all possible translations in a textbox, the learners will be provided with text excerpts from verses in the Bible with the lexeme occurring in its syntactic context. Winther-Nielsen (2017) has developed an algorithm, which makes it possible to derive the meaning of a verb from its semantic context and the learners will be 'forced' to follow this algorithm before they can fill in the meaning of a verb in a textbox. Winther-Nielsen's algorithm will be discussed in Section 4 in some detail; for now I will give simple examples of how it can be employed for learning purposes. That means before entering the meaning of the verb in the textbox the learners are asked specific questions which help them to derive the meaning of the word like

'Is the *Aktionsart* of the verb a causative?'. Here the user can choose a radio button with 'Yes' or 'No' and is directed to the next question he has to answer: 'Does the verb have the *Aktionsart* process?'. Again the learner has to click a radio button and is this way directed to the correct translation and forced to follow a clearly defined algorithm to derive the meaning of a word so that he will be able to develop knowledge of lexical rules for the Biblical Hebrew vocabulary.

4 Conceptual graphs in RRG: A formal model towards implementability

The way to go for an implementation of the RRG linking algorithm in order to develop a semantic parser is to design a computationally adequate formal model of RRG. When I am talking about computational adequacy in RRG, I mean that the theory should refer to formal systems which are computationally tractable and can be executed on a computational device in an algorithmic way in a finite time. This concept of computational adequacy supports the levels of explanatory adequacy within a Chomskyan framework (cf. Gottschalk 2014, 2012a). These two levels are interconnected because a theory which is not computationally tractable and which has a Turing complete generative power is, following Carpenter (1991), who investigated HPSG, less explanatory than a theory which is tractable, can be executed on a computational device and has less generative power in the sense of not being Turing complete (cf. Gottschalk 2014, 2012a: 130).

If a linguistic theory is computationally adequate, it can process language with low storage demands without mistakes. A crucial assumption in this context is that computational adequacy is based on the Church-Turing thesis. It claims that everything which is computable on a machine is intuitively computable and vice versa. The consequence of the Church-Turing thesis being bidirectional is that since a natural language is intuitively computable it should also be computable on a machine (cf. Blass and Gurevich 2001). This results in a computer, which can be used as a test bed for linguistic theories in order to show that a linguistic theory actually works.

It is the architecture of RRG, which raises the idea that RRG is a linguistic theory, which can be implemented with ease as it already has a design, which can be translated into a procedurally executable code and therefore can be formalized without problems. However as shown in Gottschalk (2012a, 2014) several problems occur if one attempts to directly translate the linking algorithm in either a pseudo-code meta-language or in machine code. It is a particular step in the linking algorithm, which is problematic. This step is given in (1) below:

(1) Step 2:
Determine the actor and undergoer assignments following the actor-undergoer hierarchy. . . (Van Valin 2005: 136).

In Gottschalk (2012a, 2014), this step was formalized as a program snipped in pseudo-code:

(2) **Algorithm step2**

```
if number_argument_slots = = 1 in logical structure do
   if lexical_entry_verb = = takes_undergoer do
      undergoer = referring_expression_x;
   else
      actor = referring_expression_x;
   end if.
if number_argument_slots = = 2 in logical structure do
   actor = leftmost_argument;
   undergoer = rightmost_argument;
end if.
if number_argument_slots = = 3 in logical structure do
   actor = leftmost_argument:
undergoer = new.choice( );
non_macrorole = new.choice( );
end if.
```

If the number of argument slots in the logical structure equals 1, access to the lexicon is necessary in order to determine whether the lexical entry of the verb suggests that it can be satisfied by an undergoer or by an actor. The default situation is that the verb takes actor as its only macrorole in intransitive verbs with one argument slot. There is only one situation where the AUH can apply as sole basis for the determination of macroroles: When the verb is transitive with two argument slots in the logical structure (cf. (2)). From an RRG perspective, being transitive is the default situation for verbs. This default situation results in many theory internal considerations in which the AUH is an essential part of the theory, since the assignment of macroroles in the lexicon is the marked situation (cf. Van Valin 2005: 66). The crucial situation with respect to the computability of RRG where RRG runs into a difficult problem occurs with three argument slots in the LS. RRG cannot account for three-place predicates solely based on its procedural approach and the AUH (cf. (2)). In cases with variable undergoer linking as in English three-place predicates, the linking algorithm cannot determine to

which of the possible candidates for the undergoer, the corresponding macrorole should be assigned (cf. Gottschalk 2014).

(3) a. [**do'**(Starbuck, ∅)] CAUSE [PROC & INGR **have'**(Apollo, security key)]
b. Starbuck [actor] gave the security key [undergoer] to Apollo
c. Starbuck [actor] gave Apollo [undergoer] the security key

Based on Van Valin's (2007) AUH it is not possible to decide which argument should be assigned undergoer if the computer is supposed to generate a construction as in (3a) or (3b). The reason for this is that the modified AUH leaves the choice between assigning the lowest ranking argument in the LS undergoer as in (3c) or to assign the second lowest argument in the LS undergoer as in (3b). Van Valin does not explain how to solve the problem of variable undergoer linking in RRG (cf. Gottschalk 2014). This is the reason for the use of the function 'choice' in the pseudo-code. All the function states is that it is the speaker's choice of assigning undergoer to one of the arguments in the LS. Van Valin (2007) deals with this situation in that he uses preposition assignment rules. These rules cannot apply to this part of the algorithm, however, since following Van Valin (2005: 136) the assignment of prepositions takes place in step 3 of the linking algorithm. Additionally the AUH developed in Van Valin (2007) leaves a choice with respect to the assignment of three-place predicates.

A possible solution to solve the problem that RRG is not computationally adequate is presented in Gottschalk (2012a, 2014). Here, a lexicalist approach to RRG which does not use semantic macroroles in any way is chosen. This approach rather uses a unification-based approach to RRG by employing inheritance networks and attribute value matrices to apply for the population of logical structures in the mental lexicon. It is possible to extend this approach and develop a computationally adequate model of RRG which employs conceptual graphs [CGs] (Sowa 2000) instead of RRG's Dowty-based LSs. Based on a procedure developed in Petersen (2007) it is possible to develop a computationally adequate RRG linking algorithm which is lexically motivated and can be implemented in a semantic parser.

The point of departure for using conceptual graphs to represent RRG semantics for Winther-Nielsen (2017) is that it is possible to create a formal language, which is accessible for non-linguist learners. The reason why I choose conceptual graphs as possible semantic metalanguage in RRG is based on considerations which come from theoretical computational linguistics and are related to the computability of RRG, an issue I discussed earlier in this section. Conceptual graphs are computable on a computational device and they are applicable to languages like Biblical Hebrew. Hence, they are useful for a semantic parser of Biblical Hebrew.

Conceptual graphs are directed bipartite graphs. They have two kinds of nodes: Concepts and relations. While concepts can stand alone, relations are always connected to at least one

other concept (cf. Petersen 2007; Sowa 2000). Formally conceptual graphs are equivalent to directed acyclic graphs and attribute value matrices and they can be generated from them. Concepts in conceptual graphs consist of a type and a referent. Types can either be taken from an ontology or they consist of an individual to which the concept refers. The description of concepts can either be as simple as an existential quantifier or they can be indexical. In the first case the referent is simply left blank as a short hand and in the latter case it points to a specific referent by established conventions (cf. Petersen 2003). Relations are nodes modifying one or more concept nodes. They can bring two or more concept nodes into relation with each other (cf. Petersen 2003). The relations for RRG suggested in Gottschalk (2012a) are AGNT (agent) and PTNT (patient), which refer to Fillmorian grammatical relations. In my approach to RRG, these relations replace traditional semantic macroroles as suggested by Van Valin (2005). In conceptual graphs, a relation has a relation type, which determines its signature. Such a signature dictates which types the concepts attached to a relation can have (cf. Petersen 2003). Both relations and concepts are semantically defined by an ontology.

When it comes to the current semantic representation in terms of logical structures as they are used in Van Valin (2005), Nolan (2004) has shown that they can be represented as a unification grammar in terms of attribute value matrices [AVM]. Two examples for Nolan's AVMs are given in Figure 2 and Figure 3 below.

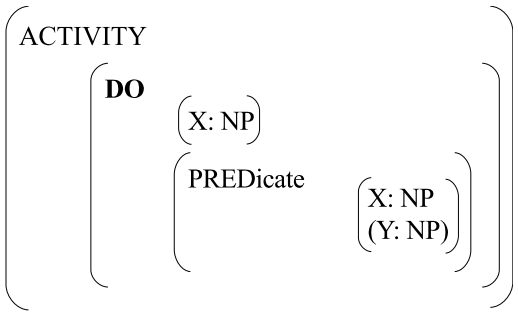


Fig. 2: AVM for Activities

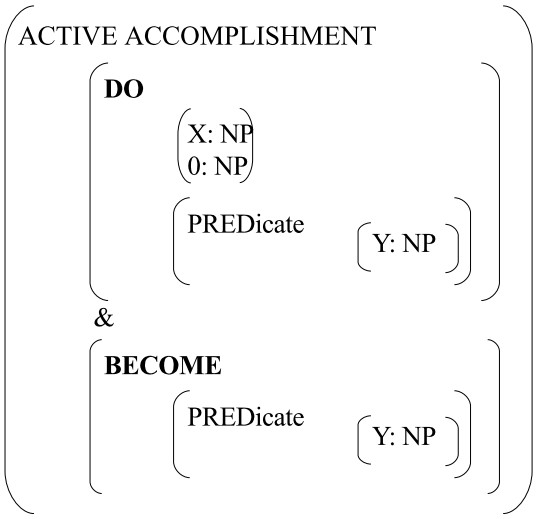


Fig. 3: AVM for Active Accomplishments

Formally AVMs are equivalent to acyclic graphs. Generally, when it comes to the computability of graphs, there are typically three possible variants regarding their complexity: a) practically computational b) belonging to context free and related kinds of grammars with cubic computing time c) computation is not possible. The AVMs representing the semantic structure in RRG can be expressed by the following formal grammar given in (4):

(4) **Non-terminals**
State
Activity
Achievement
Accomplishment
Active Accomplishment
DO
Pred
INGR
Terminals
X
Y
Z

Production Rules

Aktionsart → State
Aktionsart → Activity
Aktionsart → Achievement
Aktionsart → Accomplishment
Aktionsart → Active Accomplishment

Activity → Do
Do → X PRED
PRED → X Y

Active Accomplishment → Do BECOME
Do → X PRED
PRED → X Y
BECOME → PRED
PRED → Y

Accomplishment → Become
BECOME → PRED
PRED → Y

Starting Symbol

Aktionsart

In the example in (4), I focused specifically on a grammar, which generates AVMs for Activities, Active Accomplishments and Accomplishments. The grammar clearly shows that the data structure, which formally describes Van Valin's LSs, AVMs, can be generated by a context-free grammar. Does that mean that the whole theory of RRG is context-free and hence computable? The answer is: No, not necessarily. Certain parts of the grammar need to be expressed by a context-sensitive grammar. This is of course compatible with the model described here, but it causes RRG to have a less optimal complexity within the Chomsky hierarchy. Nevertheless the approach of using Van Valin's LSs even if they are translated into AVMs has the disadvantage that they do not account for ancient languages like Biblical Hebrew as discussed in Winther-Nielsen (2017).

The syntax to semantics linking algorithm presented by Van Valin (2005) is not a fully-fledged algorithm, it is rather a guiding principle based on which an algorithm can be generated, and this is what I will be doing in the next paragraphs.

A first step in the development of this new approach to the syntax to semantics linking, which should be implemented within the semantic parser I seek to develop to improve vocabulary learning in Bible Online Learner, is the design of the architecture of the mental lexicon, which is the heart of this lexicalist approach to RRG. It is given in Figure 4 below.

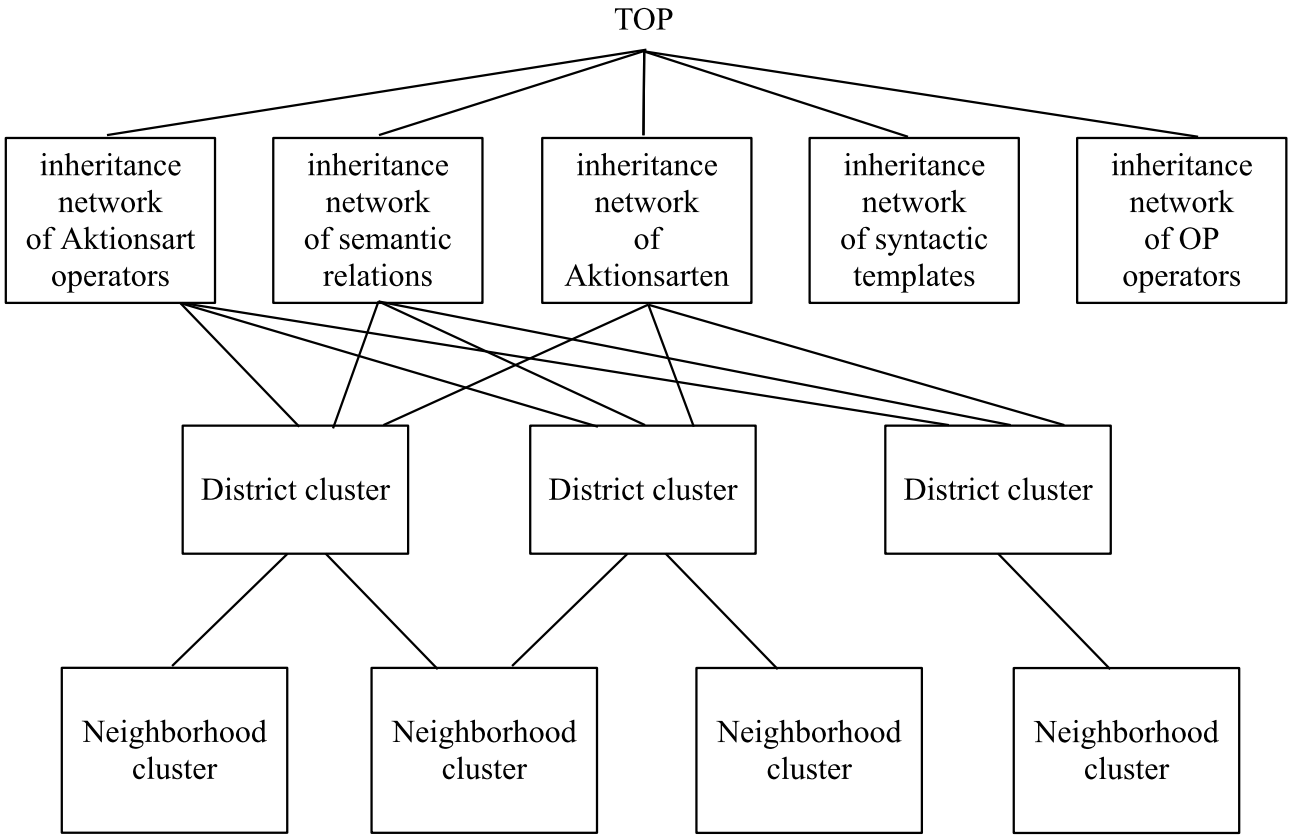


Fig. 4: Architecture of the RRG lexicon

The mental lexicon is organized as an ontology. This will make it possible to generate conceptual graphs from it. The ontology defines types for the conceptual graphs and is stored in the mental lexicon (cf. Figure 4). It is an extension of the various ontologies developed in Gottschalk (2010) and Gottschalk (2012a, 2014). In this ontology, not only the nodes from the LSC are stored and semantically defined by attribute value matrices. Ontologies for thematic relations, which replace macroroles in my approach, and the *Aktionsart* operators are also stored in it. The ontologies for thematic relations, LSC operators, *Aktionsart* operators and *Aktionsarten* themselves make up the upper ontology which is displayed in Figure 4 on page 340. Deeper in the ontology, semantic clusters, in which verbs, nouns and other parts of speech are stored by their semantic meaning, are stored in *district clusters* and *neighborhood clusters*. A district cluster contains words with less semantic content (in the sense of being less specific), like go in comparison to *amble*, while verbs as specific as *amble* are stored in neighborhood clusters.

With respect to computational complexity the choice of an ontology which has the data structure of a tree has the advantage of being searchable with specific search algorithms in reasonable time, so that access to a lexicon, which is structured in this way supports the attempt of developing a computationally adequate model of RRG.

In this model the number of positive and negative features in the AVM defines specific meaning of verbs. An AVM for the verb *receive* is given in Figure 5 below:

receive
< > == [– transfer from x argument in LS]

<selectional properties>	= =	transfer
<manner of transfer>	= =	transfer of a non. abstract entity realized as z from y argument to x argument. Describes a neutral way of transfer to x argument.
<thematic relation>	= =	*give-rel*
<pointer to construction repository>	= =	^[RP_{Agent} V RP_{Theme} {PREPPN | RP]$_{Recipient}$]
		^[RP_{Agent} V RP_{Theme} {PREP indef de N]$_{Recipient}$]

Fig. 5: Lexical Entry for the verb *receive*

In a concrete implementation of the vocabulary learning tool, it would be possible to use WordNet as reference ontology for the district clusters and the neighborhood clusters which are suggested in Gottschalk (2010, 2012a), and thus to generate glosses with the exact meaning of a Hebrew word. This is basically also what Winther-Nielsen (2017) suggests.

Together, these clusters form the ontologies for the conceptual graphs, which constitute the semantic representation in this approach to RRG. In the lexicon, which is crucial to this version of a computational approach to RRG, underspecification and AVM feature structures are employed to represent word meanings in the lexicon. The templates for the LSC are also

AVMs. They are stored in the syntactic inventory which is an extension of the semantic lexicon (cf. Nolan 2004; Van Valin 2005). Nolan (2004) has shown that it is possible to generate an AVM representing the LSC in RRG from the ontology in Figure 4.

The AVM for the LSC is given in Figure 6 below and it is based on the immediate dominance rules developed in Van Valin and LaPolla (1997). These rules are given in (5).

(5) Immediate dominance rules

SENTENCE	→	{(DP)} CLAUSE
DP	→	XP /ADV
CLAUSE	→	{(ECS)}, CORE, (PERIPHERY), {NP*}
ECS	→	XP /ADV
PERIPHERY	→	XP / ADV
CORE	→	ARG*, NUC
NUC	→	PRED
PRED	→	V /XP
ARG	→	PRP / NP / PP

When it comes to syntactic structures as in example (5), this set of rules is, however, generated by a context-sensitive grammar, at least in the case of English. Figure 6 gives an example of an AVM for an LSC.

A similar approach is suggested for the operator projection. Van Valin and LaPolla (1997) have developed ID/LP rules to represent the operator projection. The rules are given in (6).

(6)

SENTENCE	→	CLAUSE	←	IF
CLAUSE ← IF	→	CLAUSE	←	OP*
CLAUSE ← OP	→	CORE (← OP*)		
CORE (← OP*)	→	NUC (← OP*)		
NUC (← OP *)	→	V/XP		

(6) is an example of a context-sensitive part of RRG and it raises the question of how this part can be improved. For now, I will leave it to future research to investigate this question. A representation of the operator projection as an AVM is given in Figure 7.

For English, Van Valin (2005) suggests a set of syntactic templates to generate more complex syntactic templates from it. He assumes that there is a natural ordering for them within the syntactic inventory as part of the lexicon. It is possible to assume a subsumption order based on their structure and their order (cf. Nolan 2004: 214). Instead of Van Valin's approach, I use

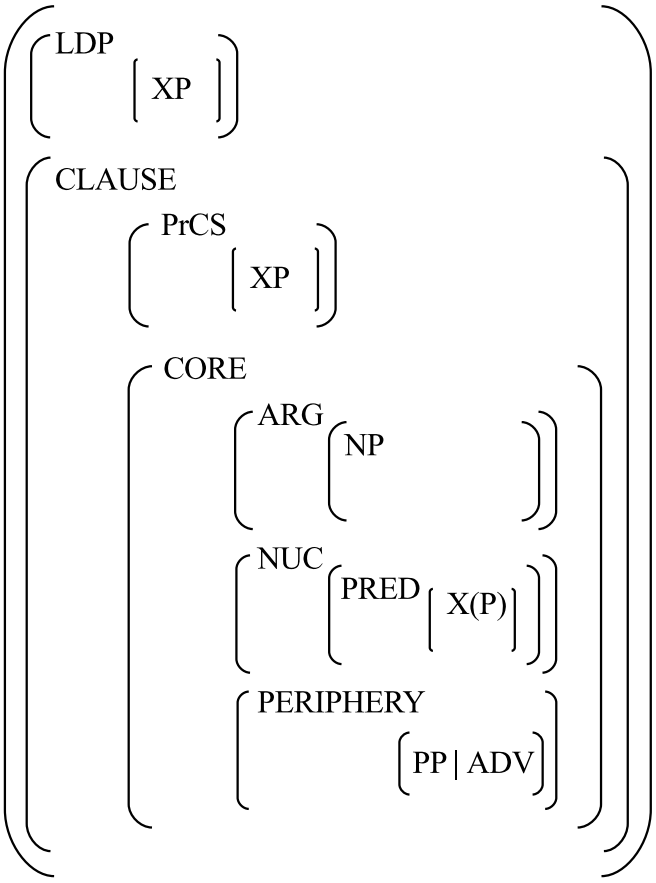


Fig. 6: AVM representing the LSC

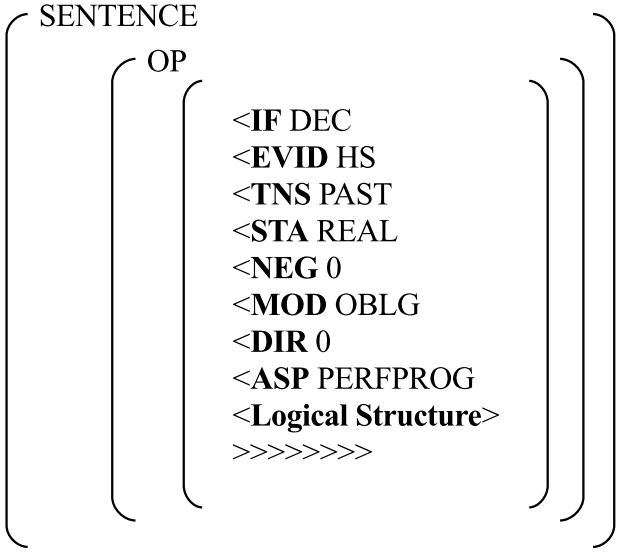


Fig. 7: AVM representing the Operator Projection

AVMs similar to the one in Figure 2 in a hierarchical order, which is attached to the ontology in Figure 1. As in that ontology, more specific AVMs, which are a subset of another AVM, are stored deeper in the ontology because they are attested to be suitable in computational implementations of RRG as shown in Nolan (2004).

The conceptual graphs for the semantic representation of RRG are formally equivalent to the AVMs representing RRG semantics suggested in Nolan (2004). The advantage of using conceptual graphs, as already mentioned, is, that they are more general data structures to represent the semantics of a text – compared to FunGramKB – as they can also be applied to ancient languages like Biblical Hebrew. Furthermore, they are already attested to be computationally derivable with the help of a computational device. An approach to generate them via a computer has been presented in Petersen (2007). I employ Petersen's approach to generate them to a limited degree in order to establish a computational approach to RRG. The conceptual graphs presented in this approach are similar to the pedagogical ones presented in Winther-Nielsen (2017). Conceptual graphs for all RRG *Aktionsarten* are given in (7) below:

(7) a. [Person: Mulder]← (Expr)← [**do'**: drive]→ (Pat)→ [vehicle: car]
activity

b. [Person: Scully]→ (Agent)→ [INGR: win]→ (Thm)→ [reward: FBI-medall]
achievement

c. [Person: Director Skinner]→ (Expr)→ [SEML: cough]
semelfactive

d. [Substance: Kryptonite]→ (Expr)→ [PROC: glim]
process

e. [Substance: Kryptonite]→ (Pat)→ [PROC & INGR: melt]
accomplishment

f. [Person: Mulder]→ (Expr)← [**do'**& INGR: drive]→ (Loc)→ [Destination: Area 51]
active accomplishment

The formal representation of the conceptual graphs in (7) uses types for concepts, which are stored in the lexicon within the neighborhood clusters and the district clusters. Concepts, in the graphs in (7), are represented in square brackets, while relations are represented in parenthesis. In the examples in (7), types are *person*, *vehicle*, *reward*, *substance* and *destination*. The relations in the parenthesis contain the thematic relations defined in Gottschalk (2012a). *Aktionsart* operators like **do'**, INGR, SEML, PROC, PROC & INGR and **do'**& INGR are stored in the ontology defining the hierarchical structure in the lexicon.

Sowa (1992) explains that conceptual graphs completely express the full semantics of common logic. In common logic, quantifiers can have scope over functions and relations but common logic retains a first-order style of model theory and proof theory. In order to support a higher-order syntax without the computational complexity of higher-order semantics, the CL model uses a single domain D that includes individuals, functions, and relations. According to Sowa, Quine (1954) was the first to suggest that the domain of quantification should be limited to a single set. Also, quantification has been used in various theorem proofers that allow quantifiers to range over relations (cf. Sowa 1992).

Sowa (1992) gives a formal grammar for conceptual graphs within Extended Backus Naur Form rules. Since the Extended Backus Naur From is equivalent to context-free grammars, and is rather a different form of expressing them, I have developed a context-free grammar generating the new conceptual graphs for RRG in (8):

(8)

Conceptual Graph	→	{concept conceptual relation} *
Concept	→	concept type referent
Conceptual relation	→	relation type
Relation type	→	signature

What the grammar rules for the new approach to semantics in RRG show, is that conceptual graphs of the type in (7) can be generated by context free grammars as in (8). They are thus not less optimal than Van Valin's (2005) approach to semantic structures in RRG. The advantage of using conceptual graphs instead of a semantic meta-language as developed in FunGramKB is that conceptual graphs can be used to create a machine-readable lexicon for languages like Biblical Hebrew, and not only for modern languages like English and Spanish. A flow chart of the syntax to semantics linking algorithm employing conceptual graphs rather than logical structures is given in Figure 8. It is based on Petersen (2007).

The AVMs representing the LSC are decomposed as CGs. This is done in two steps: In a first step by transformation rules which match each node in the AVM representing the LSC against concepts in the ontology and generate an intermediate CG. In a second step these intermediate CGs generate a semantic CG. The transformation rules operate on the word level and on the phrase level. Here, the rules are looped over the nodes in the AVM to produce a fragment of a CG. Each node in the AVM is used as input to a rule which has a specifically privileged concept called the attachment point. It is distinguished from all other nodes by having a unary relation called attach. This is the point where the CG fragment is joined with another CG fragment. These rules constitute a new linking algorithm, which now is a completely rule-based procedure which can be executed by a computational device. A table with these rules is given in Tables 1 and 2.

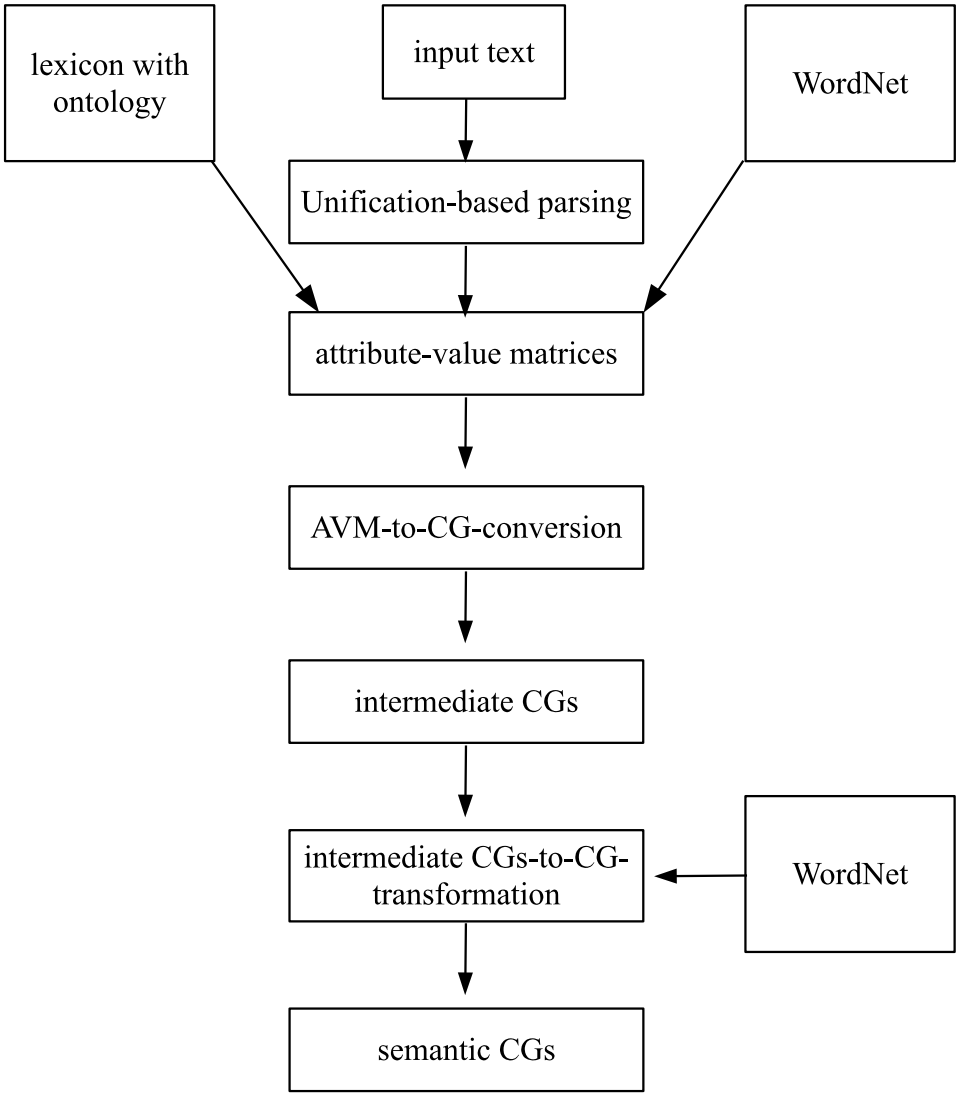


Fig. 8: Workflow for the generation of semantic structures in RRG

		input	*output*
SENTENCE	→	(DP) CLAUSE	(DP) CLAUSE ← attach
DP	→	XP /ADV	XP / ADV ← attach →
CLAUSE	→	{(ECS)}, CORE, (PERIPHERY), {RP*}	(ECS) CORE ← attach (Periphery) NP*
ECS	→	XP / ADV	XP / ADV ← attach
PERIPHERY	→	XP / ADV	XP / ADV ← attach
CORE	→	RP *, NUC	RP ← attach *, NUC ← attach

Table 1: Phrase structure rules

When it comes to the semantic parsing of Biblical Hebrew, it is the output [*verb*] ← attach where Winther-Nielsen's algorithm for the determination of *Aktionsarten* in Biblical Hebrew comes into play. The algorithm is represented in pseudo-code in (9). The input the revised algorithm for the determination of *Aktionsarten* for Biblical Hebrew receives information from, is stored in the WordNet Ontology so that based on that information it is possible to derive the correct meaning of words like נתן (*nātan*) within their morph-syntactic context and generate a gloss with the exact meaning of נתן *nātan* in the glosses in Bible Online Learner.

input	*output*
Noun, singular from operator projection	[*noun*] ← attach
Noun, plural from operator projection	[*noun*: {*}] ← attach
Verb, verbal operator from operator projection is not "be_X"	[*verb*] ← attach
Verb contains operators from lexicon and ontology, is "be_X"	attach → [state: [be]-attr → [*verb*]]

Table 2: Word rules

(9) Winther-Nielsen's algorithm for the determination of *Aktionsarten*

if

Causative:	Causation	C	+ CAUS		→ CAU
elseif					
Process (y):	Accumulated result	U	+ PROC	+TEL	→ ACC
elseif **Incident** (x):	Instantaneous result	U	+ INST	+TEL	→ ACH
elseif **Activity** (x,y):	Completed activity	A	ACTV	+TEL	→ ACTACC
elseif **Semelfactive** (x):	Instant activity	A	+ INST	-TEL	→ SEML
elseif					
Activity (x (y)):	Ongoing activity	A	ACTV		→ ACT
elseif **State** (x,(y)):	Unchanged situation	U	UNCH		→ STA

end

then assign verb-specific semantic roles: is the x A or U and is there a y?

(Winther-Nielsen 2017)

Like in the approach in Gottschalk (2012a, 2014), Winther-Nielsen does not use Van Valin's (2005) semantic macroroles actor and undergoer but rather uses a number of lexical semantic relations, which he derives based on the valence of the verb. This way it is possible to point to constructional schemas as necessary in the way it is described in Gottschalk (2012a, 2014). Winther-Nielsen's algorithm for the determination of the semantic role comes into play in the last step in Table 2. It is given in (10).

(10) Winther-Nielsen's algorithm for the determination of semantic roles

if **Single argument states** → x (U)

11 State or condition:	PATIENT	
12 existence:	ENTITY	

elseif **Two argument non-verbal states** → x, y

26 Possession:	POSSESSOR,	POSSESSED
01 Attributive:	ATTRIBUTANT,	ATTRIBUTE
02 Identificational:	IDENTIFIED,	IDENTITY
03 Specificational:	VARIABLE,	VALUE
04 Equational:	x, y=REFERENT	

elseif **Two argument verbal states** → x (A), y (U)

21 Pure location:	LOCATION,	THEME
22 Perception:	PERCEIVER,	STIMULUS
23 Cognition:	COGNIZER,	CONTENT
24 Desire:	WANTER,	DESIRE
25 Propositional attitude:	JUDGER,	JUDGEMENT
27 Internal experience:	EXPERIENCER,	SENSATION
28 Emotion:	EMOTER	TARGET

elseif **31–35 Single argument activities** → x (A)

31 Unspecified activity:	EFFECTOR
32 Motion:	MOVER
33 Static motion:	STATIC-MOVER
34 Light emission:	L-EMITTER
35 Sound emission:	S-EMITTER

elseif **41–45 One or two arguments activities** → x (A),y (U)

41 Performance:	PERFORMER	PERFORMANCE
42 Consumption:	CONSUMER	CONSUMED
43 Creation:	CREATOR,	CREATION
44 Directed perception:	OBSERVER	STIMULUS
45 Use	USER	IMPLEMENT

end

The intermediate CGs generated by the rules in Tables 1 and 2 generate an intermediate CG of the type: [SENTENCE [CLAUSE [CORE [RP: *Mulder*] , [NUC: *sees*], [RP: *Scully*]]]]. Based on the lexicon and CG-transformation rules, a conceptual graph as given in (8) is generated from the intermediate CG. With the two sets for rules, it is possible to replace the semantics to syntax linking algorithm as developed in Van Valin (2005) by a computationally tractable set of rules in order to develop a computational processing model for RRG. All these changes in the framework of RRG, as it is applied in the semantic parser, which uses RRG as its linguistic engine, lead to a new architecture of RRG. The new architecture of RRG is given in Figure 9.

In this approach to RRG, the lexicon is divided into different parts and consists of a construction repository and a syntactic repository. These two repositories are formally defined by the lexical ontology presented in Figure 4. Complex inheritance processes within the lexicon provide the linking algorithm with the required information, and make a linking from semantics to syntax and from syntax to semantics possible.

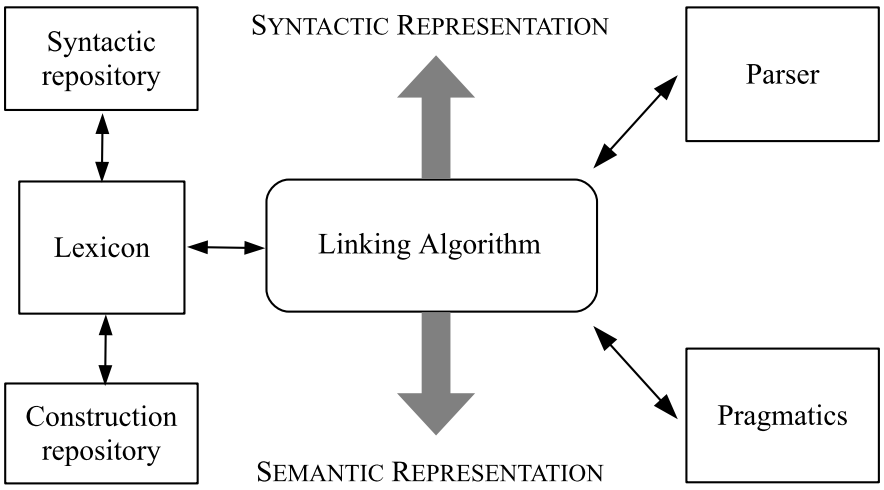


Fig. 9: The new architecture of RRG

5 Conclusion

The paper has presented a novel way to determine the meaning of verbs for learners using Bible Online Learner. With the theoretical model presented in this paper, it will be possible to develop a semantic parser based on a new approach to the syntax to semantics linking algorithm in RRG which will improve corpus-driven vocabulary learning by generating the exact meaning of a specific group of verbs in Biblical Hebrew for the glosses displayed in Bible Online Learner. Furthermore, it was shown that there are parts of RRG which are already computationally optimal, namely the generation of LSs, but which are nevertheless not suitable for ancient languages like Biblical Hebrew.

Based on these findings, it was possible to develop a computationally adequate model of the syntax to semantics linking algorithm within RRG which is also applicable to languages like Biblical Hebrew. This solution works based on a WordNet ontology which Petersen (2007) enriched with conceptual graphs, and which is extended into a computational model of RRG. This will no doubt make it possible to come up with a semantic approach to Biblical Hebrew studies and will improve corpus-driven computer-supported language learning.

References

Behringer, Reinhold, Gram-Hansen, Sandra B., Soosay, Mekala, Mikulecka, Jarka, Smith, Carl, Winther-Nielsen, Nicolai & Herber, Erich. 2013. Persuasive Technology for Learning in Business Context. *International Journal of Information Systems and Engineering*, 1(1): 94–105.

Blass, Andreas & Gurevich, Yuri. 2001. Algorithms: A Quest for Absolute Definitions. http://www.math.lsa.umich.edu/~ablass/abs.pdf

Carpenter, Bob. 1991. The Generative Power of Categorial Grammars and Head-Driven Phrase Structure Grammars with Lexical Rules. *Computational Linguistics* 17(3): 301–314.

Diedrichsen, Elke. 2014. A Role and Reference Grammar parser for German. In *Language Processing and Grammars. The role of functionally oriented computational models*, Brian Nolan & Carlos Periñán-Pascual (eds.). 105–142. Amsterdam: John Benjamins.

Dyk, Janet W. 2016. How do Hebrew Verbs Differ? A Flow Chart of the Differences. In *Contemporary Examinations of Classical Languages (Hebrew, Aramaic, and Syriac): Valency, Lexicography, Grammar, and Manuscripts. Perspectives on Linguistics and Ancient Languages 5*, T. M. Lewis, A. Salvesen, & N. Al-Jeloo (eds), 33–51. Piscataway, NJ.: Gorgias.

Faber, Pamela B. & Mairal-Usón, Ricardo 1999. *Constructing a Lexicon of English Verb*. Berlin: de Gruyter.

Fellbaum, Christiane (ed.). 1998. *Wordnet. An electronic lexical database*. Cambridge, MA: MIT Press.

Gottschalk, Judith. 2010. Storage of linguistic knowledge in the mental lexicon: An approach within Role and Reference Grammar. *ITB Journal* 19: 20–45. http://www.itb.ie/files/journal/issue-19.pdf

—. 2012a. On the computability of Role and Reference Grammar: A new approach to the analysis of three-place predicates and the hierarchical lexicon. *ITB-Journal* 22: 124–192. http://www.itb.ie/files/journal/issue-22.pdf

—. 2012b. The Persuasive Tutor: a BDI Teaching Agent with Role and Reference Grammar Language Interface. *ITB Journal* 23: 31–51. http://www.itb.ie/files/journal/issue-23.pdf

—. 2014. Three-place predicates: A computational approach. In *Language Processing and Grammars. The role of functionally oriented computational models*, Brian Nolan & Carlos Periñán-Pascual (eds), 79–104. Amsterdam: John Benjamins.

Gottschalk, Judith & Winther-Nielsen, Nicolai. 2013. Persuasive skill development: On computational surveillance strategies for modeling learning statistics with PLOTLearner. In *IWEPLET 2013 Proceedings*, R. Behringer & G. Sinclair (eds), 109–116. http://eprints.leedsbeckett.ac.uk/626/1/IWEPLET2013_Proceedings_130904_13933183.pdf

Guest, Elizabeth. 2008. Parsing for Role and Reference Grammar. In *Investigations of the Syntax-Semantics-Pragmatics Interface*, R. D. Van Valin Jr. (ed.). 435–453. Amsterdam: John Benjamins.

Guest, Elizabeth & Mairal, Ricardo. 2005. Lexical representation based on a universal semantic metalanguage. *Revista electrónica de lingüística aplicada* 4: 125–173.

Kang, Sook-Hi. 1995. The Effects of a Context-Embedded Approach to Second-Language Vocabulary Learning. *System* 23(1): 43–55.

Mairal-Usón, Ricardo. 2003. Why the notion of lexical template? *Anglogermánica online* 2002(1). https://www.uv.es/anglogermanica/2002-1/mairal.htm

Murtagh, Irene. 2011. Towards a Linguistically Motivated Irish Sign Language Conversational Avatar. *ITB Journal* 19: 72–101. http://www.itb.ie/files/journal/issue-23.pdf

Nagy, William. 1997. On the Role of Context in First- and Second-Language Vocabulary Learning. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt & M. McCarthy (eds). 64–83. Cambridge: CUP.

Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge: CUP.

Nolan, Brian. 2004. First steps towards a computational RRG. In *RRG 2004 Book of Proceedings. Linguistic theory and practice: description, implementation and processing*, B. Nolan (ed.), 194–223. [Available on RRG website]

—. 2013. Extending a lexicalist functional grammar through speech acts, constructions and conversational software agents. Plenary talk presented at the 2nd *International Conference on Meaning and Knowledge Representation in LCM and FunGramKB*. Facultad de Ciencias Económicas y Empresariales de Granada, Universidad de Granada.

Nolan, Brian & Salem, Yasser. 2009. UNIARAB: A Universal Machine Translator System for Arabic Based on Role and Reference Grammar. [Available on RRG website]

Periñán-Pascual, Carlos. 2013. Towards A Model of Constructional Meaning for Natural Language Understanding. In *The Role of Constructions in a Functional Grammar*, B. Nolan & E. Diedrichsen (eds), 205–230. Amsterdam/Philadelphia: John Benjamins.

Periñán-Pascual, Carlos & Mairal-Usón, Ricardo. 2009. Bringing Role and Reference Grammar to Natural Language Understanding. *Procesamiento del Lenguaje Natural* 43: 265–273.

Periñán-Pascual, Carlos & Arcas-Túnez, Francisco. 2010. The architecture of FunGramKB. In *Proceedings of the 7th International Conference on Language Resources and Evaluation*, N. Calzolari, K. Choukri, B. Maegaard et al. (eds), 667–674. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/LREC2010_Proceedings.zip

—. 2014. The implementation of the CLS Constructor in ARTEMIS. A computational approach. In *Language Processing and Grammars. The role of functionally oriented computational models*, B. Nolan & C. Periñán-Pascual (eds), 165–196. Amsterdam/Philadelphia: John Benjamins.

Petersen, Ulrik. 2003. Automatic Lexicon-based Ontology-creation. A methodological study. Unpublished 9th semester project report. http://people.hum.aau.dk/~ulrikp/pdf/9th-semester-pdf.zip

—. 2004. *Creation in graphs: Extracting conceptual structures from Old Testament texts*. MA thesis, University of Aalborg, Department of Communication. *Impact – an Electronic Journal on Formalisation in Media, Text and Language – Impact Theses*. http://www.impact.aau.dk/theses.html and http://ulrikp.org/MA/

—. 2007. Genesis 1:1–3 in Graphs: Extracting Conceptual Structures from Biblical Hebrew. *SEE-J Hiphil* 4. http://www.hiphil.org/index.php/hiphil/article/view/36

Quine, Willard Van Orman. 1954. Reduction to a dyadic predicate. *Journal of Symbolic Logic* 19(3): 182–184.

Salem, Yasser. 2009. A generic framework for Arabic to English machine translation. Simplex sentences using the Role and Reference Grammar linguistic model. [Available on RRG website]

Schmitt, Norbert. 1997. Vocabulary Learning Strategies. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt & M. McCarthy (eds), 199–227. Cambridge: CUP.

Sowa, John F. 1992. Semantic networks. In *Encyclopedia of Artificial Intelligence* (2nd edition), S.C. Shapiro (ed.), 1493–1511. New York: Wiley.

—. 2000. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA: Brooks/Cole Publishing.

Thompson, Jeremy. 2011. *Learning Biblical Hebrew Vocabulary: Insights from Second Language Vocabulary Acquisition*. PhD thesis, University of Stellenbosch.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

—. 2007. The Role and Reference Grammar Analysis of Three-Place Predicates. *Suvremena Lingvistika* 63(1): 31–64.

Van Valin, Robert D., Jr. & LaPolla, Randy J. (1997). *Syntax: Structure, meaning and function*. Cambridge: CUP.

Vossen, Piek. 2002. WordNet, EuroWordNet and Global WordNet. *Revue française de linguistique appliquée* VII(1): 27–38.

Wilson, Chris. 2009. Lex: a Software Project for Linguists. Technical Report *SEE-J Hiphil* 6. http://www.see-j.net/index.php/hiphil/article/view/60/

Winther-Nielsen, Nicolai. 2008. A Role-Lexical Module (RLM) for Biblical Hebrew. In *Investigations of the Syntax-Semantic-Pragmatics Interface*, R.D. Van Valin Jr. (ed.), 455–478. Amsterdam: John Benjamins

—. 2009. Biblical Hebrew parsing on display: The Role-LexicalModule (RLM) as a tool for Role and Reference Grammar. *SEE-J Hiphil* 6. 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 *Advances in Biblical Hebrew Linguistics: Data, Methods, and Analyses*. A. Moshavi & T. Notarius, 363–385. Winona Lake: Eisenbrauns.

Towards a Formalization of Role and Reference Grammar

Rainer Osswald & Laura Kallmeyer
Heinrich-Heine-Universität Düsseldorf

Abstract

We present an outline of a formalization of Role and Reference Grammar (RRG) which attempts to give full consideration to the overall architecture of RRG as a theory of grammar. The proposed formalization draws a clear distinction between declarative and procedural elements and puts emphasis on syntactic and semantic compositionality. Building on ideas from the theory of Tree Adjoining Grammars, we describe a modular way of specifying syntactic templates and introduce three general modes of syntactic composition: substitution, adjunction, and wrapping. Moreover, we sketch how to formalize the semantic representations of RRG in terms of decompositional frames and briefly discuss the issue of linking in constructional schemas.

1 Introduction

Role and Reference Grammar (RRG) has been developed as a theory of grammar which covers typologically distinct languages and which is able to capture the interaction between syntax, semantics, and pragmatics. The design of RRG was not driven by specific formal considerations. In particular, there is no formal core that plays a crucial role in RRG, as, for example, the theory of feature structures does in Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994). This is not to say that the standard formulation of RRG as presented in Van Valin (2005) has no formal elements. There are the syntactic templates of the syntactic inventory, the Dowty-inspired semantic *Aktionsart* representations, the focus domains of information structure, and, last but not least, the linking algorithm. A *formalized* theory in the sense of the present paper, however, puts stronger emphasis on mathematical and logical rigor than RRG does in its current presentation. For instance, no precise specification of the possible universal and language-specific syntactic templates is given; nor is there a formal definition of how the templates can be combined to larger syntactic structures.

The working typologist, who uses RRG as her or his framework for linguistic analysis, may not regard a thorough formalization as particularly important. In fact, one of the appeals RRG has for field linguistics is that it does not come with an overly heavy theoretical load but keeps a good balance between a rich and elaborate set of notions and explanatory mechanisms, and a semi-formal, intuitive presentation. A formalization can help to identify and eliminate possible gaps and inconsistencies of the theory and, thereby, to improve the theory. In this respect, a formalization has positive impact on every user of RRG irrespective of whether she or he is interested in formal aspects or not. Moreover, a formalization can serve as a basis for computational implementations of RRG. While a thorough formalization may not be absolutely necessary for a computational treatment from an engineering perspective (e.g. Guest 2008), it can contribute to implementations that give full consideration to the overall architecture of RRG as a theory of grammar. The type of implementation we have in mind comes with a grammar development workbench which allows one to test, for instance, whether grammar fragments of language-specific phenomena provide the desired analyses and interact properly with the universal components of RRG.

In the following, we present an outline of how such a formal framework for RRG could look like. The main focus of this article is on the syntactic templates and their composition but we also discuss briefly the formal treatment of the semantic representations and the linking in constructional schemas.

2 The organization of RRG (re)analyzed

The core component of RRG is the bidirectional linking algorithm which captures how syntactic and semantic representations are related to each other (cf. Van Valin 2005, 2010). The two directions of the algorithm, from semantics to syntax and from syntax to semantics, are formulated in a procedural way as sequences of conditional instructions (*if-do* statements) embedded in an elaborate system of case distinctions. Consider the linking from semantics to syntax (for simple sentences), which consists of five main steps. Step one of the algorithm is assumed to build the complete semantic representation of what is to be expressed by the intended utterance. The semantic representation is constructed from the logical structures associated with the selected lexical units (cf. Van Valin 2006: 281), and it also integrates parts of the communicative intentions of the speaker. The second step of the algorithm is concerned with assigning the actor and undergoer roles to arguments of the predicator. In a third step, the morphosyntactic encoding of the arguments is determined, which includes the selection of the privileged syntactic argument, and the assignment of case markers, adpositions, and agreement marking. In step four, the syntactic templates for the sentence in question are se-

lected. The final step anchors the arguments to positions in the syntactic representation of the sentence. A number of alternatives need to be considered here, such as choosing the precore slot, and the case of wh-words needs special treatment.

Kowalski (1979) succinctly characterizes an algorithm as a combination of *logic* and *control*. The logic component is about the representational structures, the compositional relations between them, and the inferential dependencies among properties of them. The control component, by contrast, defines the strategies for sequential processing and for investigating alternatives. Applied to the present context, this distinction means to keep apart the grammatical principles from their application in language production and understanding. From this perspective, the logical, declarative part of RRG is about the principles that constrain the grammatical structures and specify how these structures are related to each other. In which way the principles are put to practice is a question of how, when, and on which grounds they are applied to morphological, syntactic, or semantic information. That is, procedural elements such as assigning grammatical features to linguistic objects or selecting and combining syntactic templates come into play when the principles are applied – be it by the human cognitive system during language production and understanding, given the proposed principles are cognitively relevant, or by a field linguist who consciously employs them for analyzing her or his language data. In the current overall organization of RRG, the principles and constraints of the grammar show up only insofar as they are referred to by the linking algorithm. Examples are the macrorole assignment principle, the syntactic template selection principle, and the completeness constraint. We therefore suggest to concede the principles and constraints a more prominent place in the overall architecture by regarding them as a separate component of the grammar which is independent from procedural aspects.

Another, somewhat related issue of the current architecture is the lack of an integrative perspective on syntactic and semantic composition. The semantics-to-syntax linking algorithm assumes the semantic representation of the utterance to be fully constructed at the very beginning, based solely on the meaning of the selected words and without taking into account any of their morphosyntactic properties. Likewise, the syntax-to-semantics linking algorithm presumes that the parser, which is treated as a separate component in the architecture, is applied to the input before the linking sets in. That is, in the language comprehension process "the parser would take the input and produce a structured syntactic representation of it, identifying the elements of the layered structure of the clause and the cases, adpositions and other grammatically relevant elements in the sentence" (Van Valin 2005: 129). What kind of knowledge about grammatical principles and structures needs the parser to build on, in addition to the input sequence of words, in order to produce the required syntactic representation? Obviously, the parser needs to know about the morphosyntactic properties of the words as well as

about the syntactic templates available for the language in question and the possible modes of composition. From an information processing perspective it is not very compelling to assume that the semantic contributions of the words and the constructional schemas involved are ignored in this step. In fact, the discussion in Van Valin (2006: 283ff) about RRG and language comprehension shows that RRG does not exclude more flexible processing options and, for example, supports the use of complex templates for frequent syntactic-semantic patterns. For these reasons, we propose a view on the processing of linguistic information that allows for a stronger interlacing of syntactic and semantic composition than possible within the current architectural restrictions of RRG.

The program of formalizing of RRG starts with a formal account of the representational structures involved. In Section 3, we will do this in some detail for the syntactic representations, the compositional relations between them, and the templates of the syntactic inventory. We introduce three modes of composition, (simple) substitution, (sister) adjunction, and wrapping (substitution), which seem to be sufficient for compositionally deriving all kinds of complex syntactic constructions. The syntactic templates of the syntactic inventory, on the other hand, will be specified by means of a tree description language. In Section 4, we argue for formalizing the semantic representations of RRG in terms of decompositional frames, which are defined as base-labeled feature structures with types and relations. Section 5 presents a brief case study of how the framework developed so far can be employed for formalizing the linking of syntax and semantics in constructional schemas.

3 Syntactic structure

In RRG it is assumed that the syntactic representations are composed of syntactic templates stored in the syntactic inventory. Figure 1a shows a simple example of a syntactic representation; possible candidates of syntactic templates are shown in Figure 1b.

For a formalization we first need to decide on what kind of formal structures to use. Keeping close to the representations of RRG, we use tree structures for that purpose, where nodes can carry additional features besides the category labels (Section 3.1). The second task is to define the modes by which the syntactic representations are composed from the members of the syntactic inventory (Section 3.2). Next we need to specify the templates available in the inventory, and we need to do this in a way that allows us to capture generalizations among them within and across languages (Section 3.4).

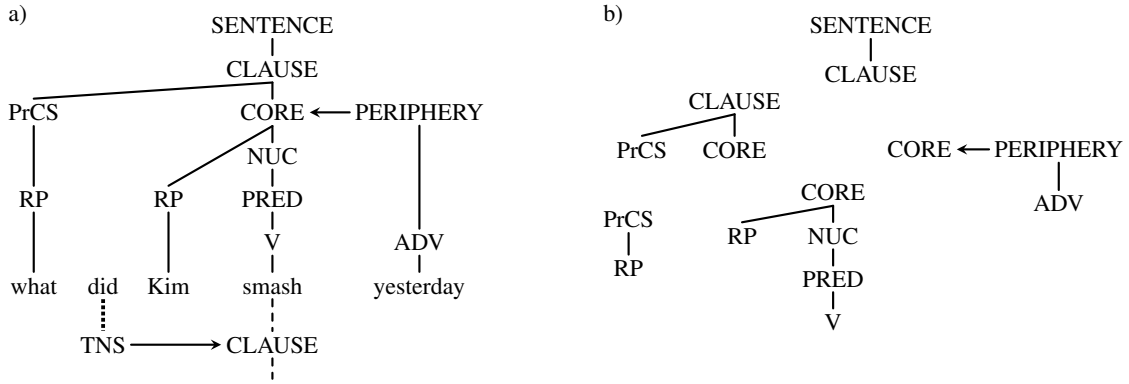


Fig. 1: Examples of syntactic representation and syntactic templates in RRG

3.1 Syntactic representations

A key component of RRG's approach to syntactic analysis is the layered structure of the clause: sentences are assumed to have an internal structural layering consisting of clause, core, and nucleus. The different layers serve as attachment sites for different types of operators: tense operators attach to the clause, modality to the core, aspect to the nucleus, etc. The core level is also the default attachment site for arguments. In the following, we will refer to the subtree of a syntactic representation consisting of the root and its non-peripheral clause, core, nucleus, and predicating descendants as the *clausal skeleton* of the representation.

The syntactic structures in RRG are basically labeled trees, and there are good reasons to use tree structures in a formalization as well. Trees provide the most natural way to analyze syntactic structures since they build on the basic relations *immediate dominance* and *linear precedence*.[1] Having decided on using trees, we need to cope with the more specific aspects of RRG's syntactic presentations such as the operator projection and the arrow notation for peripheral constituents (cf. Figure 1).

Representing the operator structure as a separate projection was originally proposed by Johnson (1987), based on the observation that the ordering among the operators is systematically correlated with their scope given by their attachment site at the clausal skeleton, whereas the surface order of the operators relative to arguments and adjuncts is much less transparent and often requires crossing branches. Johnson's idea was to use two different context free grammars for analyzing the sequence consisting of the verb plus arguments and adjuncts on

[1] There are other options for modeling constituent structures. In HPSG, for instance, constituent structures are modeled by *feature structures*, and Nolan (2004) suggests the same for RRG. But employing feature structures comes at the price of introducing features, or attributes, by which the subconstituents can be accessed. This can either be done by reconstructing tree structures as feature structure based on formal features such as FIRST and REST, or by employing functional notions like SUBJECT, DIRECT-OBJECT, etc. However, in RRG, configurational syntactic notions are *derived* concepts at best and, therefore, should not be part of the underlying modeling.

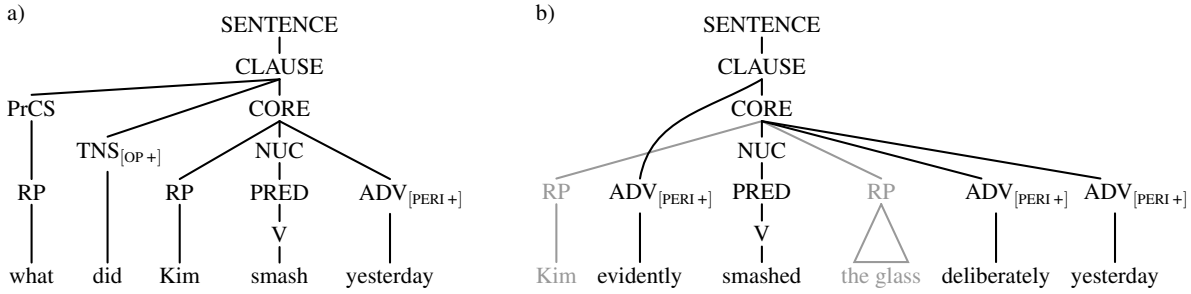


Fig. 2: Periphery and operator marking by features

the one hand, and the sequence of the verb plus operators on the other hand. The two grammars taken together then constitute a *projection grammar*. Hence the names *constituent projection* and *operator projection*. While it seems reasonable to distinguish between constituent structure and operator structure, the proposal of Johnson has the problem of being purely surface oriented. As a consequence, it does not enforce matching clausal skeletons in the two projections. However, corresponding clausal skeletons in both projections are taken for granted in the syntactic representations of RRG. For this reason, our formalization assumes a single syntactic structure in which operator components are distinguished by the feature [OP +] (cf. Figure 2a). We will see below that this representation, together with the approach to operator adjunction presented in Section 3.2.2, is sufficient to capture the scope-related ordering among the operators. The operator projection can then be defined as the subtree consisting of the clausal skeleton plus the components marked by [OP +].

In RRG's syntactic representations, peripheral structures and clause linkage markers are attached to the clausal skeleton by arrows, which is a further non-standard element to be dealt with. In our proposal, we do not posit separate PERIPHERY nodes but mark peripheral structures by a feature as indicated in Figure 2. Similar to the operators, peripheral elements are subject to the iconicity principle that their relative ordering respects the layering of their attachment sites (see Figure 2b for an illustration). The modes of combining peripheral structures with the clausal skeleton are similar to those used for operators (Section 3.2.2). As to clause linkage markers, they can be treated along similar lines. However, since they behave more like prepositions and oblique case markers than like peripheral elements or operators, we regard their proper syntactic representation as an issue that calls for further clarification before coming up with a definite proposal of their formalization.

To sum up, we formalize the syntactic representations of RRG as labeled trees, where the labels can have additional features. Since the node labels can be regarded as feature values, too, e.g., category labels as values of the feature CAT, we can assume without restriction of

generality that node labels are sets of attribute-value specifications. From this perspective, $\text{ADV}_{[\text{PERI }+]}$ is short for [CAT *adv*, PERI +]. Introducing features allowed us to get rid of the PERIPHERY nodes, whose only purpose is to mark their daughters as peripheral. By the same line of reasoning we can eliminate the PRED nodes, whose only purpose is to mark their daughters as predicating (cf. Van Valin 2005: 13). The system of node attributes and the constraints on their values are still under development at the present stage of the formalization, and we will accordingly make a rather liberal use of features without specifying the full feature set of a node and without spelling out in detail the feature co-occurrence restrictions and related constraints.

3.2 The composition of syntactic structures

The standard presentation of RRG gives only an informal description of how syntactic templates are combined to more complex syntactic structures. Van Valin & La Polla (1997: 654, note 34) suggest that a formal account of the modes of composition may show some similarity to Tree Adjoining Grammars (cf. Joshi & Schabes 1997). Our approach to formalizing the syntactic side of RRG basically follows this hint.[2]

3.2.1 Substitution

The most basic mode of composition for syntactic templates is *substitution*. In the following, a *tree with label X* is meant to be a tree whose root carries the category label *X*. A tree β with label *X* can be substituted for a leaf node labeled with *X* (the *substitution node*) of a tree α by "identifying" the root node of β with the substitution node (cf. Figure 4a). More generally, if the nodes are labeled by feature structures, then the two feature structures must be compatible, and the node of the resulting tree is labeled by the *unification* of the two feature structures.

Substitution is the main mode of composition for expanding argument nodes by the syntactic representations of specific argument realizations. In Figure 3a, the two RP structures associated with *Kim* and *the glass* are substituted at the RP nodes of an argument structure template. The template in turn is selected by the predicating lexical element *smash(ed)*, where selection means substitution of the lexical tree at the predication leaf of the argument structure template (similar to lexical anchoring in Lexicalized Tree Adjoining Grammars). Finally, the substitution of the CLAUSE node at the top turns the clause into a complete sentence.

[2] Van Valin & La Polla (*ibid.*) rightly point out that TAG is a *grammar formalism* and not a *linguistic theory* in its own right. That is, TAG *per se* does not make any commitments about what kind of categories and what kind of syntactic configurations are appropriate for linguistic analysis. However, there is one caveat: the standard adjunction operation of TAG is not well suited for modifying flat structures, which is one of the reasons why our formalization employs slightly different modes of composition.

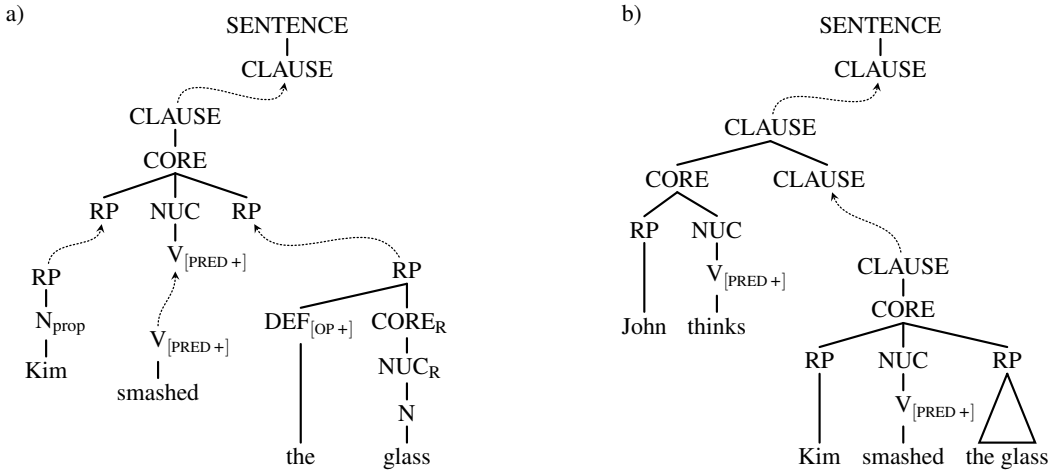


Fig. 3: Examples of substitution

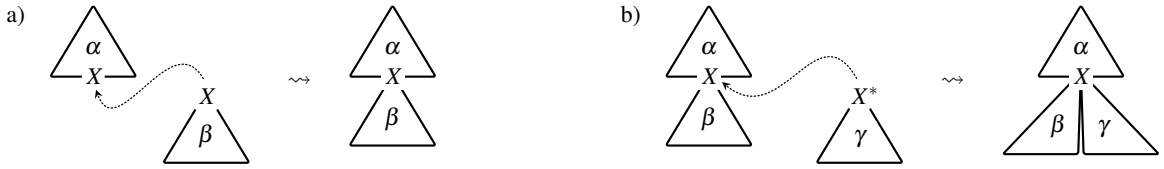


Fig. 4: Schematic sketch of substitution a) and adjunction (to the right) b)

Figure 3b shows an example in which the same clause is substituted as the complement of a clausal subordination construction associated with the verb *think*.

3.2.2 Adjunction

Peripheral structures cannot be added by substitution since they do not attach to leaves but to internal nodes in general, and the same holds for the operators; cf. Figures 1 and 2. The mode of composition we propose for these cases is *(sister) adjunction* (see also Kallmeyer *et al.* 2013).[3] As with substitution, we assume that the templates available for adjunction have a root label which coincides with the label of the target node (cf. Figure 4b). For convenience, the root of an adjunction tree is marked by an asterisk in the graphical presentations. Figure 5 sketches how the operator and periphery elements of the syntactic representations in Figure 2 can be added by adjunction. Notice that the root label of an adjunction tree specifies the attachment site at the phrasal skeleton. For example, the adjunction template selected by a temporal adverb like *yesterday* specifies the core as its layer of attachment. We also allow the substitution or the adjunction of trees at adjunction trees. For instance, the adjunctions of the

[3] Sister adjunction has been introduced as a composition operation on so-called *d-trees* by Rambow *et al.* (1995). Notice that sister adjunction differs from adjunction in TAG. In fact, the latter operation is more akin to the wrapping operation described in Section 3.2.3 below.

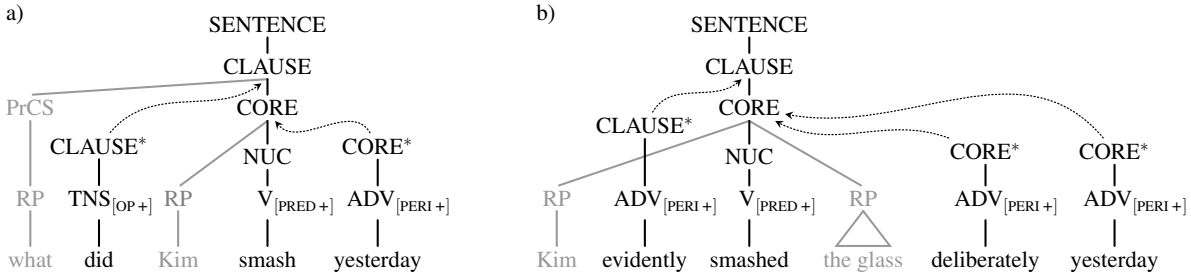


Fig. 5: Attachment of periphery and operator by (sister) adjunction

'deliberately' tree and the 'yesterday' tree to the core node in Figure 5b can be reduced to a single adjunction to the core node by first adjoining the 'yesterday' tree (to the right) of the 'deliberately' tree (or vice versa).

The examples in Figure 5 show a fairly arbitrary ordering of the adjoined substructures with respect to their sister nodes; even discontinuities can arise as shown by the placement of *evidently* within the core. However, within the operator and periphery projections, adjunction reduces to adjunction to the left or to the right of the daughters of the target node. It then remains to merge the projections with the constituent structure subject to additional ordering constraints, which are often language-specific. For instance, the tense operator must be placed between the precore slot and the core in Figure 5a, and the peripheral adverb *evidently* occurs between the first RP and the nucleus in Figure 5b. The position of the tense operator and its realization paradigm (here, by an auxiliary) is determined by the constructional schema of wh-questions in English. The placement of the peripheral adverbs is less restricted as long as the layer-related iconicity within the periphery projection is respected. For English, additional language-specific constraints would ensure that adverbials are not inserted between the verb and the following core RP argument and that phrasal adverbials go systematically to the right of the clausal skeleton.

3.2.3 Wrapping

Control constructions and extraction from complements pose a problem for the modes of composition presented so far. Consider the examples of wh-extraction in (1).

(1) a. What does John think Kim smashed?

b. What does John think Mary claimed Kim smashed?

Clearly, it would not be appropriate to assume a separate complex template for each of these constructions. The syntactic representations are to be composed of basic argument structure templates in a systematic way. There are several options for achieving this goal, depending on

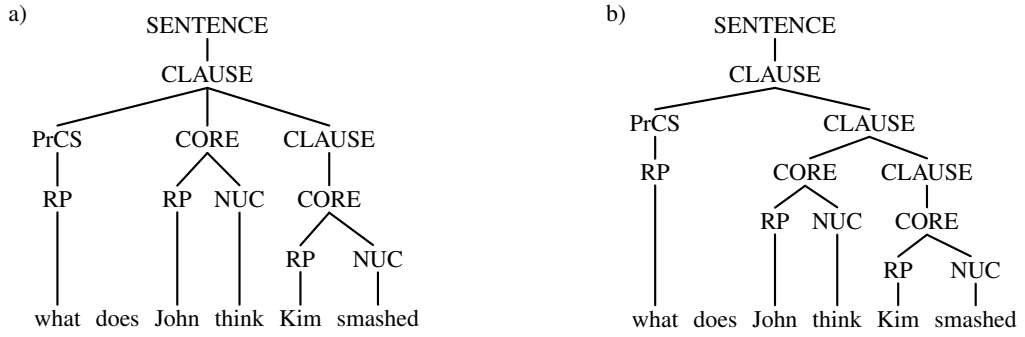


Fig. 6: Two possible syntactic representations of wh-extraction from complements

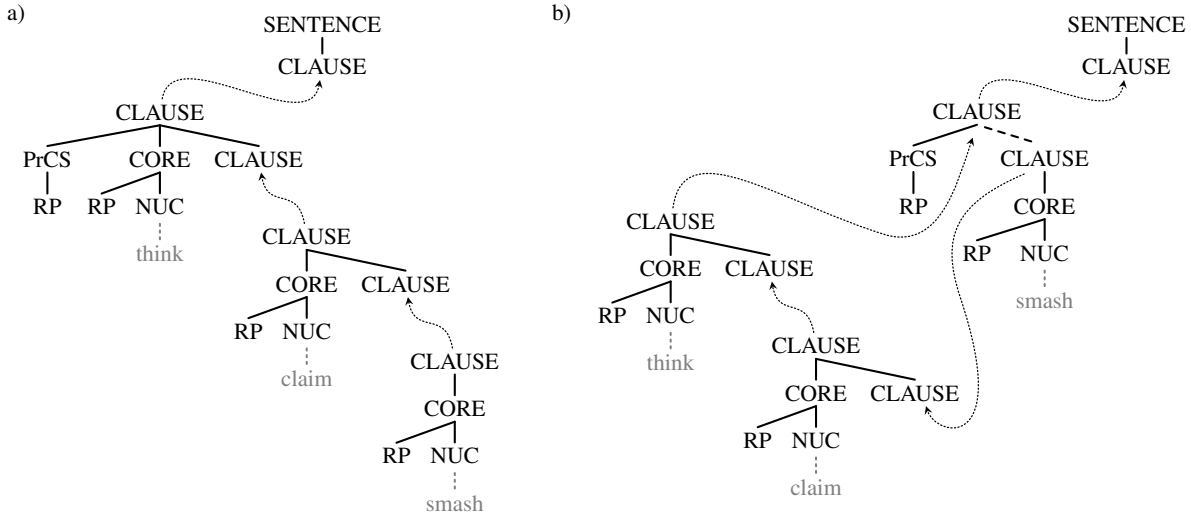


Fig. 7: Wh-extraction via simple substitution a) and wrapping substitution b)

the presumed inventory of elementary templates. First we need to decide on the proper syntactic representations of the examples in (1). While sentences of this type are discussed in the context of island constraints in Van Valin & La Polla (1997: 615) and Van Valin (2005: 273), no structural analysis is provided there. Due to the nature of the embedding constructions, the basic binary precore slot pattern [$_{\text{CLAUSE}}$ [$_{\text{PrCS}}$ ...][$_{\text{CORE}}$ ...]] shown in Figure 1 and Figure 5a does not apply to the present case. Figure 6 shows two possible alternatives. The analysis in Figure 6a assumes a precore variant of the clausal subordination pattern [$_{\text{CLAUSE}}$ [$_{\text{CORE}}$ ...][$_{\text{CLAUSE}}$ ...]]. The structure in Figure 6b, by contrast, assumes an additional clause node, and the precore slot pattern is [$_{\text{CLAUSE}}$ [$_{\text{PrCS}}$ ...][$_{\text{CLAUSE}}$ ...]]. In the following, we restrict the discussion to the first option since we regard it as difficult to come up with an independent motivation for the additional clausal node in Figure 6b.

Figure 7 sketches two ways of composing the syntactic representation of example (1b). Figure 7a employs substitution only, but at the price of assuming a special elementary template associated with 'think' that has a precore slot in addition to its normal argument slots.

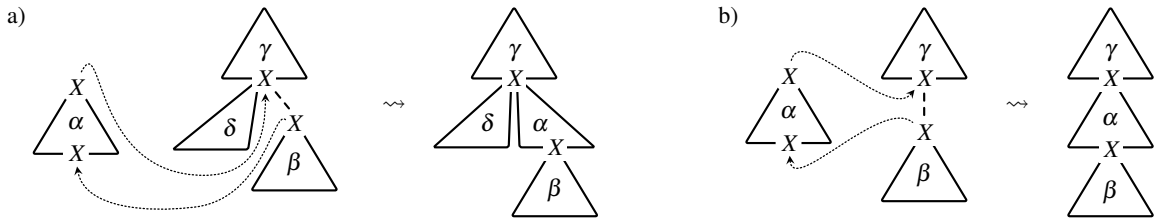


Fig. 8: Wrapping substitution (special cases)

Assuming such a template would raise the further problem of providing information about which of the arguments within the embedded clauses is referred to by the referent phrase in the precore slot. The templates in Figure 7b are more straightforward in this respect since they represent proper argument structure templates, that is, the precore slot is locally connected to the core from which the wh-word is extracted. In fact, the long-distance dependency comes about by the insertion of material between the precore slot and the corresponding core. On the downside, more flexible templates and a more complicated tree composition mechanism are required in this case. The clause node of the 'smash' structure is "split in two" and its upper part is unified with the upper clause node of the 'think' structure, thereby keeping the precore slot on the left side of the structure (similar to the adjunction mechanism introduced in the previous section); the lower part of the split node is substituted at the lower clause node of the 'claim' structure. The mode of composition just described is schematically depicted in Figure 8a and will be referred to as *wrapping substitution*.[4]

Let us evaluate the different tree composition options with respect to their applicability for linking, especially with respect to template selection. As pointed out above, it is not an option to assume that the syntactic inventory provides a single template for the complex syntactic structure of (1b). The structure has to be composed from argument structure templates which in turn are selected by the chosen lexical entries. The composition in Figure 7a has the disadvantages that the template selected by 'think' is not an argument structure template but has to be stipulated for embedded wh-questions, and that an advanced mechanism for coreference is needed. The composition scheme in Figure 7b, by comparison, has the advantage that the precore slot can be immediately linked when the template is selected. Here, the underlying assumption is that the syntactic inventory provides argument structure templates for wh-fronting.

Multiply embedded control and matrix-coding constructions may serve as another touchstone for template composition. They pose a number of challenges for RRG's syntactic analysis of such constructions as presented in Van Valin & La Polla (1997) and Van Valin (2005).

[4] Wrapping substitution is related to the concept of *flexible composition* proposed in Joshi *et al.* (2008), which allows one to interpret TAG adjunction as a wrapping operation.

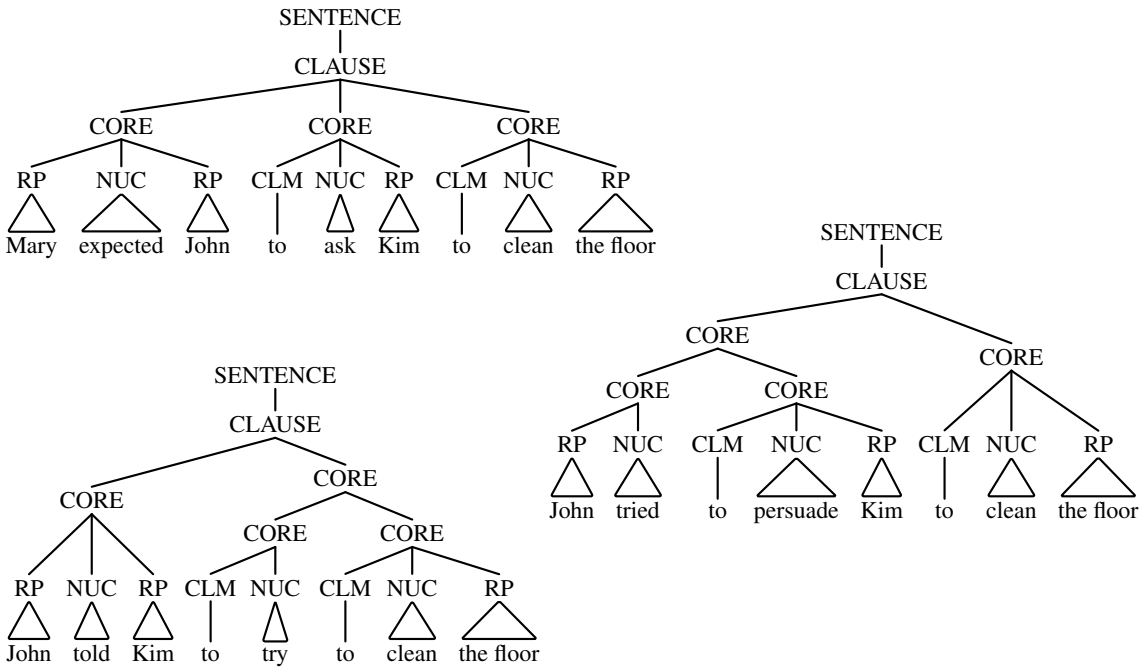


Fig. 9: Syntactic representations of the examples in (2)

Consider the examples in (2) and their syntactic representations in Figure 9.

(2) a. Mary expected John to ask Kim to clean the floor.

b. John tried to persuade Kim to clean the floor.

c. John told Kim to try to clean the floor.

Structures like the ones in Figure 9 show a mismatch between syntax and semantics in the following sense: semantically, the infinitival complements are arguments of the respective control predicates, but syntactically, they do not behave like core arguments. Example (2a) is instructive in showing how multiply embedded constructions can give rise to coordinating syntactic configurations. Notice that the clausal complements discussed in the previous section show also a mismatch between syntax and semantics in that they do not attach to the core but to the clause. But in contrast to the present case, multiple embeddings of clausal complements nevertheless correspond to embeddings on the syntactic side, though at the clause level and not at the core level. For this reason, simple substitution is sufficient for clausal complementation. This is different for the coordination structure of example (2a). Simple substitution cannot be applied here if we assume that the verbs *expect* and *ask* select elementary templates that contain a core daughter of the clause for the infinitival complement. Figure 10a sketches a possible solution that uses wrapping substitution, where we allow split nodes to carry differing top and bottom categories. Figure 10b shows that the same compositional mechanisms

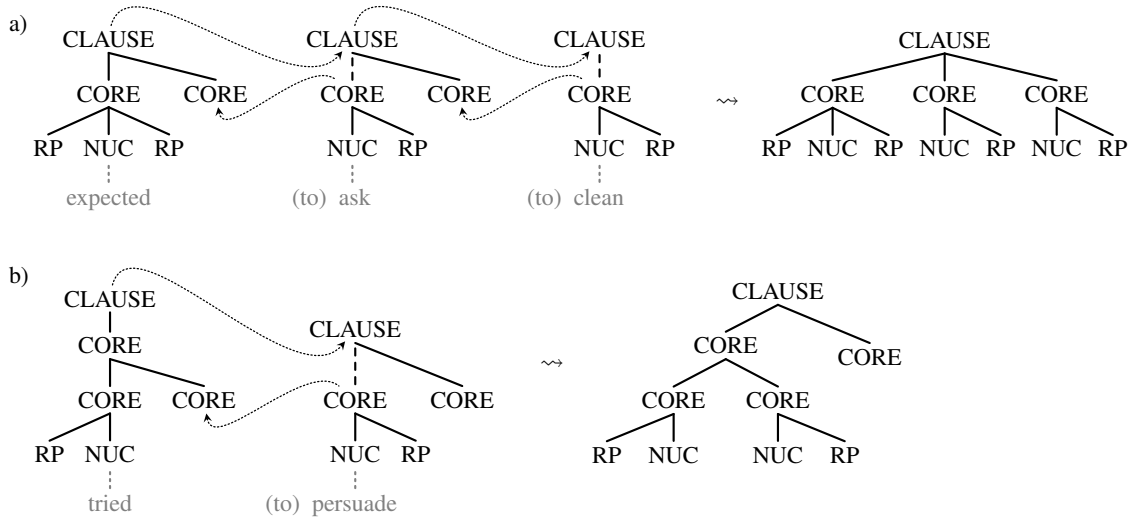


Fig. 10: Composition of templates by wrapping substitution for (2a) and (2b)

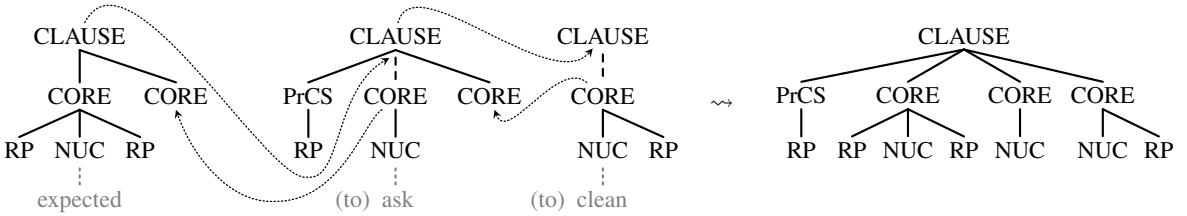


Fig. 11: Composition of the syntactic representation of (3a)

works for control verbs such as *try* which call for a core cosubordination template. Moreover, the treatment of wh-fronting introduced above nicely extends to embedded control and matrix-coding constructions like those in (3).

(3) a. Whom did Mary expect John to ask to clean the floor?
b. What did Mary expect John to ask Kim to clean?

For example, the syntactic representation of (3a) can be composed of elementary argument structure templates as illustrated in Figure 11. The general form of wrapping substitution used in this case in schematically depicted by Figure 12.

3.3 Extended domain of locality and the factoring of recursion

Consider again the composition scheme in Figure 7a. According to the traditional template selection principle of RRG, we can assume that the structure associated with 'smash' is available in the syntactic inventory. The number of core slots is reduced by one compared to the default core template of *smash* since one of the arguments occurs in precore position. As discussed above, this approach has the disadvantage that argument linking is not local; the wh-marked

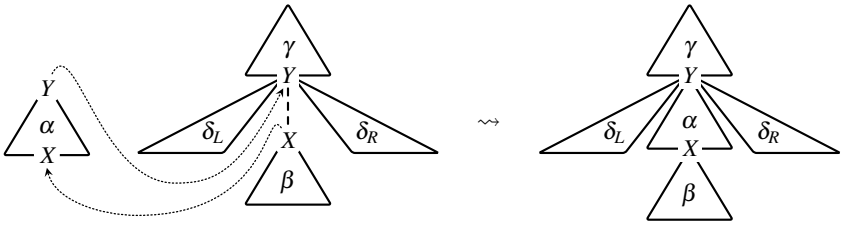


Fig. 12: Wrapping substitution (general case)

argument of *smash* can only be linked after the precore slot template is attached. But at this point of the composition, the referent of the wh-word is not locally accessible anymore because of the intervening clausal subordination structures. A possible solution is to have the information about the wh-marked argument percolate to the top of the tree. A constraint-based formalization of this percolation process could roughly work as follows: The core node of the reduced template for 'smash' carries a (set-valued) feature which contains the referential index of the participant not locally realized as well as its wh-marking. A general constraint then ensures that clause and sentence nodes collect the non-realized indices of their subordinate clauses and cores minus the indices that are realized in precore slot daughters and the like. This way of bookkeeping for modeling long distance dependencies is in fact closely related to the use of 'slash' or 'gap' features in the approaches of Sag & Wasow (1999) and Ginzburg & Sag (2001), among many others. To sum up, the approach just discussed can get along with simple substitution at the price of a considerable amount of bookkeeping. By comparison, the preferred approach exemplified in Figures 7b allows a fully local argument linking but requires a more complex method of tree composition. The former strategy is employed in GPSG/HPSG-related frameworks while the latter strategy is characteristic of approaches in the line of Lexicalized Tree Adjoining Grammars (LTAG; cf. Kroch 1987).

Two important characteristics of the LTAG formalism are the *extended domain of locality* of elementary trees and the *factoring of recursion from the domain of dependencies* by means of TAG-adjunction (Joshi & Schabes 1997: 95f). The first property means that elementary trees represent full argument projections and that they can have a complex constituent structure. The usual example given in this context is that of elementary argument structure trees associated with transitive verbs which contain an internal VP constituent. Since in RRG, the basic argument structure templates are flat, the availability of deeper hierarchical structures is less relevant in the first place. This observation hints at a crucial difference between RRG and the way LTAG is usually applied to linguistic analysis. In the LTAG literature, a standard example of the factoring of recursion by adjunction is the insertion of adverbials. Put briefly, an adverbial is added by adjoining a so-called auxiliary tree, which leads to an additional branching. That is, if a VP node or some other internal node is required as the landing site of

the adverbial auxiliary tree, then the structure gets deeper by adjunction and the verb and the direct object may become farer apart.

From an RRG point of view the described effect may be regarded as an artifact of the structural assumptions. If argument structure templates are basically flat and if modifiers are added as sister structures then there is no increase of structural depth in this case. In other words, the structure of the clausal skeleton is not affected by modification. The situation is different in the case of long distance dependencies across clausal complements. Clausal complements increase the depth of the clausal skeleton. It is this case where the notions of an "extended domain of locality" and of "factoring recursion from the domain of locality" directly apply to RRG, at least if we presume elementary argument structure templates with precore slots as illustrated in Figure 7b. The multiple coordination structure of Figure 11 is another type of example to which the factoring of recursion from the domain of locality applies. In this case, the clausal skeleton does not grow in depth but in breadth.

3.4 The syntactic inventory

3.4.1 Syntactic templates

The syntactic templates of RRG are introduced as the basic building blocks of the syntactic representations. No attempt has been made up to now to specify exactly the templates of the syntactic inventory. The discussion of the two approaches to wh-fronting in Section 3.2.3 shows that the question of which templates need to be available in the inventory is not independent of the modes of composition employed. If extraction structures are fully linked locally, as in Figure 7b, then there is no need for a precore slot template to be attached separately. In our formalization proposal, we assume that argument structure templates are clause templates and have slots for each of the arguments occurring in the clause. Let us call this the *full clause projection* assumption.[5]

For example, the templates in Figure 13 represent alternative realization patterns of transitive verbs in English. Notice the two templates for the passive, one in which only the undergoer occurs as an argument, and a second template which includes the realization of the actor by a peripheral *by*-phrase. Standard RRG would probably regard the peripheral *by*-phrase as a separate template that can be adjoined to the passive core. In the context of the framework developed so far, the question is whether the *by*-phrase should be added by (sister) adjunction like an adverbial. By listing the two realization patterns for the passive as elementary templates in Figure 13, we opt against adjunction in this case. The main disadvantage of

[5] A similar constraint on elementary trees is often proposed in TAG-based approaches to linguistic analysis; cf., e.g., Frank (2002).

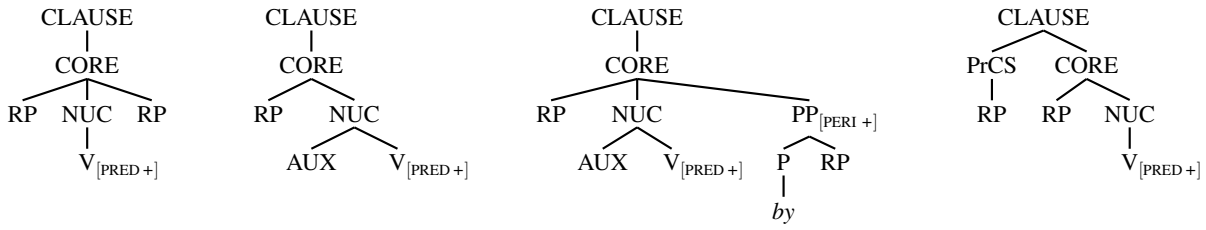


Fig. 13: Basic transitive predication template for English with variants

modeling the *by*-phrase by adjunction would lie in the constraints that need to be imposed, since adding a peripheral *by*-phrase that encodes the actor is restricted to specific circumstances. The optional peripheral *by*-phrase is part of the description of how passive voice is realized in English. Of course, a mere enumeration of the two passive templates in Figure 13 is not satisfactory for a theory of grammar. At some point, the theory should state explicitly that it is the addition of the *by*-phrase which relates the second template to the first. In our framework, this relation is encoded at the level of template *specifications* (cf. Section 3.4.2).

The rightmost template in Figure 13 represents an elementary precore slot argument structure template. This differs from the template system informally suggested in Van Valin (2005: 15), which regards the structure in question as composed of a reduced core template and a precore slot-clause fragment (cf. Figure 1b). Again, we propose that compositions of this type are best modeled at the level of template specifications. It is worth mentioning that the precore slot template in Figure 13 covers actually two different types of argument realization patterns: fronted elements in declarative sentences (as in *Bean soup I don't like*), which go along with focal stress, and fronted wh-words in interrogative sentences of the kind discussed above. The latter pattern differs from the former in the way tense is realized. The commonalities between the two templates would then again be expressed on the specification level.

3.4.2 Template specification by tree descriptions

In the previous section, we raised the question of how to characterize the (universal and language-specific) syntactic templates in a systematic way. The templates proposed above as elementary are more complex than the fragmentary templates usually assumed in RRG. The key advantage of our approach is that the composition of syntactic representations can be reduced to the three modes of composition presented in Section 3.2, namely substitution, adjunction, and wrapping. However, this leaves us with the problem of how to describe in which way the templates are built from more elementary components, and how they are related to each other. As mentioned before, it would be rather unsatisfactory if we had to regard the two

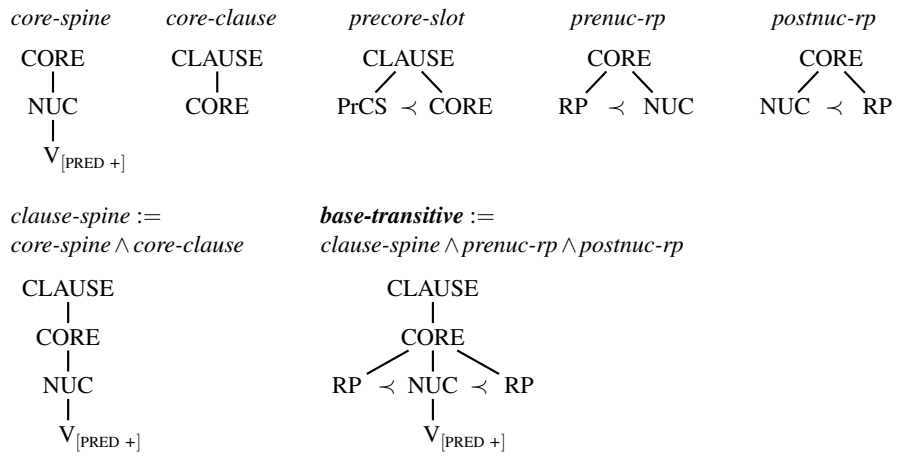


Fig. 14: Example specifications of syntactic fragments

passive templates in Figure 13 as independent units without being able to make explicit the relation between them.

The proposed solution is to treat syntactic templates as *minimal models* of tree descriptions. Relations between templates can then be captured by relating the respective descriptions. For instance, the specification of the passive template with *by*-PP consists of the specification of the simple passive template plus a specification of the *by*-PP and constraints on its position. Our use of tree descriptions for specifying syntactic templates in a modular way is inspired by the *metagrammar* approach of Crabbé & Duchier (2005), where a metagrammar is basically a system of tree descriptions that defines the syntactic inventory as the set of the corresponding minimal models.

Tree descriptions in this sense consist of dominance and precedence constraints as well as category and feature assignments. Consider the example specifications in Figure 14, which are depicted in tree-like diagrams but are to be read as tree descriptions. For instance, the specification with the name *precore-slot* says that there are three distinct nodes n_0, n_1 and n_2 labeled respectively by CLAUSE, PrCS, and CORE, where n_1 and n_2 are daughters of n_0 and n_1 immediately precedes n_2 (expressed by $\prec$). Figure 14 illustrates how the basic transitive template of Figure 13 can be defined by a piecewise combination of such specifications. The precore slot template of Figure 13 can be likewise defined by conjoining the specifications *clause-spine*, *precore-slot*, and *prenuc-rp*. In this way, common components of elementary templates are made accessible in the metagrammar.

4 From logical structures to decompositional frames

There are various ways to introduce a formal account of RRG's logical structures. Probably the most straightforward option is to regard them as expressions of a formal language, which means to give an inductive definition of these expressions by specifying their atomic components and the modes of combining them to larger expressions.[6] However, this approach has the following two disadvantages: First, there are a number of additional semantic components used within RRG's logical structures whose formal syntax seems not fully fixed yet. Among them are the internal variables (Van Valin & La Polla 1997: 117), the qualia structures, the representation of the operator semantics, and the representation of the semantics of the intra-clausal connectives. All of these representational elements would affect the syntactic make-up of the logical structures and should thus be part of their formal definition. A second issue is the interpretation of the logical structures, which is discussed in RRG expositions only at an intuitive level. From a formal point of view, the logical structures are meaningless expressions built in accordance with certain formation rules. While not absolutely necessary for a formalization of RRG as is, we think it desirable to also have a formal account of the semantics represented by the logical structures. A logically more explicit formulation of RRG's semantic representations is needed anyway, for instance, for drawing inferences.[7]

Instead of working out the formal syntax of the logical structures, we propose a slightly different approach for representing the semantic information they encode. The goal of our formalization is to keep the key properties of RRG's decompositional system without necessarily preserving the specific form of the logical structures. Moreover, we aim at a *formal semantics* account, which does not only address the form of the semantic representations but also their interpretation. For these reasons, among others, we make use of *decompositional frames* for the semantic representation. As we will see below, semantic frames, understood as feature structures, allow a fairly straightforward reconstruction of the logical structures of RRG.[8] The theory of feature structures come with a well-explored logic and a model-theoretic semantics; another advantage is the availability of *unification (under constraints)* as a universal mode of composition (cf. Kallmeyer & Osswald 2013 and the references therein).

A central purpose of the decompositional system that underlies the logical structures of RRG is the representation of *Aktionsart* types. The representational system allows the distinction between states, activities, changes of state, causative scenarios, etc., and it seems reasonable to assume that the decompositional structures are intended to reflect the internal

[6] This is what Dowty (1979: 352ff) does for his "translation language", which contains the decompositional elements Foley & Van Valin (1984) originally build on. In contrast to RRG's logical structures, Dowty's language is a full-fledged logical language with lambda abstraction which has a model-theoretic semantics.

[7] Cf. Periñan & Mairal (2009) for an argument in this direction.

[8] Cf. Nolan (2004) for a related proposal.

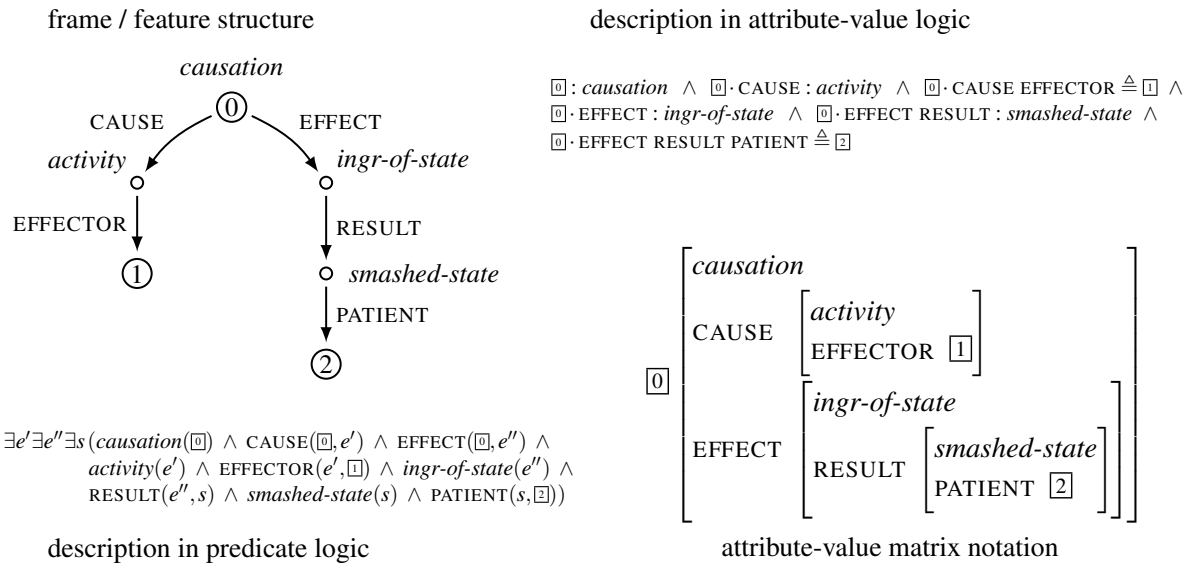


Fig. 15: Frame representation corresponding to (4)

structure of the described event or state of affairs at some level of conceptual representation. Our frame-based formalization tries to make explicit the reference to events and their subcomponents (see also Osswald & Van Valin 2014). Consider, for example, the decompositional structure shown in (4), which is associated with the causative reading of the English verb *smash*.

(4) [**do'**(x, ∅)] CAUSE [INGR **smashed'**(y)]

The logical structure basically says that an event denoted by transitive *smash* consists of an activity of someone or something *x* which causes a sudden change of state of something *y*, namely *y* getting into the state of being smashed. Put differently, the event in question is of type *causation* and has as its CAUSE component an (unspecified) activity of *x* and as its EFFECT component an ingression of a state which results in *y*'s state of being smashed.[9] This means that the event conceptualization associated with the transitive use of the verb *smash* consists of certain subcomponents (activities, changes of state, etc.), the participants involved, and various relations between these entities. The left side of Figure 15 shows a graphical representation of the structure for the example under discussion. (The numbers are labels for referring to specific components; they will become relevant below in the context of linking.) An essential property of this relational structure is that all relations are *functional*:

[9] Note that the use of CAUSE here differs from the one in (4); CAUSE denotes a functional relation that relates an event of type causation to its cause component. In (4), by contrast, CAUSE is to be interpreted as a relation between an activity and a change of state, that is, as the causation relation between events.

it is *the* cause component of the causation event, *the* effector of the activity, etc. Moreover, all components of the structure are "accessible" by functional relations from a distinguished set of *base* elements. In the example, every node is accessible from the base element labeled by 0, which, in this case, is the *root* of the structure. We refer to structure of the kind just described as *feature structures* or *frames*. More precisely, frames are defined as *base-labeled feature structures with types and relations*; see Kallmeyer & Osswald (2013: 280ff) for a formal definition.

By treating decompositional semantic representations as frame structures, we take a clear position with respect to the distinction between logical descriptions and described models. Frame structures are models; they are relational structures of a specific kind and they can be easily characterized in terms of first-order predicate logic. For instance, the frame structure depicted on the upper left of Figure 15 is a generic minimal model of the formula shown on the lower left of the Figure, where the boxed labels are used as constants and the predicates CAUSE, EFFECT, EFFECTOR, etc. are assumed to be functional. Logical representations of this kind are closely related to Neo-Davidsonian approaches to event semantics as proposed in Parsons (1990), among others. It is also possible to characterize frame structures by means of special purpose logics, so-called *feature* or *attribute-value logics*, which can be seen as variable-free fragments of first-order predicate logic. For the example under discussion, a possible attribute-value description is shown on the upper right of Figure 15. The more familiar attribute-value matrix below can be seen as a special syntactic device for writing attribute-value descriptions.

5 Linking and constructional schemas

A thorough formalization of the RRG linking system is far beyond the scope of the present paper. For this reason, and because of lack of space, we restrict the discussion on linking and constructional schemas to a case study of the English adjectival resultative construction (*kick open*, *wipe clean*, etc.). A schema for this construction can be found in Van Valin (2005: 239). The idea of constructional schemas in RRG is that "[o]nly the idiosyncratic, language-specific features of constructions are represented in constructional schemas. Hence constructional schemas, by virtue of their reference to general principles, permit the capturing of cross-linguistic generalizations, while at the same time expressing language-particular properties of grammars" (Van Valin 2005: 132).

In English (and many other languages), adjectival resultative constructions are nuclear cosubordination structures. The two predicates form a complex predicate, which means that argument linking works basically in the same way as for simple sentences (Van Valin 2005:

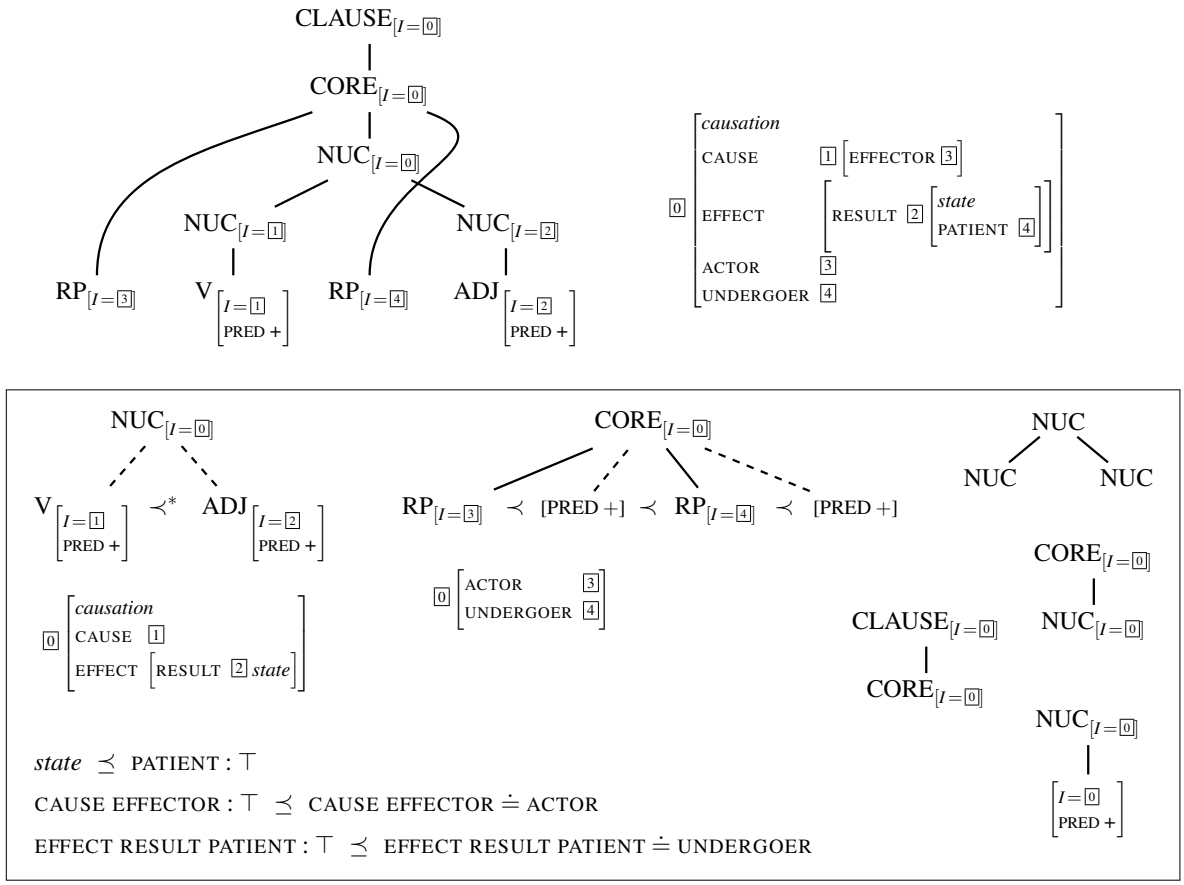


Fig. 16: Decomposition of the English adjectival resultative schema

225). The upper row of Figure 16 shows the corresponding argument structure template together with its associated causative change of state frame, which is enriched by the assignment of the macroroles. (The *I* feature of a constituent is the referential index that refers to the denoted component in the semantic frame.) We can now try to decompose this template and its associated semantic frame into general and language-specific components. Notice that this has to be done at the level of template and frame specifications since the template in question is an elementary argument structure template and not the result of compositions in the sense of Section 3.2. The lower row of Figure 16 provides a sketch of the possible components involved. On the right, we have specifications of general layer fragments and of the bare nuclear cosubordination structure. The component in the middle of the box specifies a transitive serializing multi-predicate core structure with associated macrorole assignment. (Dashed lines indicate dominance while solid lines stand for immediate dominance.) The specification on the left, which is the most language-specific component, says that a predicative verb followed by a predicative adjective can form a nucleus that has a causative change of state

interpretation, where the verb denotes the cause and the adjective denotes the result state. The constraints at the bottom of the box are responsible for the correct assignment of the macroroles.[10] For instance, the second to last constraint says that if there is an effector of a causing event then it is assigned the actor macrorole.

Although a number of details have still to be spelled out, the given example highlights a possible route to formalizing constructional schemas in a way that is compatible with the original motivation for introducing them into the architecture of RRG.

6 Conclusion

In this article, we presented a first outline of how a formalization of RRG could look like which attempts to give full consideration to all aspects of the theory. We proposed two additional desiderata with respect to the current organization of RRG: A formalization should draw a clearer distinction between declarative and procedural elements and it should put more emphasis on syntactic and semantic compositionality. The larger part of the article was concerned with the formalization of the syntactic side of RRG. To this end, we introduced three modes of composition for syntactic structures: (simple) substitution, (sister) adjunction, and wrapping (substitution). In particular, wrapping substitution has turned out to be an appropriate mode for composing embedded control and matrix-coding constructions, which have not drawn much attention in the RRG literature before. For formalizing the semantic side of RRG, we proposed the use of decompositional frames. The English resultative construction has served as an example for showing how semantic frames can be integrated with syntactic templates in a flexible way. Of course, many issues remain to be explored. For instance, the formal representation of information structure has been completely neglected in this article.

Lack of space has prevented us from saying more about how the formalism developed so far can be employed for parsing. We refer the reader to Kallmeyer *et al.* (2013) for a formal parsing scheme developed for a slightly simplified version of the above modes of composition. Moreover, Lichte & Petitjean describe ongoing work on adding a frame-semantic component to the XMG grammar development system (Crabbé *et al.* 2013), which can be seen as a first step towards the kind of grammar workbench envisaged in the introduction.

References

Crabbé, Benoit & Duchier, Denys. 2005. Metagrammar redux. In *Constraint Solving and Language Processing*. H. Christiansen, P. R. Skadhauge & J. Villadsen (eds), Lecture Notes in Computer Science 3438, 32–47. Berlin: Springer.

[10] See Kallmeyer & Osswald (2013: 287ff.) for a thorough introduction of the formal notations.

Crabbé, Benoit, Duchier, Denys, Gardent, Claire, Le Roux, Joseph & Parmentier, Yannick. 2013. XMG: eXtensible MetaGrammar. *Computational Linguistics*, 39(3): 591–629.

Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel.

Foley, William A. & Van Valin, Robert D., Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: CUP.

Frank, Robert. 2002. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: MIT Press.

Ginzburg, Jonathan & Sag, Ivan A. 2001. *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. Stanford, CA: CSLI Publications.

Guest, Elizabeth. 2008. Parsing for Role and Reference Grammar. In *Investigations of the Syntax-Semantics-Pragmatics Interface.* R. D. Van Valin Jr. (ed.), 435–454. Amsterdam: John Benjamins.

Johnson, Mark. 1987. A new approach to clause structure in Role and Reference Grammar. In *Davis Working Papers in Linguistics 2*, 55–59. Davis, CA: University of California.

Joshi, Aravind K., Kallmeyer, Laura & Romero, Maribel. 2008. Flexible composition in LTAG: Quantifier scope and inverse linking. In *Computing Meaning*. H. Bunt & R. Muskens (eds), vol. 3, 233–256. Berlin: Springer.

Joshi, Aravind K. & Schabes, Yves. 1997. Tree-adjoining grammars. In *Handbook of Formal Languages. Vol. 3: Beyond Words*. G. Rozenberg & A. Salomaa (eds), 69–123. Berlin: Springer.

Kallmeyer, Laura & Osswald, Rainer. 2013. Syntax-driven semantic frame composition in Lexicalized Tree Adjoining Grammars. *Journal of Language Modelling*, 1(2): 267–330.

Kallmeyer, Laura, Osswald, Rainer & Van Valin, Robert D., Jr. 2013. Tree wrapping for Role and Reference Grammar. In *Formal Grammar (FG 2012/2013)*. G. Morrill & M.-J. Nederhof (eds), Lecture Notes in Computer Science 8036, 175–190. Springer.

Kowalski, Robert. 1979. Algorithm = logic + control. *Communications of the ACM*, 22(7): 424–436.

Kroch, Anthony. 1987. Unbounded dependencies and subjacency in a Tree Adjoining Grammar. In *The Mathematics of Language*. A. Manaster-Ramer (ed.), 134–172. Philadelphia: John Benjamins.

Lichte, Timm & Petitjean, Simon. 2015. Implementing semantic frames as typed feature structures with XMG. *Journal of Language Modelling*, 3(1): 185–228.

Nolan, Brian. 2004. First steps toward a computational RRG. In *Proceedings of the International Role and Reference Grammar Conference 2004*. B. Nolan (ed.), 196–223. Dublin: Institute of Technology Blanchardstown.

Osswald, Rainer & Van Valin, Robert D., Jr. 2014. FrameNet, frame structure, and the syntax-semantics interface. In *Frames and Concept Types*. T. Gamerschlag, D. Gerland, R. Osswald & W. Petersen (eds), 125–156. Dordrecht: Springer.

Parsons, Terence. 1990. *Events in the Semantics of English.* Cambridge, MA: MIT Press.

Periñán, Carlos & Mairal, Ricardo. 2009. Bringing Role and Reference Grammar to natural language understanding. *Procesamiento del Lenguaje Natural*, 43: 265–273.

Pollard, Carl J. & Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Rambow, Owen, Vijay-Shanker, K. & Weir, David. 1995. D-tree grammars. In *Proceedings of the 33rd Annual Meeting*, 151–158. Association for Computational Linguistics.

Sag, Ivan A. & Wasow, Thomas. 1999. *Syntactic Theory: A Formal Introduction.* Stanford, CA: CSLI Publications.

Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

—. 2006. Semantic macroroles and language processing. In *Semantic Role Universals and Argument Linking. Theoretical, Typological, and Psycholinguistic Perspectives*. I. Bornkessel, M. Schlesewsky, B. Comrie & A. D. Friederici (eds), 263–301. Berlin: Mouton de Gruyter.

—. 2010. Role and Reference Grammar as a framework for linguistic analysis. In *The Oxford Handbook of Linguistic Analysis*. B. Heine & H. Narrog (eds), 703–738. Oxford: OUP.

Van Valin, Robert D., Jr. & LaPolla, Randy J. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: CUP.