RRG 2019

International Conference on Role and Reference Grammar

August 19-21, 2019

University at Buffalo
The State University of New York

Welcome

Welcome to the 2019 International Conference on Role and Reference Grammar at the University at Buffalo, The State University of New York. UB has been one of the centers of research in RRG, and we are delighted to be able to host the Conference. We look forward to your participation and contributions to the RRG community. We hope you will enjoy your time in Western New York.

Robert D. Van Valin, Jr.,
for the Program and Local Organizing Committees

Program Committee

Delia Bentley, University of Manchester
Daniel Everett, Bentley University
Lilian Guerrero, National Autonomous University of Mexico
Rolf Kailuweit, Heinrich-Heine-University Düsseldorf
Ricardo Mairal, National University of Distance Education
Brian Nolan, Technological University Dublin
Toshio Ohori, Keio University
Mitsuaki Shimojo, University at Buffalo
Robert D. Van Valin, Jr. Heinrich-Heine-University Düsseldorf & University at Buffalo

Local Organiaing Committee (University at Buffalo)

Joseph Baiz
Peggy Bakula
Hanno Beck
Erika Bellingham
Kiyono Fujinaga
Alexandra Lawson
Jihye Seong
Mitsuaki Shimojo
Robert D. Van Valin, Jr.

The abstracts are NOT included in the hardcopy booklet. There is, however, a PDF file containing the abstracts that can be downloaded from the **conference website**.
https://ubwp.buffalo.edu/rrg2019/

Acknowledgements

The organizers of the RRG 2019 Conference would like to thank the folowing for their generous support.

Cooperative Research Center 991
Heinrich-Heine-University Düsseldorf

Departmetn of Linguistics
University at Buffalo

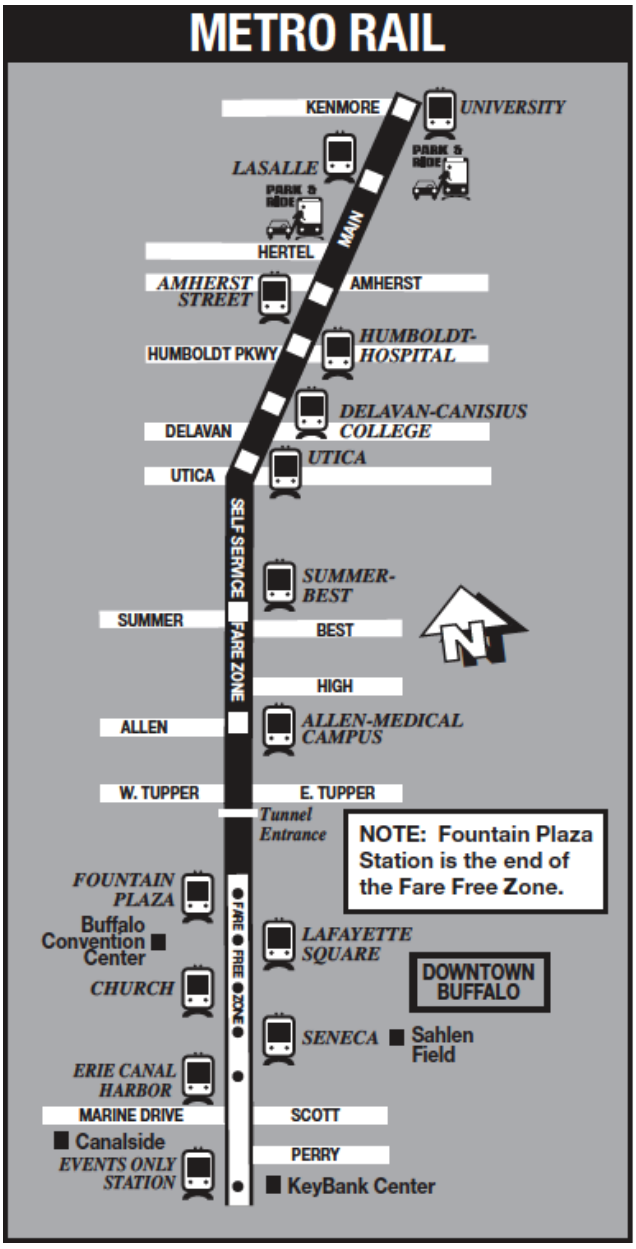
Asian Studies Program
University at Buffalo

Conference Venue

The conference is held in downtown Buffalo at the University at Buffalo's Educational Opportunity Center, 555 Ellicott Street, Buffalo, NY.

Public Transportation

The train operates between downtown Buffalo's Erie Canal Harbor Staton and the University Metro Rail Station (South Campus of University at Buffalo). You may ride free between stations on the mall (Fare Free Zone in the route map). However, if your ride takes you underground, you must pay a fare ($2/ride). The closest station to the conference venue is Fountain Plaza (about 10 min walk).
See https://metro.nfta.com/Routes/pdfs/145.pdf for train schedules and details.



Local Attractions (including nearby restaurants)

Please see https://ubwp.buffalo.edu/rrg2019/venue/
You can also search "UB educational opportunity center" in Google Map to see nearby points of interest.

Connecting to a Guest Wi-Fi Network

http://www.buffalo.edu/ubit/service-guides/connecting/wifi/guest.html

1. Select **UB Connect** from the available Wi-Fi networks on your device.
2. Open a Web browser and go to https://clearpass-portal1.cit.buffalo.edu/guest/Guest_Login.php.
3. On the **UB Network: Guest Account Login** screen, click **Create a New Guest Account**, in the green callout box.
4. Enter:
    - Your first and last name in the **Your Name** field.
    - Your non-UB email address in the **Email Address** field.
5. Check the box to **accept the terms of use**.
6. Click **Create Account**.
7. A browser will open. You have 30 minutes to check your email for your password so you can log in.
8. Go to https://clearpass-portal1.cit.buffalo.edu/guest/Guest_Login.php.
9. Decide how long you need to connect:

- **30 minutes or less**:
    1. Enter your email address in the **Username** field.
    2. Enter the password sent to your email in the **Password** field.
    3. Click **Create Device**.
- **Up to 6 months:**
    1. Click **Register This Device**.
    2. Enter:
        - **Your email address** in the **Username** field.
        - **The guest password** sent to your email in the **Password** field.
    3. Click **Log In**.
    4. Enter a unique name for your device, such as **Jane Doe's Macbook**.
    5. If the MAC Address is not filled in, enter the MAC Address for your device.
    6. Click **Create Device**.

How can I find my device's MAC Address?
All devices are different. For most devices, you can Google **Find MAC Address** and your device, e.g. **Find MAC Address iPad**, for instructions on how to find your device's MAC Address.

Program

August 19 (Mon)		
9:00-9:30	Registration & Opening	
9:30-10:00	Maria Katarzyna Prenner	Different degrees of agentivity in sentience verbs? On the decomposition of proto-agent features in Polish
10:00-10:30	Delia Bentley	The logical structure of verbs of quantized and non-quantized change
10:30-11:00	Break	
11:00-11:30	Felipe Hasler	Clausal cosubordination in Mapudungun: Relevance of referential hierarchies for the RRG Linking Algorithm
11:30-12:00	Carmen Conti	Clausal cosubordination in Spanish
12:00-12:30	Hideki Kishimoto	When Intransitives Behave Like Passive: De-causativization in Japanese
12:30-14:00	Lunch	
14:00-14:30	Christian Chiarcos, Christian Fäth & Monika Rind-Pawlowski	Towards a Role and Reference Grammar corpus for English
14:30-15:00	Tatiana Bladier, Kilian Evang, Laura Kallmeyer, Robin Möllemann & Rainer Osswald	Creating RRG treebanks through semi-automatic conversion of annotated corpora
15:00-15:30	David Arps, Tatiana Bladier & Laura Kallmeyer	Chart-based RRG parsing for automatically extracted and hand-crafted RRG grammars
15:30-16:00	Break	
16:00-16:30	Saima Hafeez	Neither dative nor instrumental as default in Urdu: Dative/instrumental marked non-MR arguments as PSA
16:30-17:00	Kiyoko Toratani	Syntax of Japanese dish names on restaurant menus: An emerging structure
17:00-18:00	**Plenary: Daniel Everett**	**Did Homo erectus use an RRG grammar?**
18:00-18:30	Reception	
18:30-19:30	John Ball, Chief Technology Officer, Pat, Inc.	Introducing Pat, a natural language understanding system using RRG

August 20 (Tue)		
10:00-10:30	Hanno Beck	Formalizing Hierarchy-Related Generalizations in RRG: The Brittle and the Supple
10:30-11:00	Robert D. Van Valin, Jr.	On the status of lexical categories in RRG
11:00-11:30	Break	
11:30-12:00	Nicolai Winther-Nielsen	A Role and Reference Grammar of Biblical Hebrew Clause Linkage
12:00-12:30	Armando Mora-Bustos & Verónica Fidencio Núñez	Operators, associated movement and subject in Mazahua (Otomanguean)
12:30-14:00	Lunch	
14:00-14:30	Kata Balogh	Information-structurally driven syntactic configurations
14:30-15:00	Mitsuaki Shimojo	On the interactions with pragmatics in Role and Reference Grammar
15:00-15:30	Anja Latrouite & Robert D. Van Valin, Jr.	Degrees of discourse configurationality and beyond
15:30-16:00	Break	
16:00-16:30	Sergio Ibáñez	Intraclausal complexity in Spanish: The case of depictive constructions.
16:30-17:00	Erika Bellingham, Pia Järnefelt, Kazuhiro Kawachi, Yu Li, Alice Mitchell, Guillermo Montero Melis, Sang-Hee Park, Anastasia Stepanova, Emanuel Bylund & Jürgen Bohnemeyer	Modeling causative complexity across languages with the Interclausal Relations Hierarchy
17:00-18:00	**Plenary: Wataru Nakamura**	**Two ways of "undoing" the linking process: Impersonal constructions in Russian and beyond**
19:00-21:00	Conference dinner	

August 21 (Wed)		
10:00-10:30	Christian Canu Højgaard	Agency and lexical decomposition of Biblical Hebrew verbs
10:30-11:00	Lilián Guerrero & Valeria Belloro	Double locatives: The case of external possession in Yaqui
11:00-11:30	Break	
11:30-12:00	Yoko Hasegawa	A Study of the Annotative Dual-Sentence Juxtaposition Construction in Japanese
12:00-12:30	Erika Bellingham	Event integration in the English by-Ving (MEANS) construction: An RRG analysis
12:30-14:00	Lunch	
14:00-14:30	Irene Murtagh	Accommodating Irish Sign Language in an Extended Role and Reference Grammar Lexicon Architecture
14:30-15:00	Felipe Hasler, Carlos González, Silvana Guerrero, Matías Jaque & Verónica Orqueda	Relevance of the meaning of operators in determining the basic meaning of a construction: The case of agarró y se fue in Spanish
15:00-15:30	Kiyono Fujinaga	The analysis of Japanese benefactive constructions in RRG
15:30-16:00	Break	
16:00-16:30	Eibhlín Ní Fhallamháin	The Genitive Case in Irish: A functionalist account of complex multifunctional syntax
16:30-17:00	Valeria Generalova	Integrating head-marking and dependent-marking properties: An RRG analysis of Halkomelem ditransitive constructions
17:00-18:00	**Plenary: Jürgen Bohnemeyer**	**Operator, information: Revisiting the operator projection in RRG, with special emphasis on tense, aspect, and finiteness**
18:00-18:10	Closing	

Plenary Lecture

Did Homo erectus use an RRG grammar?

Daniel Everett
Bentley University

In this talk I discuss the evolution of human language, based on Everett (2017) and Barham and Everett (2019). I will argue that tools of Homo erectus show the origins of symbols (in the Peircean sense) and hence that language originated nearly 2 million years ago. I will discuss the role of syntax in supporting the semiotics of language (what Peirce referred to in 1865 as "Universal Grammar") and argue that Role and Reference Grammar provides the best available model for understanding syntax at each stage of grammatical and semiotic evolution.

References

Everett, Daniel. 2017. How Language Began: The Story of Humanity's Greatest Invention. New York: W. W. Norton.

Barham, Larry and Daniel Everett. 2019. "Tools, Symbols and Evidence for Language in the Lower Palaeolithic." Ms. University of Liverpool and Bentley University

Plenary Lecture

Two ways of "undoing" the linking process: Impersonal constructions in Russian and beyond

Wataru Nakamura
Tohoku University

Since its inception, RRG has developed a linking theory with significant crosslinguistic coverage (Foley and Van Valin 1984; Van Valin and LaPolla 1997; Van Valin 2005). In this light, it rather comes as a surprise that impersonal constructions [ICs] haven't received much attention outside the area of Romance languages (e.g. Bentley 2004, 2006; González Vergara 2009). ICs receive a variety of interpretations, ranging from those with a not fully referential subject to those with no overt subject (Siewierska 2008). This syntactic classification of ICs is in contrast to the functional one in terms of agent defocusing (Divjak and Janda 2008; Malchukov and Ogawa 2011). Against this backdrop, I will propose an RRG account of two types of ICs with no overt PSA in Russian and will argue that these Russian ICs are derived from what I term "anti-linking," an operation that partially or fully undoes the two-stage linking operation at work in the interface between lexical semantics and syntax and contributes to an increase in the degree of event-orientation of the whole clause. I will show that this proposal allows us to provide a unified account of ICs and personal constructions in terms of case assignment and agreement. Finally, I will discuss a few lines of further investigation including how the above account is extended to ICs in Icelandic and some other languages.

References

Bentley, Delia. 2004. Unexpressed arguments: si-constructions in Italian. Proceedings of the 2004 Role and Reference Grammar Conference, pp.17-48.

Bentley, Delia. 2006. Split Intransitivity in Italian. Berlin and New York: Mouton de Gruyter.

Divjak, Dagmar and Laura A. Janda. 2008. Ways of attenuating agency in Russian. Transactions of the Philological Society 106(2), pp.138-179.

Foley, William A. and Robert D. Van Valin, Jr. 1984. Functional Syntax and Universal Grammar. Cambridge: Cambridge University Press.

González Vergara, Carlos. 2009. One rule to rule them all: logical structures for Spanish non reflexive se sentences. In Lilián Guerrero, Sergio Ibáñez, and Valeria Belloro (eds.), Studies in Role and Reference Grammar, pp.361-379. México: UNAM.

Malchukov, Andrej and Akio Ogawa. 2011. Towards a typology of impersonal constructions. In Andrej Malchukov and Anna Siewierska (eds.), Impersonal Constructions: A Cross-linguistic Perspective, pp.19-56. Amsterdam and Philadelphia: John Benjamins.

Siewierska, Anna. 2008. Introduction: impersonalization from a subject-centred vs. agent-centred perspective. Transactions of the Philological Society 106(2), pp.115-137.

Van Valin, Robert D., Jr. 2005. Exploring the Syntax-Semantics Interface. Oxford: Oxford University Press.

Van Valin, Robert D., Jr. and Randy J. LaPolla. 1997. Syntax: Structure, Meaning, and Function. Cambridge: Cambridge University Press.

Plenary Lecture

Operator, information: Revisiting the operator projection in RRG, with special emphasis on tense, aspect, and finiteness

Jürgen Bohnemeyer
University at Buffalo

This presentation tackles a wide range of questions surrounding the treatment of functional categories ('operators') in Role and Reference Grammar (RRG; Van Valin 2005). I begin by proposing a typology of operators and an evolutionary account of their grammaticalization. This model correctly predicts that operators are not universal, but nevertheless cross-linguistically recurrent. It also imposes severe restrictions on the possible occurrence of covert (unexpressed) operators.

Operator projections capture three types of behavioral properties of operators: (i) their taking particular syntactic units as operands (in RRG, these units are spelled out by the Layered Structure of the Clause model); (ii) the partial dissociation between their operand and their morphological realization as a result of grammaticalization; and (iii) scope relations among operators. I argue that the combination of (ii) and the complex relational semantics of some operators infuses operator projections with a certain degree of fluidity.

The remainder of the presentation probes into the complexities of the tense-aspect-mood domain as a test case, drawing mostly on data from English, Russian, and Yucatec Maya. I propose to incorporate into RRG a unified neo-Reichenbachian theory of tense and viewpoint aspect (and mood) consistent with what has emerged as a consensus model in the temporal semantics literature. The most prominent exponents of this family of approaches are Kamp & Reyle (1993: 483-690) and Klein (1994); an LF-based version was proposed by Demirdache and Uribe-Etxebarria (2004, 2007). The RRG treatment I develop follows Klein's approach most closely.

I assume an informal semantic type system in which both nuclei and cores express situation *types* (i.e., *kinds* of situations/events; Parsons 1990), whereas finite clauses encode propositions about individual situations. The nucleus expresses the relational content of the situation type description and the other core constituents add descriptions of the event participants. Sentences express speech acts that stand in an aboutness relation to the situation described by their matrix clauses; this is the sentence's 'topic situation' (Austin 1950; Klein 2009; Kratzer 2014). Topic situations are anaphorically tracked in discourse.

Following the terminology of Smith (1991), I distinguish between situation aspect, which comprises the temporal properties associated with the situation type description expressed by the nucleus and core, and viewpoint aspect. While situation aspect is not itself an operator (class), aktionsarten in the sense of Agrell (1908) can be defined as operators on situation-aspectual properties. Even though situation aspect is most commonly fully determined only at the core layer, aktionsarten are primarily nuclear-layer operators.

Viewpoint aspects are operators that map the situation type description expressed by the core to the sentence's topic situation. In doing so, they embed the situation type into a causal chain frame from which they select a portion to be instantiated by the topic situation. Following Bohnemeyer (2002), I propose eight primitive notional viewpoint aspect operators. Due to their complex relational semantics, viewpoint aspects semantically affect the core, clause, and sentence layers. However, I argue that viewpoint aspects are most closely associated with the core layer, since they occur (to a limited extent) in non-finite projections, but require application to a complete situation type description for their interpretation.

Finally, I argue for the treatment of finiteness as a kind of meta-operator in its own right. The primary function of finiteness is to map situation type descriptions to individual situations and, in the case of matrix clauses, to the topic situation. Finiteness plays a role in the verbal/sentential system of

grammar that parallels the role of determination in the nominal system. In both cases, the specific morphological categories tied up with finiteness/determination vary across languages. While tense and mood are inherently linked to finiteness, viewpoint aspect may become inextricably linked to finiteness in individual languages as well – I illustrate with Yucatec data. One advantage of the recognition of finiteness as an operator in its own right is that it offers a solution to the puzzle of so-called 'temporal anaphora' phenomena in tenseless languages such as Kalaallisut (Bittner 2008) and Yucatec (Bohnemeyer 2009). Temporal anaphora is the (quasi-)anaphoric tracking of the topic situation in discourse. Following Partee (1973), much of the literature continues to assume that temporal anaphora is a property of tense markers. The evidence from tenseless languages shows that this analysis cannot be correct. I argue instead that temporal anaphora is a property of finiteness.

From its beginnings, RRG was designed as a theory, not merely of syntax, but of the syntax-semantics interface. Operators are a key component of this interface. This lecture aspires to contribute toward greater attention to them.

References

Agrell, Sigurd. (1908). Aspektänderung und Aktionsartbildung beim polnischen Zeitwort [Aspect marking and aktionsart formation in the Polish verb]. Lund: Ohlsson.

Austin, John L. (1950). Truth. Proceedings of the Aristotelian Society 24(1): 111–129.

Bittner, Maria. (2008). Aspectual universals of temporal anaphora. In Susan Rothstein (ed.), Theoretical and crosslinguistic approaches to the semantics of aspect. Amsterdam: Benjamins. 349-385.

Bohnemeyer, Jürgen. (2002). The grammar of time reference in Yukatek Maya. Munich: Lincom Europa.

---- (2009). Temporal anaphora in a tenseless language. In Wolfgang Klein & Ping Li (eds.), The expression of time in language. Berlin: Mouton de Gruyter. 83-128.

Demirdache, Hamida, & Myriam Uribe-Etxebarria. (2004). The syntax of time adverbs. In Jacqueline Guéron & Alexander Lecarme (eds.), The syntax of time. Cambridge, MA: MIT Press. 217–234.

----. (2007). The syntax of time arguments. Lingua 117: 330–366.

Kamp, Hans & Uwe Reyle. 1993. From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory. Dordrecht: Kluwer.

Klein, Wolfgang. (1994). Time in language. London: Routledge.

---- (2009). Finiteness, Universal Grammar, and the language faculty. In Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura, & Seyda Ozcaliskan (eds.), Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin. New York: Psychology Press. 333-344.

Kratzer, Angelika. (2014). Situations in natural language semantics. In The Stanford encyclopedia of philosophy, ed. Edward N. Zaltahttp://plato.stanford.edu/archives/spr2014/entries/situations-semantics. Last accessed 02/17/2015.

Parsons, Terrence. (1990). Events in the Semantics of English. Cambridge, Mass: MIT Press.

Partee, Barbara. (1973). Some structural analogies between tenses and pronouns in English. The Journal of Philosophy 70: 601–609.

Smith, Carlota S. (1991). The parameter of aspect. Dordrecht: Kluwer.

Van Valin, Robert D. Jr. (2005). Exploring the syntax-semantics interface. Cambridge: Cambridge University Press.

Presentation Abstracts (alphabetical by first author's last name)

Chart-based RRG parsing for automatically extracted and hand-crafted RRG grammars

David Arps, Tatiana Bladier & Laura Kallmeyer
Heinrich Heine University of Düsseldorf

Introduction. During the last decade, several parsing algorithms for RRG have been proposed. (Guest, 2008) implemented a chart-based parser which uses RRG templates (van Valin, 2005) instead of tree-rewriting rules in order to better capture semantic information. (Diedrichsen, 2008) develops an RRG parser for German using a set of descriptions for syntactic constructions in German along with a richly annotated lexicon. (Cortes-Rodriguez, 2016) proposed an approach to incorporate parsing for RRG into a natural language understanding application "Artemis" based on a lexical construction model. The described approaches are, however, not able to deal with long-distance dependencies and two of them are either language- or application-specific (Cortes-Rodriguez, 2016; Diedrichsen, 2008). In this paper we propose a general parsing algorithm for RRG based on Tree-Wrapping Grammar (Kallmeyer, Osswald, & van Valin, 2013; Osswald & Kallmeyer, 2018),which shares with (Guest, 2008) the idea of a template-based chart parsing. Our parser can be applied to hand-crafted or to computationally induced RRG grammar fragments of various languages. We discuss first parsing results, and we also show an approach to automatically extract a grammar for parsing, that could then be extended with probabilities in order to be used in a data-oriented approach with a statistical chart-based parser.

RRG elementary trees and their syntactic composition. Following (Kallmeyer et al., 2013; Osswald & Kallmeyer, 2018), we adopt a formalization of RRG as a tree-rewriting grammar involving the tree composition operations *substitution* (for argument slot filling), *wrapping substitution* (for argument slot filling combined with "extraction") and *sister adjunction* (for adding operators and periphery elements among others). In other words, starting from a set of elementary trees, larger trees are generated by these operations. Fig. 1 provides an example that involves only substitution (the filling of the two NP argument slots by the elementary trees of "*average*" and "*points*" respectively, indicated by solid arrows) and sister adjunction (adding the operators "*the*" and "*had*" and the cardinal "*27*", indicated by dashed arrows). We choose to avoid crossing branches in cases of mismatches between operator projection and constituency structure (and, similarly, for the periphery). Instead, if needed, operators attach lower. We provide a feature OP on the operator tree that signals its attachment layer in the operator projection, thereby, the intended RRG structure can be retrieved. An example is "*had*" in Fig. 1a, which contributes a clausal operator (OP=cl) but comes with a CORE adjunct tree. In the last step, this operator is re-attached to its intended position as a daughter of CLAUSE. Fig. 1b gives the resulting RRG tree (with the operator projection integrated into the constituency tree).

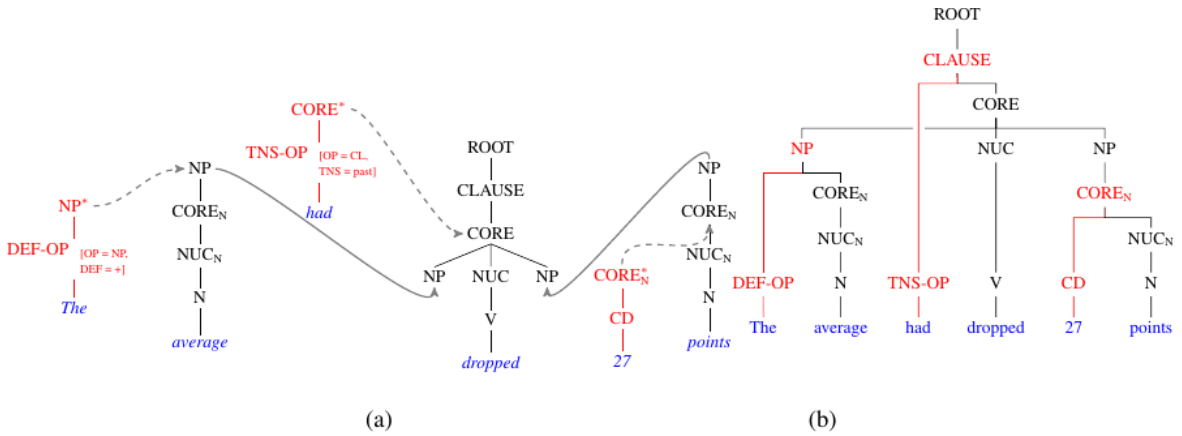


Figure 1: Composition of RRG elementary trees (a) and a resulting RRG structure (b).

Elementary trees can be enriched with *feature structures on the nodes* (see the feature [OP=NP] in Fig. 1a) and *features on the edges* that restrict adjunction possibilities during parsing. These can be used for instance to guarantee that operators attach in the correct order [see (Kallmeyer & Osswald, 2017), for more details]. The elementary trees of a grammar can be systematically described within a hand-written grammar or they can be induced from syntactically annotated data. In this paper we pursue the latter approach but the parser can also be used for hand-written grammar fragments.

Chart-based RRG parsing with elementary trees with features. In order to perform parsing with implemented RRG grammars, we developed a CYK-style bottom-up chart parser. Input to the parser are a set of elementary trees and a sentence to parse. The parser returns all derivations for the sentence that can be derived by combining the elementary trees. Node and edge features are stored during parsing. The algorithm first scans the lexical items. From these, it generates the derived RRG constituency tree bottom-up and aims at reaching one or more root nodes that dominate the whole input. Tree com-position operations trigger unification of the feature structures at the nodes at which elementary trees combine. Unification of edge feature structures is the last step of parsing. If unification fails at least once, the derivation is discarded. The parsing algorithm is implemented as an extension of TuLiPA [1], a parsing environment for Tree Adjoining Grammar (TAG) and Frame Semantics (Kallmeyer & Osswald, 2017; Osswald & Kallmeyer, 2018).

Automatic induction of RRG grammar and parsing experiments. We perform a rule-based automatic induction of elementary trees on the RRGbank (Bladier et al., 2018a), an English treebank, in order to extract an English RRG. The trees in this resource follow a special notational variant of RRG, which includes the operator projection in the constituent projection (an example of this notational variant is shown in the RRG structure in Fig. 1b; [see (Bladier et al., 2018a), for more details]). The tree-extraction algorithm uses heuristic percolation tables to distinguish arguments from modifiers along with rules for adding feature information and follows the top-down extraction approach proposed by (Xia, 1999) for TAG.

When the induced elementary trees without any feature structures are used for parsing, the parser over-generates, i.e., besides the correct analysis, it also yields ungrammatical constituency trees. In order to avoid this, we equip the elementary trees with feature structures that can be obtained automatically during elementary tree induction. These feature structures contain linguistic information, such as attachment levels of operators and peripheral elements. We evaluate the usefulness of these features by counting how many ungrammatical parsing results they eliminate. Our special interest will be (i) modelling the operator projection in complex sentences and (ii) raising and control constructions that are modelled by wrapping substitution.

Outlook. In future work, we plan to use the parser both for data-driven parsing with probabilistic large coverage grammars as well as for symbolic parsing of manually developed precision grammar

fragments. Concerning the former, we will automatically induce probabilistic grammars for various languages and apply a combination of supertagging, dependency parsing and A∗ chart parsing along the lines of (Bladier, van Cranenburgh, Samih, & Kallmeyer, 2018b; Waszczuk, 2017).We also plan to abstract away from the lexical items in the induced elementary trees. By extracting elementary tree templates based on the POS-tags provided by the treebank, the coverage of the extracted grammar can be increased. Our parsing approach can be extended to cover a variety of typologically different languages, provided the existence of sufficiently large suitable resources (such as, for example, RRGbank (Bladier et al., 2018a)). Concerning precision grammars, we will use the parser as a means to test RRG-analyses of specific phenomena via hand-written grammar implementations.

References

Bladier, T., van Cranenburgh, A., Evang, K., Kallmeyer, L., Möllemann, R., & Osswald, R. (2018a). RRGbank: a Role and Reference Grammar Corpus of Syntactic Structures Extracted from the Penn Treebank. In D. Haug, S. Oepen, L. Øvrelid, M. Candito, & J. Hajič (Eds.), *Proceedings of the 17th International Workshop on Treebanks and Linguistic Theories (TLT 2018).* Linköping University Electronic Press, Sweden.

Bladier, T., van Cranenburgh, A., Samih, Y., & Kallmeyer, L. (2018b). German and French Neural Supertagging Experiments for LTAG Parsing. In *Proceedings of ACL 2018, SRW* (pp. 59–66).

Cortes-Rodriguez, F. J. (2016). Towards the computational implementation of Role and Reference Grammar: Rules for the syntactic parsing of RRG Phrasal constituents. *Cırculo de lingüıstica aplicada a la comunicación, 65*, 75–108.

Diedrichsen, E. (2008). A Role and Reference Grammar parser for German. In R. D. van Valin (Ed.), *Studies in Language Companion Series. Investigations of the Syntax–Semantics–Pragmatics Interface* (pp. 105–142). Amsterdam: John Benjamins Publishing Company.

Guest, E. (2008). Parsing for Role and Reference Grammar. In R. D. van Valin (Ed.), *Studies in Language Companion Series. Investigations of the Syntax–Semantics–Pragmatics Interface* (pp. 435–453). Amsterdam: John Benjamins Publishing Company.

Kallmeyer, L., & Osswald, R. (2017). Combining Predicate-Argument Structure and Operator Projection: Clause Structure in Role and Reference Grammar. *Proceedings of the 13th International Workshop on Tree Adjoining Grammars and Related Formalisms*, 61–70.

Kallmeyer, L., Osswald, R., & van Valin, R. D. (2013). Tree Wrapping for Role and Reference Grammar. In G. Morrill & M.-J. Nederhof (Eds.), *Lecture Notes in Computer Science / Theoretical Computer Science and General Issues: v.8036. Formal Grammar. 17th and 18th International Conferences, FG 2012 Opole, Poland, August 2012, Revised Selected PapersFG 2013 Dsseldorf, Germany, August 2013, Proceedings* (pp. 175–190). Berlin/Heidelberg: Springer Berlin Heidelberg.

Osswald, R., & Kallmeyer, L. (2018). Towards a formalization of Role and Reference Grammar. In R. Kailuweit, L. Künkel, & E. Staudinger (Eds.), *Applying and Expanding Role and Reference Grammar* (pp. 355–378). Freiburg: Albert-Ludwigs-Universität, Universitätsbibliothek. [NIHIN studies].

van Valin, R. D. (2005). *Exploring the syntax-semantics interface.* Cambridge: Cambridge University Press.

Waszczuk, J. (2017). *Leveraging MWEs in practical TAG parsing : towards the best of the two worlds*. PhD Thesis. Unpublished manuscript.

Xia, F. (1999). Extracting Tree Adjoining Grammars from Bracketed Corpora. *Proceedings of the 5th Natural Language Processing Pacific Rim Symposium (NLPRS-99)*, 398–403.

Information-structurally driven syntactic configurations

Kata Balogh
Heinrich-Heine-Universität Düsseldorf

We propose an analysis of the syntax-information structure interface in discourse-configurational languages, without imposing functional positions in the syntax. The analysis is based on a formalized Role and Reference Grammar (RRG; Van Valin 2005, Osswald & Kallmeyer 2018), extended with frame-semantic representations (Kallmeyer & Osswald 2013, Löbner 2015).

Hungarian is a well-known *discourse-configurational* language (É. Kiss 1995, Surányi 2015) with flexible word order. As such, it is a challenging language for any formal analysis on the syntax-semantics-pragmatics interface. Hungarian has a rich morphology (with extended case marking and verbal inflection) and these morphological devices, rather than syntactic configurations, code most grammatical information. However, it is not the case that in Hungarian, the word order can be changed freely without consequences for the sentence interpretation. In Hungarian, the surface order in the preverbal field is directly linked to the information structure of the sentence. The postverbal field has free word order, the relative order of the arguments after the verb does not encode grammatical differences, the postverbal word order variations are all associated with the same semantic content. On the other hand, the word order in the preverbal field is fixed and determined by the discourse-semantic functions *topic* and *focus*. Topicalized elements occupy a clause initial position (1), and the narrow identificational focus is placed in the immediate preverbal position and triggers inverse order of the verbal particle [VPRT] and the verb (2).

(1) *What about Mary?*
Mari-nak oda-adta Péter az almá-t.
Mary-DAT VPRT-gave Peter the apple-ACC
'Peter gave Mary the apple.' (≈ To Mary, Peter gave the apple.)

(2) *What did Peter give to Mary?*
Péter az almá-t adta oda Mari-nak
Peter the apple-ACC gave VPRT Mary-DAT
'Peter gave Mary the APPLE.' (≈ It is the apple that Peter gave to Mary.)

As shown above, the topic-focus structure of the utterance determines the surface structure and the order of the preverbal elements in Hungarian, motivating the distinction between topic and focus positions in the syntactic representation, as proposed by different generativist accounts (e.g. É. Kiss 2004). These positions are reserved for restricted semantic-pragmatic functions as aboutness topic, understood in terms of Gundel (1988) and Lambrecht (1994), and narrow identificational focus (É. Kiss 1998). On the other hand, the notion of pragmatic focus (Lambrecht 1994) is also highly relevant in Hungarian. Consider, for example, the additive particle *is* 'also, too' that does not associate with the structural focus, but still considered as focus sensitive, as argued by Balogh (to appear) based on context and by Balogh & Langer (unpublished) based on prosody.

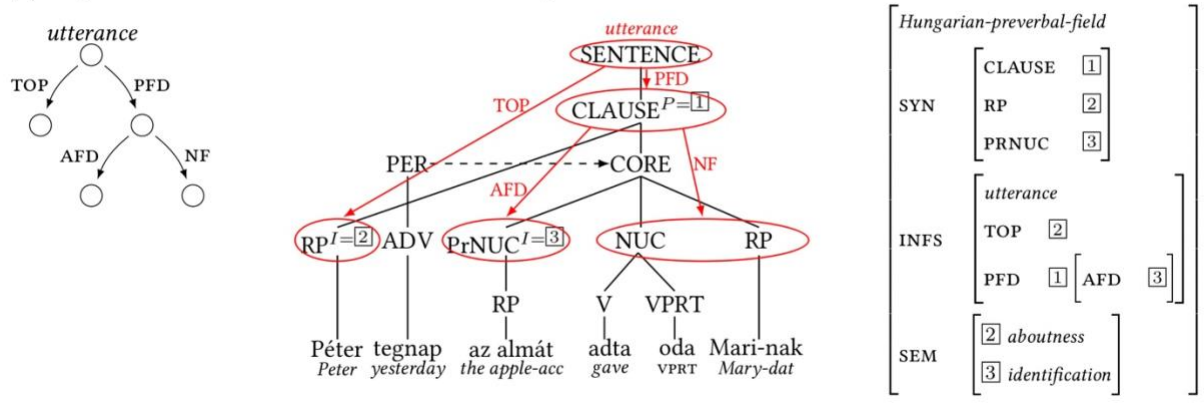
Contrary to the generativist accounts, we capture the syntax-information structure interface without syntactic functional positions, worked out in the framework of RRG (Foley & Van Valin 1984, Van Valin 2005). RRG is a highly suitable grammatical theory to capture various interface phenomena, given that the interaction of syntax, semantics and pragmatics plays a primary role in its formal architecture and theoretical principles. The main advantage of the RRG framework for the analysis of our data is that the necessary pragmatic mechanisms are already present in the theory. Therefore, it offers a straightforward implementation, as opposed to generativist accounts that require the development of a pragmatic mechanism on top of the syntax. The main goal of this work is twofold.

Firstly, we intend to propose an analysis of the Hungarian data and secondly, we propose an extension of RRG to account for discourse-configurationality across languages.

We argue that the structure-building aspect of information structure is distinct from its (semantic-) pragmatic interpretation aspect. In discourse-configurational languages, information structure has a structure-building effect. In these languages, certain information structural functions (topic and/or focus) are marked structurally via designated syntactic positions, as illustrated before in Hungarian, where the constituent in the clause initial position expresses *aboutness topic* (1) and the constituent in the immediate preverbal position expresses *identificational focus* (2). For a frame-based analysis of such structure building effects of information structure, we propose an information structure frame [IS-frame] as an overlay on the syntactic structure (3b). The overlay is understood as a direct linking of IS-functions and syntactic positions. This linking is determined by constructional schemas and replace the focus projection (Van Valin 2005).

The essence of the analysis of the Hungarian preverbal field is illustrated in (3). The generalized abstract IS-frame (3a) represents general notions of information structure: (a) the potential focus domain [PFD], the syntactic domain where the focus of the utterance can occur, (b) the actual focus domain [AFD], the syntactic domain that corresponds to the focus (domain) in Lambrecht's terms and (c) the sentence topic [TOP], often associated with aboutness. The syntactic domains AFD and PFD correspond to one or more information units, the minimal phrasal units in the syntactic representation (Lambrecht 1994). The (non-topical) part of the utterance that is inside the PFD but outside of the AFD is represented as non-focus [NF], corresponding to the notion of tail in Vallduví's (2016) terms. The constructional schema (3c) determines the linking between syntax, information structure and semantics (3b). Hereby the interface variables (represented by boxed numbers) are of great importance. The interface features appear on the syntactic nodes as well as in the representations of the semantic content and the information structure of the sentence. As illustrated in the constructional schema of the Hungarian preverbal field below, the linking is determined as: (a) the referent of the constituent in the sentence initial RP is considered as the topic and interpreted in terms of aboutness and (b) the referent of the constituent in the PrNUC position is considered as the focus and interpreted in terms of identification (also deriving exhaustivity). We argue that the IS-overlay analysis is cross-linguistically applicable, to represent both universal notions and language specific configurations.

(3) a. generalized IS-frame b. IS-overlay c. constructional schema



References

Balogh, K. to appear. Focus sensitivity in Role and Reference Grammar. In Van Hooste et al. (eds). *Proceedings of the 2015 RRG Conference*.

Balogh, K. & C. Langer. unpublished. *Additive particles: syntax-prosody interface and interpretation*. Talk at the 14th International Conference on the Structure of Hungarian. June 11-12, 2019. Potsdam, Germany.

É. Kiss, K. (ed). 1995. *Discourse Configurational Languages*. Oxford: University Press.
É. Kiss, K. 1998. Identificational focus versus information focus. *Language* 74.
É. Kiss, K. 2004. *The Syntax of Hungarian*. Cambridge: University Press.
Foley, W. A. & R.D. Van Valin, Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: University Press.
Gundel, J. 1988. Universals of topic-comment structure. In Hammond et al. (eds). *Studies in Syntactic Typology*. Amsterdam: John Benjamins.
Kallmeyer, L. & R. Osswald. 2013. Syntax-Driven Semantic Frame Composition in Lexicalized Tree Adjoining Grammars. *Journal of Language Modelling*. 1(2).
Lambrecht, K. 1994. *Information Structure and Sentence Form*. Cambridge: University Press.
Löbner, S. 2015. Functional Concepts and Frames. In Gamerschlag et al. (eds). *Meaning, Frames, and Conceptual Representation*. Düsseldorf University Press.
Osswald, R. & L. Kallmeyer. 2018. Towards a formalization of Role and Reference Grammar. In Kailuweit et al. (eds). *Applying and expanding Role and Reference Grammar*. Freiburg: Albert-Ludwigs-Universität, Universitätsbibliothek.
Surányi, B. 2015. Discourse-configurationality. In Féry & Ishihara (eds). *The Oxford Handbook of Information Structure*. Oxford: University Press.
Vallduví, E. 2016. Information structure. In Aloni & Dekker (eds). *The Cambridge Handbook of Formal Semantics*. Cambridge: University Press.
Van Valin, Jr., R. D. 2005. *Exploring the syntax-semantics interface*. Cambridge: University Press.

Formalizing hierarchy-related generalizations in RRG: The brittle and the supple

Hanno T. Beck
University at Buffalo

I demonstrate several instances where sentences that "should" be ungrammatical according to conventional syntax, are not. Their acceptability seems to hinge on something more than syntax alone. Problematic cases of this sort can be optimally handled in a framework such as Role and Reference Grammar (Van Valin and LaPolla 1997, Van Valin 2005), that is sensitive to known hierarchies, for instance a thematic relations hierarchy (Van Valin 2005, pp. 53-61). Predictions of a sentence's acceptability, if they can be partly informed by a hierarchy or continuum, are more realistic than those based on a brittle black-and-white "can the grammar generate this" criterion. I propose a formalization of continuum-sensitive semantic generalizations, as a part of the RRG constructional schema system.

Under many approaches, syntax-only analyses make generalizations that are unavoidably brittle – any string is either licensed by the syntax, or is not, and other considerations have no way to play a role. For such an inflexible analysis, exceptions can spell disaster. But it is these exceptions that can yield informative data, pointing us toward a more nuanced understanding that embraces not only syntax but also semantic and pragmatic insights.

A syntactic framework such as Role and Reference Grammar, by availing itself of descriptive hierarchies/continua, often of a semantic sort, is able to express generalizations that are less brittle and more supple.

For a typical area where purely syntactic generalizations fail, consider the tough construction in English as in (1):

(1) This jar is tough to open.

With the tough construction, various syntactic claims have been made that underestimate or even ignore the role of semantic/pragmatic considerations. I will present examples of several published syntactic generalizations that, understood in black-and-white terms, are simply false – they have exceptions. But all of them involve preferences based on some meaningful continuum or hierarchy. Where brittle generalizations fail, supple ones can succeed.

Here is a nonexhaustive sampling of some of these brittle claims:

(a) Tough constructions are supposed to be ungrammatical if the embedded clause has an expletive "there" subject (e.g. Bresnan 1971);
(b) Tough constructions are alleged to require intentionality or animacy on the part of the (covert) embedded subject (e.g. Lasnik & Fiengo 1974);
(c) Tough constructions are supposed to be incompatible with a nongeneric indefinite matrix subject (e.g. Postal 1971);
(d) Tough constructions are supposed to be limited to a certain group of adjectives (and some nominals and verbs) that subcategorize for it (discussed in Fleisher 2015);
(e) Tough constructions are alleged to be incompatible with embedded non-intentional *be* (discussed in Jones 1991).

In each case, the brittle generalization can be improved by aligning it with an independently motivated semantic continuum, predicting which examples will sound better or worse, thus embracing the apparent exceptions. The exceptions that I have found to (a)-(e) above stem from and reflect speaker preferences that are describable in terms of these continua. Several continua or hierarchies are in play in (a)-(e); animacy, thematic relations, definiteness, topic vs. focus, etc., as I will show in detail.

RRG's constructional schema system gives us a way to let hierarchies contribute. In the semantics section of a constructional schema, reference can be made to relevant considerations, including

hierarchies that play a role. The result is an account augmented by the ability to distinguish good-sounding from bad-sounding examples on the basis of a hierarchy – that is a supple account.

I propose that we begin work on formalizing the semantics section of a constructional schema. In so doing, we can constrain and categorize the kinds of statement permitted in that section, making new predictions about what is possible in grammar, and hinting that some of these semantically-informed constraints may be universal.

With a better mechanism for characterizing generalizations – to include supple, semantically sensitive tendencies – RRG offers a framework for grammar that corresponds more naturally and closely to actual linguistic usage facts. We can perhaps even extend this idea to new areas where relevant hierarchies have not yet been proposed. Whenever we can reconceptualize a generalization's exceptions (or other "unruly data") as exemplifying sensitivity to a continuum or hierarchy, that generalization can be drawn in more supple and more comprehensive terms.

Hierarchy-sensitive grammatical facts should be encoded in the form of preference tendencies; semantic preferences in the grammar are systematic, not a random or haphazard collection. For example, in an RRG system where, for a given construction, the PSA (privileged syntactic argument) must not be Actor (hence, ~A), a useful semantic constraint in some cases will be "and the less Actor-like, the better." Expressing a preference tendency for one end of a continuum versus the other, may turn out to be a common format for such constraints. In any case it can be spelled out so that explicit claims are being made and can be evaluated more easily crosslinguistically. A proposal envisioning hierarchies and preference statements about them to involve partially-ordered sets and set-theoretic functions, will be offered. Such a notation system enables further predictions about what sorts of semantic constraints are expected or unexpected in the world's languages. The supple triumphs over the brittle.

References

Bresnan, J. W. 1971. Sentence Stress and Syntactic Transformations. Language 47:2, pp. 257-281.

Fleisher, N. 2015. Rare-class adjectives in the tough-construction. Language 91:1, pp. 73-108. Jones, C. 1991. Purpose Clauses. Dordrecht: Kluwer.

Lasnik H., & Fiengo, R. 1974. "Complement Object Deletion," Linguistic Inquiry 5, 535-571. Postal, P. 1971. Cross-Over Phenomena, Holt, Rinehart & Winston, New York.

Van Valin, R. D., Jr. 2005. Exploring the syntax-semantics interface. Cambridge: Cambridge University Press.

Van Valin, R. D., Jr. & LaPolla, R. J. 1997. Syntax: structure, meaning & function. Cambridge: Cambridge University Press.

Event integration in the English *by-Ving* (*MEANS*) construction: An RRG analysis

Erika Bellingham
University at Buffalo

Multi-predicate argument-structure constructions (e.g. caused motion, resultative) allow speakers to build descriptions of complex events by integrating the semantics of each predicate with the more abstract semantics of the construction (Goldberg, 1995:5; Langacker, 1991:293-304). The highly productive English *means* construction, exemplified in (1), integrates two (potentially complex) events (**main event** + means event) with the construction's semantics (a *means* relation between the two events). This paper investigates the semantic and syntactic properties of the English *means* construction and considers its position in the Interclausal Relations Hierarchy.

(1) a. **He angered Democrats** by threatening to veto his own budget.
b. **Archy wrote** by hurling himself at the typewriter keys one at a time.
c. **His own wife survived** by clinging to a couple of chairs until she was rescued.
Corpus of Contemporary American English (Davies, 2008-)

Semantics: Although often used to paraphrase the semantics of other constructions (e.g. Goldberg & Jackendoff, 2001:538; Levin, 1993:98; Jackendoff, 1997:554-5), the *means* construction itself has received little attention. The pattern of event integration typically exploited in these paraphrases (in which the means event specifies the underspecified causing subevent of the main event, e.g. (1a)) is however only one of the possibilities for the means construction. A corpus study of this construction reveals three distinct patterns of event integration (summarized in Table 1), distinguished based on the component of the main event frame that they target. The pattern can be predicted based on the main event frame's causal and aspectual structure as well as the salience/cultural relevance of a larger force-dynamic structure (containing the main event frame). A second pattern (1b) requires a durative main event (an activity or accomplishment, or a phase of a larger event). Here, the means event is a fine-grained construal of the entire main event: the relationship between the two events is similar to the discourse-level *elaboration* coherence relation (c.f. Hobbs, 1979). In the third pattern (1c), the main event is restricted to events which are not themselves causally complex, but exist within a highly salient/culturally relevant force-dynamic structure (e.g. *dying, surviving, escaping, learning*). The means construction directs attention to (and the means event provides a fine-grained construal of) the event which precedes the main event in the salient force-dynamic structure: this preceding event would otherwise not have been explicitly mentioned.

Variant	**Main event properties**	**Semantic integration of events**
Specify cause (1a)	Causally complex, with an underspecified causing event	Means event specifies the causing event in the main event
Elaborate activity (1b)	An activity or accomplishment, or a phase of an event	Means event is an elaboration of the activity or phase
Add cause (1c)	Achievement or accomplishment; associated with salient force-dynamic structure	Means event describes the cause of the main event

Table 1: Three variants of the English* means *construction.

I demonstrate that each of three existing semantic analyses of the English *means* construction (Van Valin, 2005; Ohori, 2001; Balkanski, 1992) are insufficient to account for the range of possible meanings that the means construction can contribute. One possible analysis would be to posit three subtypes of the means construction, each with a slightly different semantics, with the appropriate subtype selected based on the causal/aspectual properties of the main event. I argue, however, that these three meanings can be unified with a single semantic analysis, in which the primary semantic contribution of the means construction is to portray the means event as a fine-grained construal (c.f. *scalar adjustment* in Croft & Cruse, 2004; Croft, 2012) of a component of the main event frame. This component must be such that its realization realizes the main event (1b) or ensures its realization in inertial worlds (1a, 1c).

Syntax: I analyse the English means construction as a *core peripheral subordination*, with the *by-Ving* constituent (which expresses the means event) occurring in the periphery of the main core (which expresses the **main event**). This analysis is based on five main pieces of syntactic evidence: it is headed by a preposition *by* (a hallmark of the core periphery); the *means* constituent exhibits ordering freedom (2); other elements in the periphery of the matrix core may intervene between the two units (3a-b); two clauses in a cosubordination construction (e.g. English conjunction reduction, cf. Van Valin, 2005: 231) can each participate in the means construction (4a); and finally the main core in the means construction may be non-finite (5), which is evidence against clausal juncture. The construction is syntactically asymmetrical, as the *means* constituent may itself be a complex clause: in (6), the *means* constituent contains two independent cores.

(2) By analyzing the bomb, **investigators began to develop its profile**.
(3) a. **This number was calculated** *by the authors* by dividing the population by 7.5.
b. **He angered Democrats** *on Wednesday* by threatening to veto his own budget.
(4) **Sandy studied** by reading her notes and **will celebrate finishing her exam** by burning them.
(5) Sandy wanted **to study for her exam** by reading her notes.
(6) On Monday, Sandy **celebrated her graduation** by burning her notes in the morning and drinking beer all afternoon.

***Means* in the Interclausal Relations Hierarchy:** The interclausal relations hierarchy (Van Valin & LaPolla, 1997; Van Valin, 2005) iconically maps a morphosyntactic complexity scale to a conceptual complexity scale of semantic relations. Semantic relations at one end of the scale show a high degree of cohesion, such that the linked units describe facets of the same event or action (rather than two distinct events). The asymmetry of the *means* construction offers an interesting twist on iconicity in clause linkage: from the perspective of the juncture, the two units do indeed describe two facets of a single event (one is a construal of a subpart of another), yet within the means constituent, one of the facets may be construed as multiple distinct events, potentially with separate time-positional modifiers as in (6).

References

Balkanski, Cecile T. (1992). "Action relations in rationale clauses and means clauses". In: *COLING 1992 Volume The 15th International Conference on Computational Linguistics*. Association for Computational Linguistics.

Croft, William. (2012). *Verbs: aspect and causal structure.* Oxford: Oxford University Press.

Croft, William & D. Alan Cruse. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Davies, Mark. (2008-) *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present.* Available online at https://corpus.byu.edu/coca/.

Goldberg, Adele E. (1995). *Constructions: A construction-grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldberg, Adele E., & Jackendoff, Ray. (2004). The English resultative as a family of constructions. *Language 80(3)*: 532-568.
Hobbs, Jerry R. (1979). Coherence and coreference. *Cognitive Science 3*: 67-90.
Jackendoff, Ray. (1997). Twistin' the Night Away. *Language 73(3)*: 534-559.
Langacker, Ronald W. (1991). *Foundations of cognitive grammar*. Volume 2. Stanford, CA: Stanford University Press.
Levin, Beth. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
Ohori, Toshio. (2001) Some thoughts on a new systematization of interclausal semantic relations. Paper presented at 2001 Role and Reference Grammar Conference, University of California, Santa Barbara.
Van Valin, Robert D., Jr. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
Van Valin, Robert D., Jr. and Randy J. LaPolla. (1997). *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Modeling causative complexity across languages with the Interclausal Relations Hierarchy

Erika Bellingham, Pia Järnefelt, Kazuhiro Kawachi, Yu Li, Alice Mitchell, Guillermo Montero-Melis, Sang-Hee Park, Anastasia Stepanova, Manne Bylund & Jürgen Bohnemeyer
University at Buffalo, Stockholm University, National Defense Academy of Japan, University at Buffalo, University of Bristol, Max Planck Institute of Psycholinguistics, University at Buffalo, University at Buffalo, Stockholm University & University at Buffalo

We present a multivariate analysis of acceptability judgment data on causative constructions collected from speakers of nine languages spoken on four continents. The central research question of this study is which semantic variables determine the minimal morphosyntactic complexity necessary for acceptable descriptions of a given causal scenario (or 'chain'). To measure the morphosyntactic complexity of the stimulus descriptions, we employed the juncture-nexus-type scale inherent in the Interclausal Relations Hierarchy (IRH) of RRG.

Motivation of the study - It has long been hypothesized that the minimal morphosyntactic complexity required of a causative description iconically mirrors the conceptual complexity of the represented causal chain (Haiman 1983, McCawley 1978, *inter alia*). For example, (1) is a good description of an event in which Floyd took a hammer to the vase with the goal of breaking it, whereas (2) is more suitable in case Floyd sneezed so hard that the vase fell off a shelf.

(1) *Floyd broke the vase*
(2) *Floyd caused the vase to break*

What is less clear is which semantic factors influence the speaker's selection of syntactic complexity level (e.g., Dixon 2000 proposes 9 semantic variables), how these factors rank in importance, and the extent to which the answer to the first two questions is language-specific. These questions have so far been addressed in studies based on secondary data (Escamilla 2012, Levshina 2016a) and in corpus studies (Haspelmath 2008; Levshina 2015, 2016b, 2017). The sole attempt to date at tackling them on the basis of primary crosslinguistic data has been a small production study (Bohnemeyer et al. 2010). Here, we present (to our knowledge) the first study of the semantic typology of causatives based on acceptability ratings data. This method presents all participants with the same descriptions and described scenarios, captures inter-speaker variation, does not depend on the availability of corpora for the sample languages, and provides not only positive evidence of what speakers actually say in a given situation, but also negative evidence of what they consider unacceptable as a representation of a given causal chain.

Study design - 12+ speakers per language rated descriptions, elicited from L1 speakers of each sample language, of 43 video clips for their goodness of fit. The descriptions crossed the stimulus scenes with a language-specific set of response types featuring a range of causative constructions (cf. Table 1 below). During a training phase, the raters had learned to express through their ratings whether the descriptions were (i) ungrammatical, (ii) grammatical and interpretable but semantically inaccurate of the described scene, (iii) accurate but pragmatically misleading for the scene, or (iv) semantically and pragmatically appropriate for it. The scenes had been designed to manipulate ***causer type*** (intentional actor vs. accidental actor vs. natural force), ***causee/affectee type*** (controlled vs. psychologically impacted vs. physically impacted vs. inanimate), ***mediation*** (the presence vs. absence of an intermediate subevent/participant between cause and effect), and further variables not included in the analysis we present here.

Coding the stimulus descriptions based on the IRH - To assess the manner and level of structural integration of the expressions of causing and resulting events in the descriptions, each description was assigned a juncture-nexus type (JNT) in the Layered Structure of the Clause model (Van Valin 2005). The juncture level of the descriptions (see Table 1) was then used as the basis for the dependent

variable in a series of language-specific analyses. The JNT scale is unique in projecting the relevant aspects of morphosyntactic complexity into a single scale. In phrase structure grammars, the complexity of the causing and resulting event expressions and that of the combination of the two are all independent of one another.

Analysis - For each language, we trained language-specific classification trees using the CART algorithm (Breiman 1984) to predict the juncture of the most compact construction type to receive ceiling ratings given a particular combination of independent variable levels. A Random Forest analysis (Tagliamonte & Baayen 2012) was also performed to provide a more reliable ranking of variable importance. These methods permit an assessment of the relative impact of the candidate predictors without being susceptible to overfitting due to collinearity or sparsely populated cell issues as standard regression models are.

Table 1. *Construction types by language and juncture (AC – Adjunct causer/reason ('because of x'), CC – Causal connective, CV – Converb, MC – Morphological causative, PC – Periphrastic causative, RV – Resultative construction (incl. resultative-type serial verb construction), SC - Scalar Connective construction ('So x that y'), TC – Transitive causative verb)*

Juncture level Language	Field site	Simplex or nuclear-layer	Core-layer	Clause-layer
Datooga (Nilotic)	Tanzania	MC, TC	AC, PC, SC	CC
English (Germanic)	United States	RV, TC	PC	AC, CC, SC
Japanese (Japonic)	Japan	MC, TC	AC	CC
Korean (isolate)	South Korea	MC, RV, TC	PC	CC, CV
Russian (Slavic)	Russia	TC	PC	AC, CC, SC
Sidaama (Cushitic)	Ethiopia	MC, TC	AC, PC	CC
Swedish (Germanic)	Sweden	RV, TC	PC	CC, SC
Yucatec (Mayan)	Mexico	MC, TC	PC	CC
Zauzou (Loloish)	China	RV	CC, CV, PC	CC

Preliminary findings and implications - The sample languages form a continuum in terms of the proportion of scenes for which clausal junctures are the only acceptable response type for most participants. For the Japanese and Korean participants, this was the case for more than half of the scenes tested. Conversely, for the Datooga, Sidaama, Yucatec, and Zauzou speakers, this was not the case with even a single scene. The three European groups assumed intermediate positions along this cline. The participants also differed by language in terms of the conditions under which they would accept compact descriptions (simplex-nucleus or nuclear junctures): speakers of the European languages and Yucatec (and, in first approximation, Korean) rejected such descriptions whenever an intermediate event/participant was involved in the scene. In contrast, Japanese speakers accepted compact descriptions of such scenes, but tended to reject them when the causer was accidental or a natural force. It seems plausible that Japanese speakers prefer to avoid direct causal attribution when referring to such scenes (cf. Fausey et al. 2010). (Inter-speaker variation precludes a clear picture of the Datooga, Sidaama, and Zauzou data in this regard.) Tentatively, English and Yucatec emerge as overall mediation-dominant, Sidaama as causer-type-dominant, Datooga, Japanese, and Korean as causee/affectee-type-dominant, and Russian and Swedish as being dominantly sensitive to both mediation and causee/affectee-type. Our study also showcases the usefulness of the IRH as a tool for measuring morphosyntactic complexity, including in, but not restricted to, typological research.

References

Bohnemeyer, J., N.J. Enfield, J. Essegbey, & S. Kita. (2010). The Macro-Event Property: The segmentation of causal chains. In *Event representation in language: Encoding events at the*

language-cognition interface, eds. Jürgen Bohnemeyer and Eric Pederson, 43–67. Cambridge: Cambridge University Press.
Breiman, L. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.
Dixon, R.M. (2000). A typology of causatives: form, syntax and meaning. In *Changing valency: Case studies in transitivity*, eds. Robert M. W. Dixon and Alexandra Y. Aikhenvald, 30--83. Cambridge: Cambridge University Press.
Escamilla Jr, R.M. (2012). *An updated typology of causative constructions: Form-function mappings in Hupa (Californian Athabaskan), Chungli Ao (Tibeto-Burman) and Beyond*. PhD Dissertation, University of California, Berkeley.
Fausey, C. M., B. L. Long, A. Inamori, & L. Boroditsky. (2010). Constructing agency: the role of language. *Frontiers in Psychology* 1: 162. doi:10.3389/fpsyg.2010.00162.
Haiman, J. (1983). Iconic and economic motivation. *Language* 59(4): 781–819.
Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1): 1-33.
Levshina, N. (2015). European analytic causatives as a comparative concept: Evidence from a parallel corpus of film subtitles. *Folia Linguistica* 49(2): 487–520.
Levshina, N. (2016a). Finding the best fit for direct and indirect causation: a typological study. *Lingua Posnaniensis* LVIII(2): 65–82.
Levshina, N. (2016b). Why we need a token-based typology: A case study of analytic and lexical causatives in fifteen European languages. *Folia Linguistica* 50(2): 507–542.
Levshina, N. (2017). Measuring iconicity: A quantitative study of lexical and analytic causatives in British English. *Functions of Language* 24(3): 319–347.
McCawley, J. (1978). Conversational implicature and the lexicon. In S*yntax and semantics IX: Pragmatics*, ed. Peter Cole, 245-258. New York, NY: Academic Press.
Tagliamonte, S. A. and Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language variation and change*, 24(2):135–178.
Van Valin Jr, R. D. (2005). Exploring the syntax-semantics interface. Cambridge: Cambridge University Press.

The logical structure of verbs of quantized and non-quantized change

Delia Bentley
The University of Manchester

It has long been known that there are two principal classes of accomplishment verbs, which are sometimes referred to as verbs of quantized and non-quantized change (Hay et al. 1999; Beavers 2013; see also Dowty's (1979: 88-90) degree achievements and Bertinetto & Squartini's (1995) verbs of gradual completion). With verbs of quantized change the progressive does not entail the perfect (cf. 1a), whereas with verbs of non-quantized change it does (cf. 1b).

(1a) The soldier is dying ≠> the soldier has died.
(1b) The child is growing => the child has grown.

The verbs exemplified in (1a) are *telic*, in the sense that they entail a specific final goal state (Beavers 2013). The same is true of *die, be born, fall, arrive*, and many verbs of inherently directed motion (Levin 1993). Telic verbs are incompatible with *completely*, which is redundant, and with *not completely*, given that they have a specific final state as one of their entailments.

(2)The soldier has died #completely / #but not completely.

According to Hay et al. (1999), with verbs of non-quantized change telicity can arise as an implicature, when the adjectival base of a deadjectival verb (e.g., *empty, straighten*, etc.) entails a bounded difference value, which is a measure of the amount to which an argument of the verb changes with respect to a gradable property. The verbs that have this kind of adjectival base test out as telic in accordance with some diagnostics (cf. 3a), although they are also compatible with *completely* and *not completely*.

(3a) Mary is straightening the antenna ≠> Mary has straightened the antenna.
(3b) Mary has straightened the antenna completely / but not completely.

Telicity can also arise as an implicature based on a conventional property of the undergoer. This can be the case with deadjectival verbs whose basis does not entail a bounded difference value. These verbs test out as non-telic (4a) and are compatible with *not completely* (4b). The implicature of telicity, in the above sense, can be cancelled (cf. 4b).

(4a) Mary is shortening the rope => Mary has shortened the rope.
(4b) Mary has shortened the uniform, but not completely.

We address the question of how these differences in the encoding of scalar change (Rappaport Hovav 2008) ought to be represented in the RRG system of lexical decomposition. Our point of departure is Van Valin's (2005) proposal that the operator BECOME should be broken down into the process operator PROC and the result state operator INGR. This proposal allows RRG to capture the difference between, on the one hand, verbs of scalar change that describe a change with no end result and, on the other, verbs that lexicalise an end result as part of their meaning. This contrast has a transparent morphological manifestation in some languages, for example, Mparntwe Arrernte and Lakhota. A verb that describes a process of cooling with no lexicalized end result is represented as in (5a), whereas one that describes a telic process of cooling is represented as in (5b).

(5a) PROC **cold'** (x) (see Mparntwe Arrernte *irrernte*+*irre*) (Van Valin 2005: 44)

(5b) PROC **cold'** (x) & INGR **cold'** (x) (see Mparntwe Arrernte *irrernte+arle+irre*)

The same principle is applied in the representation of active accomplishments put forward in Van Valin (2018). Thus, activity *run*, which does not entail an endpoint, is represented as in (6a), whereas the corresponding active accomplishment is represented as in (6b), which includes the end state of being at a location (z).

(6a) [**do'** (x, [**run'** (x)]) ∧ PROC **cover.path.distance'** (x, (y))]
(6b) [**do'** (x, [**run'** (x)]) ∧ PROC **cover.path.distance'** (x, (y))] & INGR **be-at'** (z, x)

Adopting this system, the class of *die* can be represented as in (7), whereas the representation of *straighten* could in principle vary between the two alternatives in (8a-b), with (8b) including the representation of a final goal state arising from the adjectival base.

(7) PROC **die'** (x) & INGR **dead'** (x)
(8a) PROC **straight'** (x)
(8b) PROC **straight'** (x) & INGR **straight'** (x)

Our talk addresses some challenges to the analyses in (7)-(8a-b) and proposes a tentative solution. We ask if we should we only break down BECOME into PROC+INGR in the cases illustrated in (5a-b, 6a-b) and if any verbs of scalar change require the operator PROC *alone* in their lexical representation (cf. 5a). Lastly, we address the issue of whether we should differentiate quantized vs. non-quantized change in LS.

We discuss evidence in favour of the breaking down of BECOME into PROC+INGR in the lexical representation of *all* verbs of scalar change. Important evidence is provided by the selection of the perfect auxiliary in Italian. In this language, 'be' is selected with monovalent states, achievements and accomplishments (cf. 9a), whereas 'have' is selected with activities (cf. 9b) (Centineo 1986, Van Valin 1990, Bentley 2006).

(9a) Il treno è salito (su per la collina in due ore). (Italian)
the train be.3sg gone.up up for the hill in two hours
'The train went up (the hill in two hours).'
(9b) Il treno ha sobbalzato (per ore).
the train have.3sg jolted for hours
'The train jolted (for hours).'

Given that monovalent verbs of quantized and non-quantized change select 'be', INGR **pred'** must be in the LS of both types.

(10a) La bomba è scoppiata. [Quantized change] (Italian)
the bomb be.3sg exploded
'The bomb exploded.'
(10b) La paura è cresciuta. [Non-quantized change]
the fear be.3sg grown
'The fear has grown.'

Given that some verbs of non-quantized change allow activity (process) readings, in which case they select 'have', PROC must be part of their LS.

(11) La frutta ha marcito per/*in una settimana. [Process]
the fruit have.3sg rotted for in a week
'The fruit has rotted / has been rotting for a week.'

Therefore, we put forward the following proposal: the LS in (5a) (cf. 12a) should only be adopted for verbs describing a process and entailing the absence of an end result. The LS in (7) (cf. 12b) should be adopted for verbs of quantized change, where **pred'** (x) represents a lexicalized *specific* final goal state. Instead, we propose the LS in (12c) for the verbs of non-quantized change, where INGR **pred α'** (x) indicates that a final state exists, but the verb does not entail a specific final goal state.

(12a) PROC **pred'** (x) [verbs which only lexicalize the process]
(12b) PROC **pred'** (x) & INGR **pred'** (x) [verbs of quantized change]
(12c) PROC **pred'** (x) & INGR **pred α'** (x) [verbs of non-quantized change]

References

Beavers, John (2013), Aspectual classes and scales of change, *Linguistics* 51(4), 681-706.

Bentley, Delia (2006), *Split intransitivity in Italian*. Berlin: Mouton de Gruyter.

Bertinetto, Pier Marco. & Squartini, Mario (1995), An attempt at defining the class of gradual completion verbs, in P.M. Bertinetto, V. Bianchi, J. Higginbotham & M. Squartini (eds), (1995) *Temporal Reference, Aspect and Actionality*, Turin: Rosenberg & Sellier, 11-28.

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

Centineo, Giulia (1986), A lexical theory of auxiliary selection in Italian. *Davis Working Papers in Linguistics 1*, 1-35.

Hay, Jennifer, Kennedy, Christopher & Levin, Beth (1999), Scalar structure underlies telicity in "Degree Achievements", in T. Matthews & D. Strolovitch (eds), (1999), *Proceedings of the 9th Semantics and Linguistic Theory (SALT 9)*, Ithaca, NY, CLC Publications: Cornell University, 127–144.

Levin, Beth (1993), *English Verb Classes and Alternations: A Preliminary Investigation*, University of Chicago Press: Chicago, IL.

Rappaport Hovav, Malka (2008), Lexicalized meaning and the internal structure of events, in S. Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*., Amsterdam/Philadelphia: John Benjamins, 13–42.

Van Valin Robert D. Jr. (1990), Semantic parameters of split intransitivity. *Language* 62, 221-260.

Van Valin, Robert D. Jr. (2005), *Exploring the Syntax-Semantics Interface*, Cambridge: Cambridge University Press.

Van Valin, Robert D. Jr. (2018), Some Issues Regarding (Active) Accomplishments, in R. Kailuweit, L. Künkell, E. Staudinger (eds), *Applying and Expanding Role and Reference Grammar*, FRIAS, Freiburg Institute for Advanced Studies, Albert Ludwigs Universität Freiburg, 71-93.

Creating RRG treebanks through semi-automatic conversion of annotated corpora

Tatiana Bladier, Kilian Evang, Laura Kallmeyer, Robin Möllemann & Rainer Osswald
Heinrich Heine University Düsseldorf

Introduction. Wide empirical coverage is an important issue for any grammatical theory. A strongly data-driven approach toward achieving this goal is the analysis and annotation of a sufficiently large collection of sentences which, ideally, captures all grammatical phenomena of a given language. An annotated corpus of this kind could then be used for testing the validity and the degree of coverage of hand-written grammar fragments as well as for the data-driven extension of such fragments. In this paper, we present ongoing work on developing treebanks for RRG, that is, corpora annotated with RRG compliant syntactic structures. Since large-scale syntactic annotation is a highly time-consuming task, our approach builds on existing annotations, which are transformed automatically into RRG structures. Our automatic conversion is developed on a small set of manually annotated sentences and combined with an additional manual correction cycle. We present two conversion algorithms: the first takes as input constituent structure annotations as used in the Penn Treebank (PTB) (Marcus, Marcinkiewicz, & Santorini, 1993), the second starts with dependency annotations in accordance with the Universal Dependency framework (UD) (Nivre et al., 2016).

Conversion procedure and annotation format. Our tree conversion process is iterative and error-driven, alternating between improving the conversion algorithm and comparing its output to manually validated RRG trees. We apply the conversion algorithm to bootstrap samples of new RRG trees, which are then checked and corrected by the annotators using the click/drag/drop-based web-interface we developed for the RRGbank (https://rrgbank.phil.hhu.de).

The usual notation of the RRG structures differs from the tree notation format typically used in treebanks, in that the RRG structures use an operator projection and represent periphery nodes and clause-linkage markers disconnected from the main constituency structure [cf. van Valin, 2005]. To avoid the discrepancy between the annotation formats, we adopted a notational variant for the RRGbank in which every RRG structure is represented as a single connected tree. We merge the operator projection (which is usually represented in the lower part of the RRG structures) with the constituency projection (see, for example, the position of the label for the tense operator *were* in Fig. 1b). We also attach peripheries (PERI) and clause-linkage markers (CLM) as daughters to the corresponding parent nodes preserving information contained in the original RRG structure. An example of our notational variant is shown in Fig. 1b, in which three boxes mark nodes represented differently in our notation.

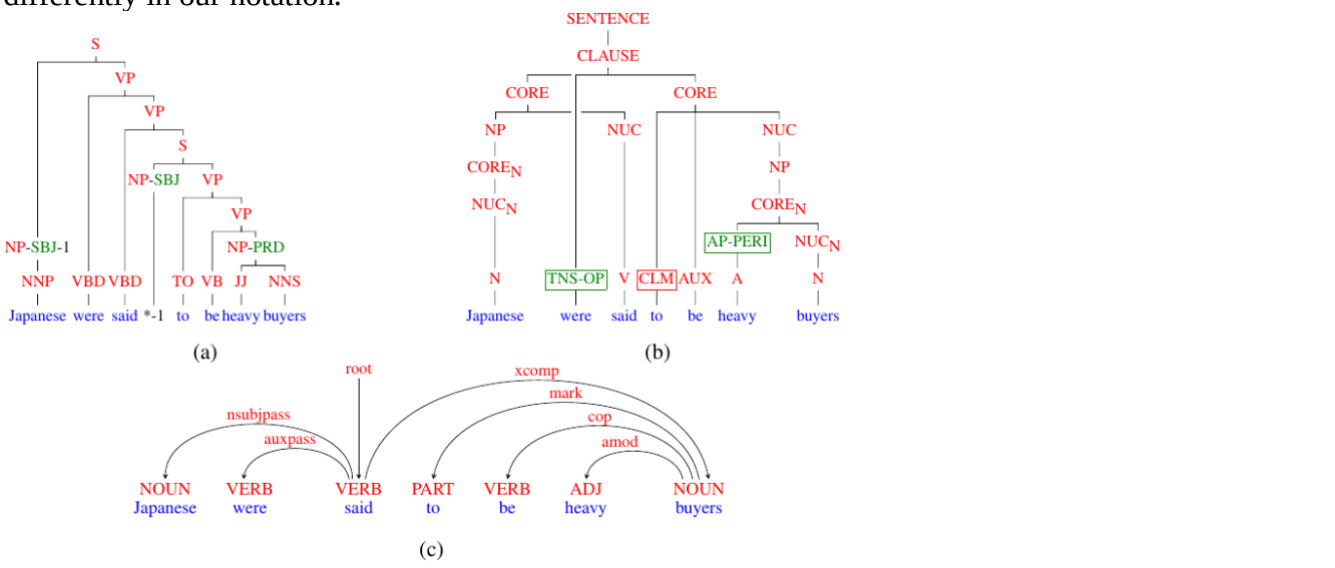


Figure 1: A tree from PTB (a) and PTB-UD (c) treebanks converted to an RRG structure (b).

Transformation of PTB to RRG structures. We chose the Wall Street Journal (WSJ) Sections of the Penn Treebank for conversion, which contains about 50.000 syntactically annotated sentences from WSJ articles. We randomly chose 500 sentences which contain different types of constructions and are no longer than 25 tokens. 100 sentences in this set were manually validated by at least two annotators with RRG expertise and the remaining 400 sentences have been manually corrected by one annotator. Our next goal is to increase the number of corrected sentences to 500 and add another 3000 shorter sentences (≤10 tokens). For transformation, we created a set of conversion rules, each applicable to constituents of a specific type. An example of a PTB tree transformed to an RRG structure is shown in Fig. 1b. We evaluated the performance of our conversion algorithm in terms of *completeness* and *correctness*. *Completeness* of the transformation was measured on the percentage of nodes in converted trees which have a label in the RRG label set. Since PTB and RRG share some labels (for example, PP, NP), the measured completeness amounted to 25.0% before conversion and 97.1% after conversion. We measured *correctness* by comparing converted trees with the manually annotated trees (i.e. our "gold trees"). The overall EVALB F1-score for the first 205 gold trees is 93.02.

Transformation of UDs to RRG structures. Universal Dependencies (Nivre et al., 2016) is a set of annotation guidelines for dependency trees. Like RRG, it emphasizes cross-linguistic applicability of all its structures (here: head-dependent arcs) and categories (here: functional labels of these arcs). The universal guidelines and the wide variety of existing treebanks for more than 70 languages make UD a promising starting point for creating a multilingual RRG resource where many conversion rules can be uniformly applied to corpora of different languages. An example UD tree is shown in Fig. 1c. We are currently in the process of developing an algorithm to convert UD trees to RRG structures by mapping each possible *local tree* (a node with its dependents and arc labels) to an RRG fragment and recursively converting the tree, starting from the root. The overall EVALB F1-score for the first 205 gold trees is currently 86.41.

Problematic and interesting cases during conversion. The main challenge for automatic conversion lies in phenomena for which RRG proposes different analyses, but which are not (explicitly) distinguished by PTB or UD. For example, PTB distinguishes prepositional arguments from prepositional adjuncts only in some cases that are marked with function labels such as -CLR. UD does not distinguish them at all. The control verb constructions are uniformly marked with traces in PTB and with the *xcomp* label in UD, while RRG requires two different analyses for these constructions depending on the nexus choice of the verb (i.e. co-subordination or coordination, see an example from PTB in Figure 2). In such cases, automatic conversion needs heuristic rules—e.g., based on lexical properties—to produce the correct RRG annotation. Comparing PTB and UD as input, we find that while UD trees are uniform across languages and less complex (e.g., they have no VP nodes), they are also slightly less informative, e.g., concerning adjunct vs. argument PPs and the nesting of coordinated NPs. Our converter still has to look at the original PTB tree to resolve these cases, illustrating that conversion from UD still needs some language-specific and treebank-specific rules. It remains to be seen which approach is ultimately superior. The conversion process also revealed several open questions in the RRG theory, which should be studied further, among which are for example RRG analysis of quantifier phrases such as "*more than 60%*". Finally, annotation errors exist in PTB, which need to be corrected manually.

Discussion and future work. In our future work, we plan to test and improve our conversion algorithm on other languages included in the UD corpus, starting with Russian and Tagalog. The ultimate goal is to be able to convert all corpora in the UD corpus to RRG corpora. We plan to make our converted corpora publicly available or available via the Linguistic Data Consortium depending on the original treebanks we use for the conversion.

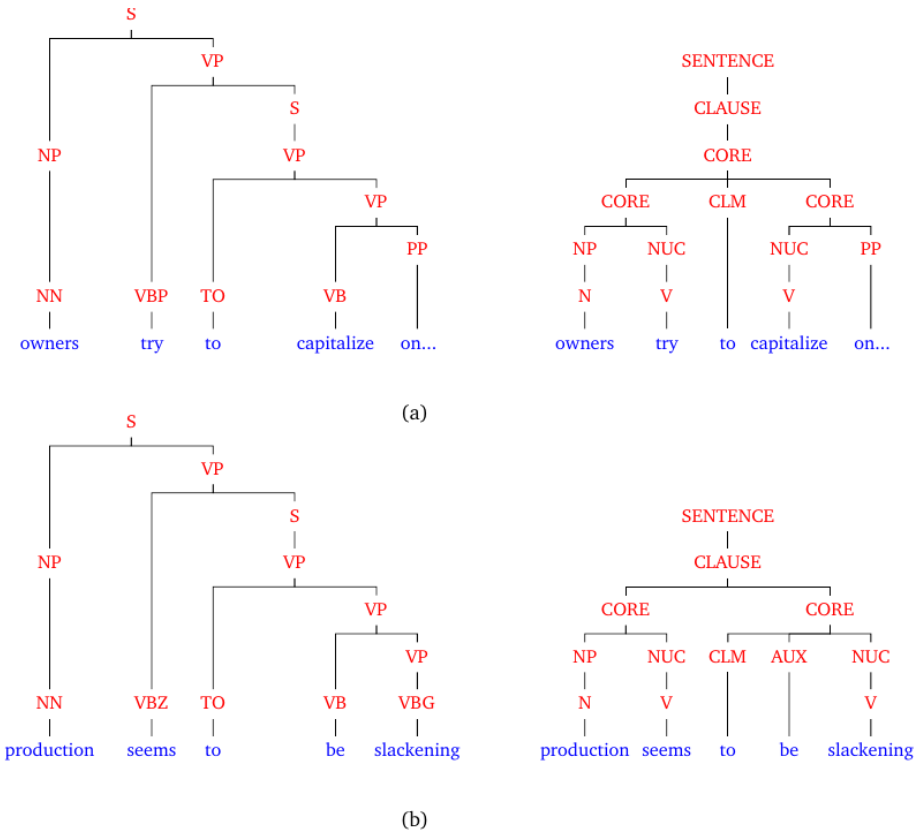


Figure 2: Core cosubordination (a) and core coordination (b) analysis in RRG for one construction in PTB.

References

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In K. Bontcheva & Z. Jingbo (Eds.), *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 55–60). Stroudsburg, PA, USA: Association for Computational Linguistics.

Marcus, M. P., Marcinkiewicz, M. A., & Santorini, B. (1993). Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, *19*(2), 313–330.

Nivre, J., Marneffe, M.-C. de, Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., et al. (2016). Universal Dependencies v1: A Multilingual Treebank Collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (pp. 1659–1666).

van Valin, R. D., Jr. (2005). *Exploring the syntax-semantics interface.* Cambridge: Cambridge University Press.

Agency and lexical decomposition of Biblical Hebrew verbs

Christian Canu Højgaard
Fjellhaug International University College Copenhagen; & Vrije Universiteit, Amsterdam

Role and Reference Grammar (RRG) offers a syntax-semantics interface for determining semantic roles of linguistic references (Van Valin 2005). Following a strict procedure, the linguist begins by analysing the internal temporal aspect of the verb (the *Aktionsart*), then constructs so-called logical structures according to the *Aktionsart*, and, finally, determines the degree of agency by the position of the argument in the logical structure.

The syntax-semantics interface has proven effective for numerous languages. One basic assumption of the interface is the researcher's intuitive knowledge of the language he/she investigates. The verbal *Aktionsart* is determined by applying a set of interrogative questions (cf. Dowty 1979) by which the researcher can classify the verbs by excluding awkward constructions. In other words, the researcher has the freedom to match any verb and any adverb to judge the adequacy of any particular construction.

But what if the researcher had no direct access to the language under investigation? And what if there were no living language-users to consult? Biblical Hebrew is one such language. The entire corpus of the canonical Hebrew Bible comprises no more than 400,000 words and we cannot expect all possible verb-adverb matches to actually exist in the ancient corpus. For these reasons, the RRG syntax-semantics interface is hard to apply to Biblical Hebrew using a traditional procedure.

In a new project, the syntax-semantics interface is reconsidered. If the aim is to predict the degree of agency of literary participants in the Biblical corpus, which parameters are important for this task? In the paper it is argued that at least two inherent verbal features are important: dynamicity (stative vs. active) and causativity, as also evidenced by the morphology of Biblical Hebrew. The finer distinctions of causativity was not developed in the original interface (Van Valin 2005, 42 n. 5) but has later received much more attention (e.g. Nolan, Rawoens, and Diedrichsen 2015; Copley and Martin 2014). The task is two-fold: First, to explore quantitative methods as to their efficiency of predicting dynamism and causativity on the basis of syntactic patterns. And second, to create logical structures of Hebrew verbs and determine the degree of agency of the semantic arguments.

To carry out the research, an open corpus of the Hebrew Bible is used, namely the ETCBC database developed at Vrije Universiteit, Amsterdam, in the Netherlands (Roorda et al. 2018). Open source technologies, such as Jupyter Notebooks and Python 3, are used to extract syntactic constructions from the ETCBC database and to carry out statistical analyses.

References

Copley, Bridget, and Fabienne Martin, eds. 2014. *Causation in Grammatical Structures*. First edition. Oxford Studies in Theoretical Linguistics 52. Oxford, United Kingdom: Oxford University Press.

Dowty, David R. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Synthese Language Library 7. Dordrecht ; Boston: D. Reidel Pub. Co.

Nolan, Brian, Gudrun Rawoens, and Elke Diedrichsen, eds. 2015. *Causation, Permission, and Transfer: Argument Realisation in GET, TAKE, PUT, GIVE and LET Verbs*. Studies in Language Companion Series (SLCS), volume 167. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Roorda, Dirk, Cody Kingham, Christiaan Erwich, and SeHoon Park. 2018. "Etcbc/Bhsa: Default Text Format For Lexemes." Zenodo. https://doi.org/10.5281/zenodo.1446393.

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

Towards a Role and Reference Grammar corpus for English

Christian Chiarcos, Christian Fäth & Monika Rind-Pawlowski
Goethe University Frankfurt

We describe an effort to create an openly available RRG treebank. At the time of writing, no publicly available RRG corpus seem to be in existence, although several initiatives to create rule-based and dictionary-based parsers have been around and are still being continued. Yet, without annotated data, no basis for their systematic evaluation is available, and neither, it would be possible to explore the usability and suitability, advantages or challenged that a semantically oriented formalism to syntax presents for purposes of state-of-the-art natural language processing or natural language understanding techniques which are largely based on machine learning. However, no machine learning without training data.

At the same time, the difficulties in creating RRG corpora have been identified, e.g., for a Quechua-Spanish corpus, whose creators, Rios and Göhring (2012), found that `the annotation process with RRG is too complex and error-prone' and thus shifted to simpler dependency-based syntax formalisms. As an alternative, we suggest deriving RRG annotations from existing, manually created annotations. Because of the specifics of RRG, and especially the great importance of semantics and verbal frames for its structure, no single resource is currently in existence from which am RRG corpus could be just derived. However, we argue that it is possible to derive valid RRG annotations from the intersection of existing annotation efforts in syntax on the one hand, and on verbal frames on the other hand.

We thus render the creation of an RRG corpus in terms of *annotation transformation* rather than *annotation*, and build on two cross-linguistically applied frameworks for syntax and semantics, respectively, the Universal Dependencies (UD) initiative (http://universaldependencies.org/) which currently provides syntactic annotations for more than 90 languages, and PropBank (PB, http://propbank.github.io/), a framework for the annotation of semantic frames of verbal (and nominal) predicates currently applied to 11 languages. Both resources are integrated at a deep conceptual level, where we

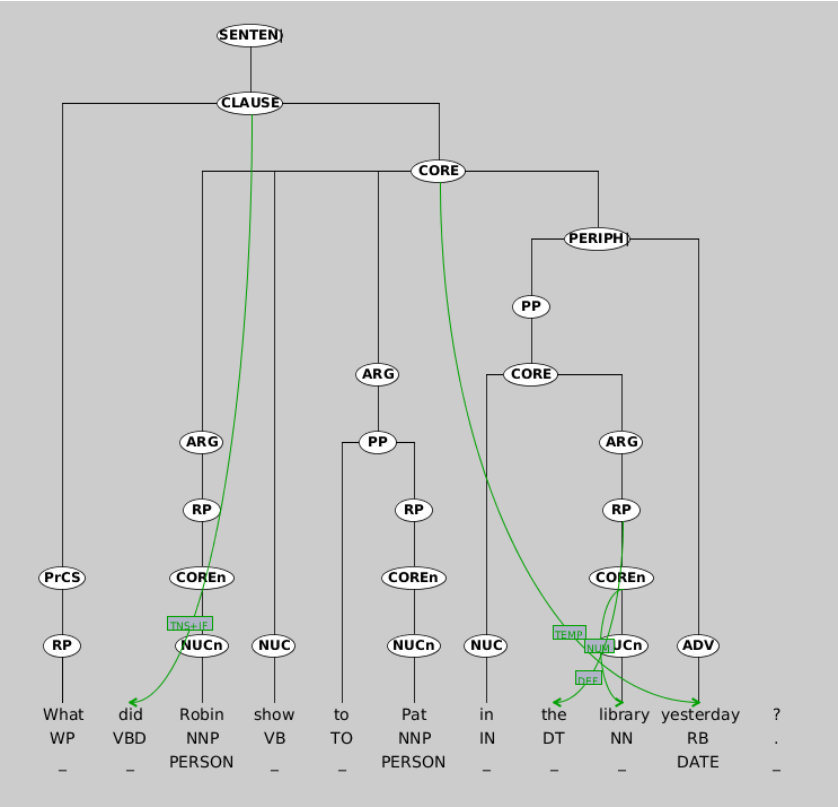
(1) derive CORE, arguments and periphery from semantic frames (PB),
(2) extrapolate the operator projection from morphosyntactic annotations (UD), applied to CORE and CLAUSE,
(3) derive nexus and juncture from shared semantic arguments (PB), operators (UD,PB), and clause linkage markers (UD),
(4) extrapolate remaining syntactic structures from UD.

Note that our approach is not a transformation, but a full decomposition and recomposition of various pieces of linguistic annotation according to RRG assumptions about their interaction. The resulting representation is thus richer than any annotation adopted as source. Also note that the underlying technology, CoNLL-RDF (Chiarcos and Fäth 2017), allows to consult external resources during the transformation, a functionality we use for disambiguating clause linkage markers and prepositions with verbal NUC, which are not distinguished in the underlying annotations.

A sample parse produced using automated annotations for Van Valin (2005, p.7, Fig.1.3) is illustrated below. Note that this visualization is produced with off-the-shelf tools for corpus querying developed by Lezius (2002), so it does not reflect the visual characteristics of Role and Reference Grammar but presents operator projection (green) and constituent projection (black/grey) in a compact, consolidated fashion. As for representing RRG analyses, we adopt a hybrid representation: For example, we keep RRG 1997 ARG labels as these can be reliably predicted from PB, but not automatically disambiguated without a designated RRG frame dictionary (which is not publicly available).

Over our GitHub repository (https://github.com/acoli-repo/rrg), this data is available for download and for use with the corpus tool TIGER Search (Lezius 2002). We provide two data sets under, a gold

corpus comprised of all English examples from Van Valin and LaPolla (1997) and Van Valin (2008). This corpus was automatically annotated using the Stanford parser v. 1.6 by Manning et al. (2014), and the PropBank-compliant SRL system MATE by Björkelund et al. (2010), its annotations manually refined and both annotations then automatically transformed into an RRG representation.
The second dataset is a full transformation of the English Web Treebank, a corpus of 250,000 words with manual annotations for Universal Dependencies and PropBank semantic roles. This corpus was automatically converted and manually evaluated on the answers/dev section of the English Web Treebank. At the moment, we are awaiting copyright clearance for the original text, and, for the moment, provide the transformation scripts, instead, so that interested users can re-build the annotations locally.



The underlying technology is described with greater level of detail in a recent paper by Chiarcos and Fäth (2019). The interested reader may also want to compare a related effort recently described by Bladier et al. (2018), a second RRG corpus of English that was published shortly after our RRG Treebank. A key difference is that their data is created by converting from an existing syntax annotation, so that ambiguities that arise from RRG-specific information missing in the source annotation could not be resolved adequately. In comparison, we integrate information from two sources, semantic and syntactical, to arrive at a representation that is richer than either source annotation.

Our main interest in attending the International Conference on Role and Reference Grammar 2019 is to present our efforts and to discuss our preliminary achievements, but also our limitations. In particular, we would like to explore possible directions for the validation of the generated treebank and discuss this with the wider RRG community. At the moment, we evaluate our efforts by matching the generated trees against known patterns originally devised for RRG *parsing*. However, it is evident that this form of evaluation is not only weak, but also limited to known patterns - whereas many of the constellations we produce are plausible (and in parts, they correspond to text book examples), they do not seem to conform to any of the previously published patterns. Furthermore, we see our efforts as

complementary to rule-based and dictionary-based approaches for RRG parsing, with high potential for mutual synergies.

Acknowledgements
The research described in this paper was conducted in the context of the Early Carreer Research Group "Linked Open Dictionaries (LiODi)", funded by the German Ministry for Education and Research (BMBF).

References

Björkelund, A., B. Bohnet, L. Hafdell, and P. Nugues (2010). A high-performance syntactic and semantic dependency parser. In *Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations*, pp. 33-36. Association for Computational Linguistics.

Bladier, T, A. van Cranenburgh, K. Evang, L. Kallmeyer, R. Möllemann, R. Osswald (2018), RRGbank: A Role and Reference Grammar Corpus of Syntactic Structures Extracted from the Penn Treebank. in *Proceedings of the 17th International Workshop on Treebanks and Linguistic Theories (TLT 2018), December 13–14, 2018, Oslo University, Norway*, pp. 5-16.

Chiarcos, C. and C. Fäth (2017). CoNLL-RDF: Linked corpora done in an NLP-friendly way. In *Proceedings of the First Conference on Language, Data, and Knowledge - First International Conference (LDK 2017), Galway, Ireland, June 19-20, 2017*, pp. 74-88.

Chiarcos, C. and C. Fäth (2018). Graph-Based Annotation Engineering: Towards a Gold Corpus for Role and Reference Grammar. In *Proceedings of the 2nd Conference on Language, Data and Knowledge (LDK 2019), Leipzig, Germany, May*

Chiarcos, C. and N. Schenk (2018). The ACoLi CoNLL libraries: Beyond tab-separated values. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Lezius, W. (2002). TIGERSearch - ein Suchwerkzeug für Baumbanken. In *Proceedings of the 6. Konferenz zur Verarbeitung naturlicher Sprache (6th Conference on Natural Language Processing, KONVENS 2002), Saarbrücken, Germany*.

Manning, C., M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.

Rios, A. and A. Göhring (2012). A tree is a Baum is an árbol is a sach'a: Creating a trilingual treebank. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012*, pp. 1874-1879.

Van Valin, Jr., R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge University Press.

Van Valin Jr, R. D. (2008). *Investigations of the Syntax Semantics Pragmatics Interface*, Volume 105. John Benjamins Publishing.

Van Valin, Jr., R. D. and R. J. LaPolla (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

Clausal cosubordination in Spanish

Carmen Conti
Universidad de Jaén

This paper examines a set of structures containing finite clauses in Spanish that have been traditionally analyzed as coordinate (1-2) or juxtaposed (3). In contrast to conventional analyses, we will argue that the clauses underlined in examples (1)-(3) are cosubordinate (in accordance with Conti's 2018 work on *o… o* 'either… or'). To achieve this end, we will focus in particular on the patterns of grammatical dependency shown by these clauses:

(1) *O no venía o llegaba tarde*.
'Either (s)he wasn't coming or (s)he was arriving late'.
(2) *Ni veía ni quería ver si contenía algo en su interior*.
'(S)he neither saw nor wanted to see whether there was something inside it'.
(3) *Íbamos al cine casi a diario, leíamos, paseábamos.*
'We went to the movies almost every day, we read, we walked'.

Unlike coordinate clauses, cosubordinate clauses are not integrated (they are not part of the constituent they are joined to), but are grammatically dependent (Van Valin 1984: 546; Foley & Van Valin 1984: 242; Van Valin & LaPolla 1997: 454; Van Valin 2005: 188; Hengeveld & Mackenzie 2008: 157). More specifically, as Van Valin (2005: 187) points out, the property that distinguishes cosubordination from coordination is operator dependency. In contrast to coordinate clauses, cosubordinate clauses depend on the same matrix of operators, namely tense, status –epistemic modality, external negation–illocutionary force –the type of speech act– and evidentiality (Van Valin 2005: 9).

As this paper will show, the clauses underlined in examples (1)-(3) above, which are fully inflected, share several restrictions with respect to illocutionary force that point to their grammatical dependency. In contrast to coordinate clauses (4b, 5b, 6b), they reject illocutionary force asymmetries (4a), cannot express a command via the imperative mood (5a) and cannot be asserted independently (6a):

(4) a. **Luisa o se casó con Pedro ¿o se casó con Luis?*
'Luisa either married Pedro… or did she marry Luis?'
b. *Luisa se casó con Pedro ¿o se casó con Luis?*
'Luisa married Pedro, or did she marry Luis?'
(5) a. **O estudia o trabaja*.
'Either study or work'.
b. *Estudia o trabaja*.
'Study or work'.
(6) a. *Los prisioneros o enfermaron o murieron en el mar, ¿verdad?*
'The prisoners either got sick or died at sea, didn't they?'
—*Sí* 'yes' (the listener agrees with this assertion).
—*No* 'no' (the listener disagrees with this assertion).
b. *Los prisioneros enfermaron o murieron en el mar, ¿verdad?*
'The prisoners got sick or died at sea, didn't they?'
—*Sí* 'yes' (Yes, they did).
—*No* 'no' (The listener disagrees with this assertion or the listener disagrees just with one of the alternatives).

This paper will also show that the behavior of other clausal operators, such as tense and status, varies among these structures. The juxtaposed clauses in (3) seem to depend on the same tense and status operators (they cannot be negated independently and the tense must be the same in all of the clauses in the sentence), whereas clauses with *o… o* can show different tenses and can be negated independently. In structures with *ni… ni*, both clauses must be negated (they seem to depend on the same status operator), but they can show differences with respect to tense.

In sum, we will argue that the structures in (1)-(3) show patterns of operator dependency similar to those observed for cosubordinate clauses, in spite of the fact that the clauses being analyzed are finite and seemingly independent.

To this end, this paper will be structured as follows. First, we will summarize the existing literature on *o… o*, *ni… ni* and juxtaposed clauses. Next, we will present the sample used in our study, which comprises more than 200 examples from texts written in European Spanish (*Corpus del Español del Siglo XXI*, RAE 2013-present). Finally, we will analyze the features of grammatical dependency observed in the structures studied, paying particular attention to clause operators.

References

Conti, Carmen (2018), Disjunctive clauses with *o… o* 'either… or' in Spanish and clausal cosubordination, *Folia Linguistica* 52(1), 75-105.

Foley, William A. and Van Valin, Robert D., Jr. (1984), *Functional syntax and universal grammar*, Cambridge: Cambridge University Press.

Hengeveld, Kees and Mackenzie, J. Lachlan (2008), *Functional discourse grammar. A typologically-based theory of language structure*, Oxford: Oxford University Press.

RAE (2013-present), *Corpus del Español del Siglo XXI (CORPES XXI)*, available at www.rae.es.

Van Valin, Robert D., Jr. (1984), A typology of syntactic relations in clause linkage, *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 10, 542-558.

Van Valin, Robert D., Jr. (2005), *Exploring the syntax-semantics interface*, Cambridge: Cambridge University Press.

Van Valin, Robert D., Jr and LaPolla, Randy J. (1997), *Syntax. Structure, meaning and function*, Cambridge: Cambridge University Press.

*This research has been supported by the project *Cosubordinación en español* (FFI2017-85429-P), Ministerio de Ciencia, Innovación y Universidades de España.

The analysis of Japanese benefactive constructions in RRG

Kiyono Fujinaga
University at Buffalo

The paper analyzes the Japanese benefactive construction using RRG framework. First, the paper claims that the Japanese benefactive constructions involve three types of linkage: (1) cosubordination, (2) nuclear subordination, and (3) ad-nuclear subordination. An example of each type is given below. In nuclear subordination, the giving and receiving verbs are predicating nuclei, contributing to the clausal meaning with a recipient argument, which is absent in the semantic representation of the matrix verb. In the ad-nuclear subordination, a non-predicating nucleus is used as a modifier simply to indicate directional orientation of participants and the physical meaning of giving and receiving (transferring objects) are bleached.

(1) Josh cooked (me or my in-group person) a dish. (recipient)
Joshuga (watashi ni) ryoori o tsukut-te kure-ta.
Josh NOM to me dish ACC cook-TE give-PAST
[[**do'** (Josh)] CAUSE [BECOME **cooked'** (dish)]] PURP[BECOME **have'** (I, dish)]

(2) Josh showed (me or my in-group person) his homework. (metaphorical transfer)
Joshuga (watashi ni) shukudai o mise-te kure-ta.
Josh NOM to me homework ACC show-TE give-PAST
[**do**' (Josh, I) CAUSE [**see**' (I, homework)]

(3) Josh went for shopping (for me or my in-group person). (deputive)
*Joshuga (*watashi ni) kaimono ni it-te kure-ta.*
Josh NOM to me shopping forgo-TE give-PAST
[**do'** (Josh, [**go**' (Josh)]) & INGR **be-at**' (shopping, Josh) PURP[NOT [**do**' (I, [**go**' (I)]) & INGR **be-at**' (shopping, I)]]]

Secondly, I will discuss the understood beneficiary in benefactive construction with the verb *kureru* 'give to me'. Unlike the other giving and receiving verbs-- *ageru* 'give' and *morau* 'receive'-- the beneficiary of benefactive *kureru* `give (to me or my in-group person)' is by default understood as the speaker or someone the speaker feels closely associated with (Kuno, 1987; Wetzel, 1985) and the beneficial nominal is often not overtly expressed as in (1) and (2).

Because this construction always assumes the speaker or an in-group member as the recipient, it has a specified argument in the logical structure-- 'I'-- and this argument is directly linked from discourse representation when not overtly expressed in the sentence.

Lastly, in the analysis of 268 occurrences of giving and receiving verbs occurring in natural conversations, I found that the speakers produced the recipient type (1) most frequently. I claim these variations as an evidence that Japanese benefactive constructions are typically used to create beneficiary that are absent in the verb's meaning.

References

Hasegawa, Yoko (1996), *A Study of Japanese Clause Linkage*. Stanford: CSLI Publication.

Kuno, Susumu (1987), *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago: The University of Chicago Press.

Traugott, Elizabeth. C., & Trousdale, Graeme (2014). Constructionalization and constructional changes. In *Oxford Studies in Diachronic and Historical Linguistics*. Oxford: Oxford University Press.

Wetzel, Patricia. J. (1985). In-Group/Out-Group Deixis: Situation Variation in the Verbs of Giving and Receiving in Japanese. *Language and Social Situation*, 141–157.

Integrating head-marking and dependent-marking properties:
An RRG analysis of Halkomelem ditransitive constructions

Valeria Generalova
Heinrich Heine University of Dusseldorf

Halkomelem is a Salish language which is subject to various typological studies. It is claimed to be head-marking, however, it's index of markedness according to Nichols (1986) is -2 (while prototypical head-marking languages can score up to -7). In other words, although Halkomelem is predominantly head-marking, it demonstrates dependent-marking properties in some constructions.

The present paper focuses on ditransitive constructions in Halkomelem (as defined in Gerdts, 2010). These constructions comprise three semantic arguments, which can be realized as RPs or verb affixes. In order to license the presence of three arguments, a verb must comprise a transitivizing and a dative (or beneficiary) suffix. An example of a ditransitive construction is shown in (1) (taken from Gerdts, 2010, 577).

(1) = (66a)

ni?	ʔam–əs–θaḿš–əs	łə	słeni?	ʔə	kʷθə	pukʷ
AUX	give-DAT-TR.1S.OBJ-3ERG	DT	woman	OBL	DT	book

'The woman gave me the book.'

Example (1) demonstrates a combination of dierent argument-marking patterns present in Halkomelem. The Theme argument 'the book' is realized as an independent RP without being marked by a verb affix (although the transitivizer on the verb is present and opens a valence for this argument). In contrast, the Recipient argument 'me' is realized as a verb sux only. Meanwhile, third person Recipients occur as independent RPs. Interestingly, a third-person Agent is doubled: it is realized as a verb affix (glossed *3ERG*) and as an RP 'the woman'.

Dierent combinations of persons for the three arguments lead to a large variety of syntactic templates. It appears that they cannot be analyzed consistently neither in the traditional RRG way, nor with use of specic structures for head-marking languages (as suggested in Van Valin, 2013).

The present paper aims to describe Halkomelem ditransitive constructions using the RRG framework and suggest a novel analysis, combining various RRG decisions.

Our analysis is based on a formalized version of RRG by Osswald and Kallmeyer (2018) and makes use of various features in order to perform the syntactic analysis and link the syntactic structure to the logical one.

The novel idea is to use a "request-response" feature scheme. Axes on the verb licensing the presence of arguments (-TR and -DAT) are treated as features that request the lling of argument slots. Argument suxes and RPs are syntactic constituents that bear the response values of respective features. Argument suxes are part of the CORE*W* structure, which, respectively, is a descendant of the CORE. Feature values from CORE*W* are propagated to the CORE. All arguments must be lled within the CORE. If there are RPs that specify arguments that have already been lled, they appear in extra-core slots [ECS] (see Van Valin, 2013).

As a result, sentence (1) will be parsed as shown in Fig. 1. The Recipient is lled within the CORE*W*, the Theme is lled within the CORE, but outside of the CORE*W*. The Agent is lled with an ax within the COREW and specied with an RP, which is moved to the ECS.

In other words, the necessary information must be lled at the CORE level, while optional argument specication is done within the CLAUSE. It is in line with the principles of the formalization by Osswald and Kallmeyer (2018). The COREW and the CORE are specied in the metagrammar, and ECS are added to the trees through the operation of sister adjunction.

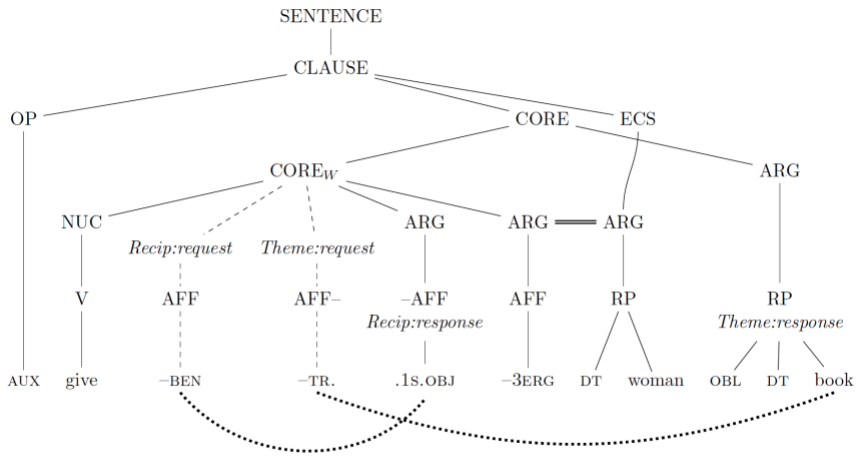


Figure 1: Suggested analysis of Example 1.

Another phenomenon is the lack of overt marking of certain arguments as in (2) (Gerdts, 2010, 576). Within the COREW there are two features "requesting" for arguments. But only one of those arguments is filled: the Recipient is expressed with a first plural pronoun-like verb affix.

(2) = (60)

ni? xʷayəm–əs–t–a̓lxʷ–əs
AUX sell-DAT-TR-1PL.OBJ-3ERG
'He sold it to us.'

The Theme is not lled neither within the COREW with an affix, nor within the CORE with an RP. For these cases, we postulate the presence of a null element "responding" the request of the feature. Although RRG does not encourage null syntactic elements, they are undoubtedly present in some morphological paradigms. Postulating a null element in (2) would help to keep the metagrammar compact and unied without inserting additional trees for sentences like (2).

The conference talk is going to comprise more detailed information about our analysis and more examples of Halkomelem ditransitive constructions. This abstract shows that our analysis is able to deal with head-marking and dependent-marking trends encountered in Halkomelem ditransitive constructions. Although we considered only Halkomelem so far, we anticipate this approach to be adequate for other languages demonstrating both head-marking and dependent-marking properties.

References

Gerdts, D. (2010). Ditransitive constructions in Halkomelem Salish: A direct object/oblique object language. In A. L. Malchukov, M. Haspelmath, and B. Comrie (Eds.), Studies in ditransitive constructions: A comparativehandbook, pp. 563-610. Walter de Gruyter.

Nichols, J. (1986). Head-marking and dependent-marking grammar. Language, 56-119.

Osswald, R. and L. Kallmeyer (2018). Towards a formalization of Role and Reference Grammar. In R. Kailuweit, E. Staudinger, and L. Kunkel (Eds.), Applying and expanding Role and Reference Grammar (NIHIN Studies), pp. 355-378. Freiburg: Albert-Ludwigs-Universität, Universitätsbibliothek.

Van Valin, Jr, R. D. (2013). Head-marking languages and linguistic theory. Language typology and historical contingency: In honor of Johanna Nichols, 91-124.

Double locatives: the case of external possession in Yaqui

Lilián Guerrero & Valeria A. Belloro
Universidad Nacional Autónoma de México & Universidad Autónoma de Querétaro

It is well known that languages have more than one structure for encoding the relation between a possessor and a possessee, and that their distribution usually correlates with different semantic relationships, e.g., kinship terms, body parts, garments, natural entities, domestic animals, and the ownership of other personal objects (Heine 1997; Stolz et al. 2008). The Yaqui language is not the exception (Uto-Aztecan, Mexico). At the phrase level, a possessive relation can be expressed by juxtaposition (1a), possessive phrases (1b) and genitive phrases (1c); in the later, the possessor takes the accusative case marker. These are cases of internal possession (the two participants form a single constituent). At the clause level, Yaqui also has the so-called possessive clauses taking a possessive verb (1d).

(1) a. waka beea
cow skin
'The cow's skin'

b. em koba
2sg.pos head
'your head'

c. Joan-ta kari
Juan-acc house
'John's house'

d. Empo bwe koba-ta jippue
2sg.nom big head-acc have
'You have big head.'

This paper focuses on external possessor constructions involving body parts (Heine 1997; Payne & Barshi 1999; Haspelmath 1999; Lødrup 2009). In the constructions in (2), the possessor occurs within a constituent separate from that which contains the body, there is no possessive verb, the body part is marked by the (general) locative postposition *-po,* and the coding of the possessor varies. The basic (unmarked) structure involves the external possessor as the nominative subject and the body part as a locative argument; the literal meaning of (2a) would be 'Mary burned on [her] hand'. In (2b), there is an applicative morpheme, and the possessor takes accusative case, 'He/she squashed me the fingers'. The intriguing examples are (2c-f). In (2c), the possessor is coded by an accusative pronoun but there is no applicative morpheme, 'the boy scratched me on the face'; in (2d) the possessor takes the directional postpositional marker *-u* (i.e., dative-like marker), something like 'the woman took the scarf on the neck to her'; in (2e-f), the possessor is marked by the locative postposition *-(e)t* 'on, over', 'I put the necklace on the neck on her'. Thus, there is an additional participant that is not a part of the valency of the verb: the possessor counts as a core argument and the verb does not carry any valency-changing morpheme.

(2)a. [Maria-Ø] [mam-po] taja-k
María-nom hand-loc burn-pfv
'Mary burned her hand.' (lit. burn on the hand)

b. Aapo [mam-pusiam][nee] pueta-po pitta-ria-k
3sg.nom hand-finger.pl 1sg.acc door-loc squash-appl-pfv
'He/She squashed my fingers in the door.' (lit. squashed me the fingers)

c. U ili miisi-Ø [pujba-po] [nee] witta-k
det little cat-nom face-loc 1sg.acc scratch-pfv
'The little cat scratched my face.' (lit. scratched me on the face)

d. U jamut-$Ø_i$ tajoorim [kuta-naa-po] [a-u_i] u'ura-k
det woman-nom cloth.acc.pl neck-around-loc 3sg.obl-dir take-pfv
'The woman took off the scarf on her neck.' (lit. took the scarf on the neck to her)

e. Inepo [kutanaa-po] koka-ta [a-et] yecha-k
1sg.nom neck-loc necklace-acc 3sg.obl-loc put-pfv
'I put the necklace on her neck.' (lit. put on her the necklace on the neck)

f. U baiseebolim-Ø [mam-po] [ne-t] yejte-k
det butterfly-nom hand-loc 1sg.obl-locc stand-pfv
'The butterfly stood in my hand.' (lit. stood on the hand on me)

We first analyze whether the Yaqui data belong to the well-known 'possessor raising' or 'dative' external possession found in European and Mesoamerican languages (e.g. *yo le corté el cabello* 'I cut his hair', *le pegué en la pierna* 'I hit him on the leg'; Conti 2011; Palancar & Zavala 2013). We demonstrate that Yaqui data satisfy some of the affectedness conditions proposed by Haslpemath (1999), namely the referential, inalienability and situational hierarchies, and the fact that the possessor and the possessee do not belong to the same phrase. However, the syntactic relation hierarchy is disfavored: although the possessee is expressed as a locative PP, the coding of the possessor as a dative-like argument (2d) is unusual. In fact, Yaqui prefers to demote both, the possesee and the possessor. We also examine how Yaqui external possessors can be accommodated in Role and Reference Grammar (Van Valin 2005). First, the construction in (2c) can be represented by the logical structure [LS] in (3a); however, (2d-f) cannot be captured by the regular LS for a three-place predicate (3b) because in it there are three argument positions only; instead, we propose the LS in (3c). The second problem to solve is the assignment of the Undergoer macrorole. In Yaqui, only accusative arguments can serve as the passive subject, hence the Undergoer (Guerrero & Van Valin 2004). In external possessive constructions, only the accusative possessor in (2b-c) shows the same syntactic privilege (Undergoer). The constructions in (2d-f) do not take Undergoer macrorole (i.e., are M-intransitive). A third problem is how the locative coding of the possessor and possessee can be predicted from the LS of these constructions. In (3c) we propose the component become **be-at'** (x, y) in order to highlight the locative meanings over the possessive meaning in (3b). A final problem is the assignation rules for the two locative postpositions: the general locative *-po* 'on' marking the body part (the first argument of the first locative component), and the specific locative *-(e)t* 'on, above' and the directional *-u* 'to' marking the external possessor (the first argument of the second locative component).

(3)a. [**do'** (cat, Ø)] cause [become **scratched'** (1sg, face)]
b. [**do'** (1sg, Ø)] cause [become **have'** (neck, necklace)] possessor?
c. [**do'** (1sg, Ø)] cause [become **be-at'** (neck, necklace)] & [become **have'** (3sg, necklace)]

References

Conti, Carmen (2011), Possessive dative revisited: another view of external possession in Spanish, *Studia Linguistica* 65 (1): 170-197.

Guerrero, Lilián & Robert Van Valin, (2004), Yaqui and the analysis of primary object languages, *IJAL* 70 (3): 290-319.

Heine, Bernie, (1997), *Possession. Cognitive sources, Forces, and Grammaticalization*. Cambridge: CUP.

Haspelmath, Martin, (1999), External possession in a European areal perspective, in *External Possession*, D. Payne & I. Barshi (eds.). Amsterdam: John Benjamins, 109─135.

LØdrup, Helge, (2009), External and internal possessors with body part nouns: the case of Norwegian, *SKY Journal of Linguistics* 22: 221-250.

Palancar, Enrique & Roberto Zavala, (2013), *Clases léxicas, posesión y estructuras de cláusula compleja en Mesoamérica.* Una introducción. Available at hal-01493982.

Payne, D. & I. Barshi. 1999. *External Possession*. Amsterdam: John Benjamins.

Stolz, Thomas, et al, (2008). *Split Possession: An Areal-linguistic Study of the Alienability Correlation and Related Phenomena on the Languages of Europe.* Amsterdam: JB

Van Valin, Robert, (2005), *Exploring the syntax-semantic interface*. Cambridge: CUP.

Neither dative nor instrumental as default in Urdu: Dative/instrumental marked non-MR arguments as PSA

Saima Hafeez
University at Buffalo

Dative case has been established as being a default case due to its occurrence in oblique arguments which represent a number of different thematic roles across languages (Van Valin, 2018). This paper argues that an instrumental case also performs a distinct kind of idiosyncratic behavior in occurring on human effector argument NPs in representations of accidentally happening events in Urdu. It further postulates that, like dative marked oblique argument NPs, the instrumental marked argument NPs also occur as the Privileged Syntactic Argument (PSA) in Urdu, other than ergative or nominative marked nominals. The paper posits that neither ergative nor dative are default case markers in Urdu, as dative marks the non-macro-role (non-MR) experiencers and instrumental marks non-MR effectors.

The assignment of dative case to the argument NP representing an experiencer is supported by the view that the nominative or ergative case is blocked from occurring on the PSA due to a lack of evidence for agentivity. This explanation is supported from data from different languages in which agentivity evidence is completely blocked. This analysis can be extended to the occurrence of instrumental oblique arguments as PSA in monoclausal constructions of Urdu.

In Urdu, in terms of an effector's potentiality continuum, the agents with high potentiality instantiated by ergative marked NPs are highly ranked, whereas, the instrumental marked nominals are ranked the lowest. Moreover, the human referents of dative-marked nominals outrank the instrumental arguments but are subsumed under the ergative marked arguments. The continuum can be formed as in (1):

(1)Ergative nominal > dative nominal > instrumental nominal
Prototypical agentivity > induced potentiality > reduced potentiality

The paper offers an RRG analysis to explain the behavior of human effectors with reference to varying potentiality levels in the causality domain. The highest level can be aptly termed as 'prototypical agentivity' and is expressed through ergative-marked nominals. Contrary to this, the occurrence of instrumental nominals expresses the lowest potentiality level, which is assigned to human effectors who are allocated a 'reduced potentiality status'. Finally, in obligation constructions with a dative nominal, the potential to carry out an action is neither optimal nor marginal, rather it is somewhere in between these. Hence it is treated in the proposed analysis as having an induced potentiality status.

The occurrence of dative and instrumental markers and their alternations is also discussed with respect to constructions with psych predicates, consumption predicates and transfer predicates. These domains are explored to show that causality is an apt domain to study the variation in the agentivity or potentiality levels of human effectors, which manifest in causative constructions as causers or causees.

After establishing that instrumental causers in Urdu/Hindi act as PSAs, evidence is presented to show that the instrumental acts in a manner similar to applicatives, in that it adds a non-macro-role oblique argument to the core.

References

Van Valin, R. D. Jr. (2018). Dative case and oblique subjects. In J. BarSdal, N. Pat-El & S. M. Carey (Eds.), Non-Canonically Case-Marked Subjects, The Reykjavik- Eyjajjallajokull papers (pp. 117-131). John Benjamins Publishing Company.

A Study of the *Annotative Dual-Sentence Juxtaposition* Construction in Japanese

Yoko Hasegawa
University of California, Berkeley

This study investigates a type of parenthetical expression in Japanese called an **annotative dual-sentence juxtaposition** (*chūshakuteki nibun-renchi*, ADSJ), exemplified in bold in (1), and explores how it can be analyzed and represented in the RRG theory of clause linkage.

(1) *Kono uchi ni chigai-nai ga, **doko kara hait-te ii ka**,*
this house COP different-NEG but where from enter-TE good INT
katte-guchi ga nakat-ta.
kitchen-door NOM not.exist-PST
Lit. 'It must be this house, but, **where one can enter**, there was no kitchen door.'

This sentence poses no comprehension difficulty to native speakers of Japanese, but is likely to be incomprehensible to most non-native speakers. Domenico Lagana (1975) writes that it appeared at first glance easy to translate. However, his analysis of its structure and recognition of all lexical meanings provided him with absolutely no clue as to how to interpret the sentence as a whole. He eventually gave up hopelessly with this interpretation:

(2) There is a house (or possibly houses) in a certain place. Currently, the house is not dissimilar to something else, probably to other houses (or looks long unchanged). Someone asks someone else a question: Where can someone (or who) or something (or what) go into it? A LOGICAL LEAP. In the past, there was no backdoor.

Structurally, this construction resembles an indirect question, e.g. (3):

(3) ***Doko kara hait-te ii ka** wakar-anaka-tta.*
where from enter-TE good INT find-NEG- PST
Lit. 'I couldn't find **where one could enter** (i.e. where to enter).'

Indirect questions occur only with certain types of predicates: e.g. communication (*disclose*, *inform*), conjecture (*estimate*, *guess*), decision (*control*, *decide*), inquisitive (*ask*, *be interested in*), knowledge acquisition (*discover*, *learn*), knowledge retention (*forget*, *know*), relevance (*be important*, *matter*). In RRG, indirect questions in Japanese are analyzed as clausal subordination with the interrogative particle *ka* as a complementizer (Van Valin and LaPolla 1997: 669). As Lagana was baffled, however, the *ka*-marked interrogative constituent in (1) has no apparent syntactic or semantic relation to the host-constituent, *there was no kitchen door*.

ADSJ is by no means a tangential construction; it has existed perpetually since Old Japanese (700-800 CE). Today, the most common function of the sentence-medial *ka* is to mark an indirect question, as in (3). However, ADSJ is not derived from this common function. Rather, the indirect question was developed in the 14th to 16th centuries and became popular only in Early Modern Japanese (i.e. 17th to 19th centuries). Therefore, ADSJ should be regarded as an instantiation of a more essential (central) function of the particle *ka*.

Due to the lack of syntactic mechanism to connect the *ka*-constituent and the main (sentence-final) constituent in (1), the former is deemed parenthetical. *Parentheticals* are expressions that are interpolated in, but seemingly independent structurally of, the host sentence. Although enormously diverse in form and complexity, parentheticals generally remark on various aspects of verbal communication, commonly to convey "additional" information, e.g. the speaker's attitude, certainty, or

endorsement towards the statement made by the utterance. Semantically and pragmatically, the typical function of the *ka*-constituent in ADSJ is to conjecturally *annotate* how the situation denoted by the main-constituent is brought about, as exemplified in (4).

(4) *Hiro wa, **netsu ga aru no ka**, ase o kaite-ita.*
TOP fever NOM exist NMLZ INT sweat ACC push.out-was
Lit. 'Hiro, **if he has a fever**, was sweating.'

RRG provides a rich apparatus for the study of clause linkage by positing four levels of juncture (nuclear, core, clausal, and sentential) and three nexus relations (coordination, subordination, and cosubordination). Because the linked units in cosubordination are dependent upon the matrix unit for one or more of the operators for that level, and because the matrix unit of sentential juncture is *text*, which does not have its own operator, sentential cosubordination is impossible (Van Valin 2005: 201). Combinations of four juncture and three nexus types, then, yield eleven linkage types in universal grammar. As can be seen in (4), the *ka*-junct in ADSJ can have its own subject and predicate that can be distinct from those of the main constituent; therefore, the juncture level is either clausal or sentential, but not nuclear or core.

Nomura (1995), who named this construction *chūshakuteki nibun-renchi*, analyzes it as consisting of two *sentences* being juxtaposed (i.e. coordinated) because the main junct cannot host the *ka*-junct as subordination. However, RRG does not sanction it as sentential coordination because the two constituents cannot have independent illocutionary force (IF) operators. The *ka*-junct is interrogative in form, but, as with the case of indirect questions, it does not carry the illocutionary force of inquiry. That is, there is only one IF operator of the sentence, and it must be declarative. Therefore, the juncture level is clausal, and the nexus type cannot be coordination.

Two subtypes of clausal subordination are recognized in RRG: complements and adverbials. As discussed above, the *ka*-clause in ADSJ is not a complement of the matrix predicate. On the other hand, it can be analyzed as an instance of adverbial subordination. For example, the semantic function of the *ka*-clause in (4) appears to be comparable to a *because*-clause, i.e. "because Hiro (probably) had a fever, he was sweating." However, this type of adverbial modification is hitherto unknown; therefore, such an analysis then needs to posit *ka* as a clausal subordinator (vis-à-vis a complementizer) with independently motivated evidence. The present study advocates that the nexus type involved in ADSJ is clausal cosubordination – an analysis that can account for the difference between (1, ADSJ) and (3, indirect question) without adding extra complexity to the theory.

Although ADSJ has drawn some attention from researchers in Japan (e.g. Hattori 1992, Nomura 1995, Ishii 2003, Takamiya 2003, 2004, Shiba 2016), their works mostly focus on its historical origins and subsequent development, and articles on ADSJ in English are scarce (Some notable exceptions are Serafim and Shinzato 2005 and Kinuhata 2012). The present study will contribute to enrichment of the RRG theory of clause linkage by reporting a new type of clausal cosubordination, which is often illustrated by switch-reference constructions.

References

Hattori, Tadasu. 1992. Gendaigo ni okeru *-ka* no aru shuno yōhō ni tsuite [On a certain usage of *-ka* in Modern Japanese. *Tokushima Daigaku Kokugo Kokubungaku* 5: 57-65.

Ishii, Yasuo. 2003. Interrogative adjuncts in Japanese: A preliminary study. In *Empirical and Theoretical Investigations into Language: A Festschrift for Masaru Kajita*, ed. Shuji Chiba et al., 281-96. Tokyo: Kaitakusha.

Kinuhata, Tomohide. 2012. Historical development from subjective to objective meaning: Evidence from the Japanese question particle *ka*. *Journal of Pragmatics* 44: 798–814.

Lagana, Domenico. 1975. *Nihongo to watashi [The Japanese language and myself]*. Tokyo: Bungei Shunju.

Nomura, Takashi. 1995. *Ka* ni yoru kakari-musubi shiron [A study of *kakari-musubi* with *ka*]. *Kokugo kokubun* 64 (9): 1-27.
Serafim, Leon , and Rumiko Shinzato. 2005. On the Old Japanese kakari (focus) particle *koso*: Its origin and structure. *Gengo Kenkyu* 127: 1-49.
Shiba, Ayako. 2016. Kindai-nihongo no kansetsu-gimon-kōbun to sono shūhen: Jūzoku *ka*-setsu o motsu kōbun no nettowāku [The indirect Interrogative and its relational constructions in Late Modern Japanese: A network between constructions with the embedded *ka*-clause]. *Kokuritsu Kokugo Kenkyujo Ronshu* 10: 193-220.
Takamiya, Yukino. 2003. Gendai nihongo no kansetsu-gimonbun to sono shūhen [The Indirect question construction and its periphery in Modern Japanese]. *Mie Daigaku Nihongogaku Bungaku* 14: 104-16.
Takamiya, Yukino. 2004. *Yara(u)* ni yoru kansetsu-gimonbun no seiritsu: Futeishi gimon o chūshin ni [Origins of the indirect question construction with *yara(u)*: Focusing on wh-questions]. *Mie Daigaku Nihongogaku Bungaku* 15: 124-11.
Van Valin, Robert. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
Van Valin, Robert, and Randy LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Relevance of the scope of operators in determining the basic meaning of a construction: The case of « *took and went away*» (*"agarró y se fue"*) in Spanish

Felipe Hasler, Carlos González, Silvana Guerrero, Matías Jaque & Verónica Orqueda
Universidad de Chile, Universidad Católica, Universidad de Chile, Universidad de Chile & Universidad Católica

As stated in Jaque et al. (2018), in Spanish there is a type of constructions that combines two finite verbs in a single predicate and, therefore, constitute the nucleus of a single clause, as shown in the following example:

(1) Mi hijo agarr-ó y se fu-e
My son AUX.took-3SG.PST and went.away-3SG.PST
Out of the blue, my son went away

This type of constructions bears similarities with serial verbs in semantic-grammatical terms (as they constitute a single predicate), but it differs from them in formal characteristics, such as that they can be interrupted (My son took, and after hesitating a moment, went away/*Mi hijo agarró y, luego de dudarlo un momento, se fue*), suffer the individual denial of the finite verbs that compose them (My son, does not took and leaves without saying me nothing?/*¿Mi hijo, no agarra y se va sin decirme nada?* / Even having made many undercover threats, finally my son took and did not leave/*Aunque había hecho muchas amenazas encubiertas, al final mi hijo agarró y no se fue*) and show a linking element in the conjunction *and* (*y)*. Likewise, it bears similarities with the so-called verbal periphrasis of Spanish grammatical studies in which only the second of the verbs conveys a lexical meaning, while the first expresses a grammatical value. This is the case, for example, in *My son had to leave* (*Mi hijo tuvo que irse)*, in which the verbal periphrasis *had to* (*tener que*) + *infinitive* expresses a deontic modal meaning. It differs from these, however, in that periphrasis is always constructed with the lexical verb in one of its non-finite forms. Considering the similarity between these constructions, serial verbs and verbal periphrasis, we call them *multiverbal constructions*, in the sense proposed by Aikhenvald (2011).

Although multiverbal constructions of the type of (1) have not received much attention in Spanish grammatical studies, those who have studied them do not agree on what is meaning of the construction. It has been proposed, fundamentally, five explanations that justify it: (a) the first verb is an expletive (Valdés, 1553, Correas, 1631, Cuervo, 1867); (b) the construction expresses a completive aspectual meaning, in opposition to the consideration of the event in its phases (Keniston, 1936, Kany, 1951, Coseriu, 1966); (c) the construction has an ingressive aspectual meaning (Fraenkel, 1926, Havers, 1927, García Sánchez, 2007); (d) the construction expresses mirative meaning (Keniston, 1936, Coseriu, 1966); (e) the construction has a topic informative function (Arnaiz and Camacho, 1999) or a focalizing one (Seco, 1961, Aleza Izquierdo and García-Mendall, 1986, Garachana, personal communication).

Our proposal is that the multiverbal constructions of the type of *took and went away* (*agarró y se fue*) express essentially an aspectual value, although not ingressive or completive, but interruptive, and that the value of mirativity and focus indicated in the bibliography is derived from this main value by an inference process. In this framework, the objective of this research is to characterize the scope and the combinatorial of the operators as a test to determine the basic meaning of a construction, for which we will focus on the distinction between the basic interruptive meaning and two of the meanings proposed in related literature: the ingressive, which can be discarded, and the mirative, which is proposed as a meaning inferentially derived from the basic meaning of the construction.

Regarding the ingressive meaning, we propose that it is not the basic meaning because it is possible to combine it with periphrasis that express that meaning. In contrast, the ingressive periphrasis do not combine well between them, as shown in (2), which is expected, given that they try to establish focus

on the same section of the event and, consequently, occupy the same place in the projection of operators.

(2) a. *It started to begin to rain (Se puso a empezar a llover)
b. Arrived and it started to rain (Llegó y se puso a llover)

In fact, it seems that a sentence like *arrived and it started to rain* (*llegó y se puso a llover*) is even preferable to the simplest one *arrived and rained* (*llegó y llovió*), which indicates that the construction in question, although it does not transmit an ingressive meaning, is highly compatible with it.

As for the mirative meaning, since mirativity is one of the broadest operators in the structure of the clause and has effects all over it, then, this construction should not be able to appear under the control of an operator of narrower effects, such as, for example, a deontic modal operator. However, as seen in the example in (3), the desemantised verb *come* (*llegar*) may be under the domain of a deontic modal verb as *can* (*poder*), compatible with an aspectual meaning, since the aspect has one of the narrowest reach in the structure of the clause.

(3) Tú no pued-es lleg-ar y entr-ar
You NEG can-2SG AUX/come-INF and go.in-INF
sin pagar
without paying
You can't go in without paying, unexpectedly

As we have indicated, the mirative meaning arises inferentially in more specific contexts, such as the combination with stative verbs, usually imperfective, as can be seen in (4)

(4) Encontraron un guardapelo cerrado y va y era un
They found a closed locket and go.3SG and was.3SG a
Horrocrux
Horrocrux
They found a closed locket and it came to be a Horrocrux

References

Aikhenvald, A. (2011). Multi-verb constructions: setting the scene. A. Aikhenvald y P. Muysken (Eds.), *Multi-verb constructions. A view from the Americas*. Leiden: Brill.

Aleza Izquierdo, M., y García-Medall, J. (1986), "Funcionalidad de la perífrasis «cojo y me voy» en español", *Cuadernos de Filología II: Studia Lingüística Hispánica*, 3, 5-17.

Arnaiz, A., y Camacho, J. (1999), "A Topic Auxiliary in Spanish", J. Gutiérrez-Rexach y F.M. Gil (eds.): *Advances in Hispanic Linguistics: Papers from the 2nd Hispanic Linguistics Symposium (Vol. 2)*, Somerville, MA, Cascadilla Press, 317-331.

Correas, G. ([1631] 1906). *Vocabulario de refranes y frases proverbiales y otras fórmulas comunes de la lengua castellana: en que van todos los impresos antes y otra gran copia que juntó el maestro Gonzalo Correas*. Madrid: J. Ratés.

Coseriu, E. ([1966] 1977). Tomo y me voy. Un problema de sintaxis comparada europea. *Estudios de lingüística románica*. Madrid: Gredos.

Cuervo, R. J. (1867). *Apuntaciones críticas sobre el lenguaje bogotano con frecuente referencia al de los países de Hispano-América*. Paris: A & R, Rogers y F. Chernoviz, Editores

Fraenkel, E (1926). "Zur Parataxe und Hypotaxe im Grieschischen, Baltoslawischen und Albanischen", en *Indogermanische Forschungen*, 24, 290-315.Havers

García Sánchez, J. (2007). “Tomo y me voy”. Expresión plena y elipsis. *Actas del XV Congreso de la Asociación Internacional de Hispanistas*. México D.F.: Fondo de Cultura EconómicaArnaiz y Camacho
Jaque, M., C. González Vergara, S. Guerrero, F. Hasler y V. Orqueda. 2018. Es llegar y llevar: construcciones multiverbales de verbo finito coordinadas en español. *Boletín de Filología*.
Kany, Ch. ([1945] 1970). *Sintaxis hispanoamericana*. Madrid: Gredos
Seco, M. ([1961] 2002), *Diccionario de dudas y dificultades de la lengua española*. Madrid: Espasa.
Valdés, J. de. ([1535]1990). *Diálogo de la lengua*, (ed. C. Barbolani). Madrid: Cátedra

Clausal cosubordination in Mapudungun: Relevance of referential hierarchies for the RRG Linking Algorithm

Felipe Hasler
Universidad de Chile

Mapudungun is a non-genealogically classified Andean language that is spoken with different degrees of vitality in the center and south of Chile and Argentina. In the plane of the simple clause, it is a head-marking language with an integrated inverse alignment system governed by a saliency hierarchy based on the inherent topicality associated with the status of each participant in the speech act participant ranking and in the discursive topicality related to the opposition third person proximate/obviative (Golluscio 2010). Regarding complex syntax, dependent constructions are characterized by the use of nominalizers, with the personal reference indexed in the possessive adjective that precedes the nominalized verbal form.

This research, framed in the projects FONDECYT 11180078 and FONDECYT 1180071, aims to explore the relevance of referential hierarchies in the Linking Algorithm within the framework of interclausal relations in Mapudungun, following the RRG approaches presented in Van Valin (2005). For thus, we describe the cosubordinate clausal construction used in this language for the expression of relations of cognition, exemplified in (1), since it presents a mismatch between semantics and syntax that highlights the phenomenon to be studied:

(1) kom che kim-a-f-e-**n**-ew
all people know-FUT-IR-INV-IND.1SG.P-OBL
[**ta=ñi** ketro-le-n ta-ñi longko]
ta=POS.1SG obtuse-ESTAT-NR1 ta-1SG.POS head
'All the people would have known me to be without head.' (Salas 2006, [1992]:267)

It is a cosubordinate construction since, first of all, the dependent clause is not a syntactic argument of the main one - in fact, it does not correspond to a semantic argument: the place that corresponds to the non-PSA argument of the construction is assumed by the PSA of the dependent clause. In addition, it depends in terms of operators of the main one, since it can not receive evidential operators (2). In contrast, if it can receive tense operators, such as the suffix of future *-a* (3), which allows us to identify the construction as a clausal juncture.

(2) *kim-e-yu [ta-mi amu-rke-el wiya]
know-INV-IND.1A.2P DET-2SG.POSS ir-EVID-NR2 yesterdat
'I (to you) know they say you went yesterday.'

(3) kim-e-**yu** [**ta-mi** amu-a-el wule]
know-INV-IND.1A.2P DET-2SG.POSS ir-FUT-NR2 tomorrow
'I (to you) know that you leave tomorrow.'

According to the Linking Algorithm, in a clausal juncture each related clause should link semantics and syntax independently (Van Valin 2005: 228). However, as seen from the contrast between (4) and (5), it is interesting to note that the link between semantics and syntax of the independent clause can only be resolved once it has been resolved in the dependent clause.

(4) a. Juan kim-e-**n**-ew [ta=**ñi** kewa-fi-el Pedro]
Juan know-INV-IND.1SG.P-OBL ta=POS.1SG punch-3.P-NR2 Pedro
'Juan knows that I punched Pedro.' (Golluscio and Hasler 2017: 72)

(5) b. Juan kim-e-**n**-ew [ta=**ñi** kewa-e-t-ew Pedro]
Juan know-INV-IND.1SG.P-OBL ta=POS.1SG punch-INV-NR3-OBL Pedro
'Juan knows that Pedro punched me.' (Golluscio and Hasler 2017: 73)

Thus, the PSA of the dependent clause appears in the main clause as Undergoer: its syntactic expression is evaluated again according to its relation of salience with the Actor. In example (4), the PSA of the dependent clause, which assumes the Actor macro role, is encoded in the main clause as PSA; in contrast, in the example (5), the PSA of the dependent clause, which assumes the Undergoer macro role in this clause, is also coded as PSA in the main clause. This occurs because the Mapudungun has, as we mentioned before, an opposition between direct and inverse voice that is resolved according to the place of the arguments in the hierarchy of salience, detailed in (6) (Golluscio, 2010: 714): if the Actor is ranked higher in the hierarchy than the Undergoer, then the construction will have a direct voice; if the opposite occurs, the construction will exhibit inverse voice.

(6) 1SG / PL> 2SG / PL> 3SG / PL Proximate> 3SG / PL Obviative

According to this, both in (4) and (5), the PSA of the main clause is the 1st person, an argument that does not belong to the argumental structure of that clause. Then, unlike what was predicted in the model, our proposal is that in these constructions the argument of the dependent clause that is encoded in the main one is determined by its PSA function in the dependent clause, which requires that macro roles have been assigned to the arguments and these have been evaluated according to their place in the empathy hierarchy. Then, once the semantics and the syntax of the dependent clause have been completely linked, the argument of this clause encoded in the main clause assumes the Undergoer macro-role and is evaluated again in its relation with the Actor of that clause within the framework of salience hierarchy, receiving a new syntactic function. In this way, we propose that the link in these constructions occurs sequentially and not independently, even when it is a clausal juncture.

References

Golluscio, L. 2010. "Ditransitives in Mapudungun", en: Malchukov, A., M. Haspelmath y
B. Comrie (eds.). *Studies in Ditransitive Constructions. A Comparative Handbook*, 711– 756). Berlín: De Gruyter Mouton.

Golluscio, L and Hasler, F. 2017. Jerarquías referenciales y alineamiento inverso en mapudungun. RASAL- *Revista de Lingüística 1*: 69-93. (Referential hierarchies and inverse alignment in Mapudungun)

Salas, A. 2006. *El mapuche o araucano*. Santiago: Centro de Estudios Públicos.

Van Valin Jr, R. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Intraclausal complexity in Spanish: The case of depictive constructions

Sergio Ibáñez
Universidad nacional Autónoma de México (UNAM)

Depictive secondary predicates involved constructions where two events are coded as simultaneously happening, as in the following Spanish examples:

(1) a. *Herminia llegó **contenta** a la escuela*.
'Herminia arrived to the school happy'.
b. *Tolomeo trabajó todo el día **cantando***.
'Tolomeo worked all day singing'.

In (1a) it is reported that Herminia arrived to the school and that, at the same time, she was happy when that happened. In the same fashion, in (1b) the speaker says that Tolomeo worked all day, and that when he did so, he was singing. In this way, these examples imply semantically complex constructions, where the adjective and the gerund in bold letters function as subsidiary predicates in respect to the finite verb, a fact that is captured in the linguistic tradition under the label of secondary predication. Despite the recognition of this predicative identity, in the Hispanic Linguistic bibliography secondary predicates are considered as a kind of adjunct modifiers inside simple clauses (Demonte and Masullo 1999; Palacar and Alarcón 2007, among many others). That is, in syntactic terms they are not treated as predicates. This mismatch between their syntactic and semantic identity is understood if one follows the traditional idea that syntactic complexity only exists between or among clauses, but no inside clauses, because they are considered the basic or smallest grammatical units that express a complete thought. In this sense, the syntactic status of secondary predicates as a kind of adjunct is a consequence of the lack of a theoretical and conceptual frame that allows to treat them syntactically as predicates, and the clauses where they appear as complex structures.

This situation is perfectly solved by RRG (Van Valin 2005), a theoretical model that starts from the idea of the *layered structure of the clause* (LSC) inside which to more basic syntactic units are recognized: The *nucleus* and the *core*. The first one is the primary constituent of the clause and it contains the predicate (usually a verb); the core subsumes the nucleus and its arguments; the core plus the non-arguments (the periphery) forms the clause.

In correspondence with the LSC, RRG posits that each one of these basic units can form complex linkages, so two or more nuclei can constitute one complex nucleus, and two or more cores can be united to give rise to a complex single core, just the same way as two or more clauses form a complex sentence. In RRG the nature of the units involved in a complex relation is called *juncture* (nuclear juncture, core juncture and clausal juncture).

Additionally, RRG considers that there are three types of linkages or nexus: coordination, subordination and a third type that is usually not recognized in traditional, structural and generative grammars: cosubordination. On the one hand, following the traditional characterizations, coordination can be defined as a relation between two (or more) independent and free units (-dependency, -integrated), while subordination is a linkage that implies that one unit, the subordinate one, syntactically depends on the other, the matrix, and that it is embedded or integrated inside this last one (+dependency, + integrated). On the other hand, cosubordination is a mixed relation where there is no integration of one unit to the other, but where one unit clearly depends on the other (+dependency, -integrated); typically the dependent predicate is non-finite and obligatorily share one argument with the main verb.

In this paper, I want to use these RRG conceptual and analytic tools to properly describe the Spanish depictive construction exemplified in (1). I argue that the secondary predicate is not syntactically an adjunct but a predicate, and that the constructions where they appear are complex constructions, much

in the same way, as it has been proposed for the same type of constructions (depictive and resultative) in English and other languages (Van Valin and LaPolla 1997; Van Valin 2005). I also argue that the constructions in (1) have the cosubordinate nexus type. Nevertheless, I will show that, in a different way to what has been proposed for English, the Spanish adjectival construction illustrated in (1a) implies a core level juncture and not a nuclear juncture, as it is also the case of the example in (1b) with a gerund as the secondary predicate.

In this way, the main contributions of this work are, firstly, to prove the validity of the RRG tools for describing a phenomenon that has not yet receive an accurate description in the current Hispanic Linguistic bibliography; and secondly, to show that similar constructions (the adjectival ones) in different languages (Spanish and English) can imply a distinct level of juncture (core vs nuclear).

References

Demonte, V. and P. J. Masullo. 1999. "La predicación: Los complementos predicativos", in Bosque I. and Demonte V. (eds.), *Gramática descriptiva de la lengua española*, vol. II, Madrid, Espasa-Calpe, pp. 2.461-2.523.

Palancar, E. and L.J. Alarcón. 2007. "Predicación secundaria depictiva en español". *Revista Española de Lingüística*, 37, pp. 337-370.

Van Valin. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.

Van Valin, R. and R. J. LaPolla, 1997, *Syntax, Structure, Meaning and Function*. Cambridge: Cambridge University Press.

When intransitives behave like passive: De-causativization in Japanese

Hideki Kishimoto
Kobe University

Cross-linguistically, verbs like 'break', 'burn', 'melt', and 'open' typically participate in inchoative/causative alternation, but verbs like 'dance' and 'work' do not. One salient semantic property of the former 'inchoative/causative verb' pairs is that an agentive meaning is not encoded in the inchoative (or intransitive) verbs.

(1) a. John broke the vase.
b. The vase broke.

Due to the difference in the meanings of the pairs of verbs, it is often assumed that two distinct semantic representations, i.e. decomposed Logical Structures in RRG, are related to yield the transitivity alternation (see e.g. Van Valin 1993).

(2) **[do'** (x, ϕ)] CAUSE [BECOME **broken'** (y)] ←→ [BECOME **broken'** (y)]

Since the LS on the left contains two variables, which are realized as marcrorole arguments, it represents the meaning of a transitive clause like (1a). The LS on the right does not have part of LS representing an activity, so that the LS is related to an intransitive clause like (1b).

Nevertheless, it is observed (e.g. Haspelmath 1993) that in certain languages, intransitive (inchoative) verbs sometimes include an agentive meaning, while English lacks this option entirely. Japanese has such intransitive verbs, as in (3).

(3) a. Kodomo-ga nantoka (*karera-ni) tasukat-ta.
child-TOP somehow them-by be.rescued-PST
'The child was somehow rescued (by them).'
b. Kodomo-ga nantoka (karera-ni) tasukee-rare-ta.
child-TOP somehow them-by rescue-PASS-PST
'The child was somehow rescued (by them).'

In Japanese, intransitive verbs are related to transitive verbs via morphological affixation of an intransitivizing or a transitivizing suffix. The intransitive *tasukar-u* in (3a) differs morphologically from the passive verb *tasuke-rare-ru* in (3b), derived by combining the transitive *tasukaru* with the passive *rare*. The described event in (3a) (in the intended sense) cannot be realized unless some agentive action is involved (and thus the meaning of the intransitive clause in (3a) can only be expressed by a passive clause in English). Despite the verb's carrying an agentive meaning, the agent, which is implied by the meaning of the verb, can never be realized.

On the other hand, there is also a class of intransitive inchoative verbs that allow an agent to be manifested with morphologically oblique marking, as in (4a).

(4) a. Gootoo-ga (keikan-ni) tukamat-ta.
burglar-NOM police-by be.caught-PAST
'The burglar was caught (by the police).'
b. Gootoo-ga (keikan-ni) tukame-rare-ta.
burglar-NOM police-by catch-PASS-PAST
'The burglar was caught (by the police).'

(4a) is an intransitive clause, but has a passive-like form because the agent is obliquely marked in a way similar to the passive clause in (4b).

The facts of the intransitive verbs in (3a) and (4a) raise the theoretically interesting question of why the agent can be realized in one class of intransitive verbs but not in the other class, even though both classes of verbs carry agentive meanings. To account for the facts on the class of intransitive verbs including *tasukaru* 'be rescued' in (3a), Kageyama (1996) suggests that verbs like (3a) include the meaning of an agent act but its participant is not realized. While Kageyama does not fully specify how the agent is prevented from being realized, I suggest that intransitive verbs carrying agentive meanings have the LS expressing agentivity, where a variable is replaced by a constant.

(5) **[do'** (C, ϕ)] CAUSE [BECOME **rescued'** (y)]

In (5), the constant saturates the slot in the activity LS, and hence is not linked to an argument, the result of which is that the agent cannot be realized. Since (5) has only one argument to be realized as a macrorole argument, it follows that even if the agent can be identified contextually, it cannot appear as an argument in the clause.

This analysis faces a challenge in accounting for the facts of (4a), where an obliquely-marked agent is present. In regard to (4a), I propose that the agent is made available by virtue of equating the constant agent with a locative argument, as represented in (6).

(6) **[do'** (C_i, ϕ)] CAUSE [BECOME **captured'** (y) & **be-at'** (z_i, y)]
└──────────────────────────────────────┘

The line under the LS indicates an identity relation between the constant C and the variable *z*. When this relation is established, the agent can be linked to a variable in the LS, and hence, it can be realized in the clause. In (6), the variable *y* is the only macrorole argument, and hence is realized as the subject. On the other hand, the variable *z* is a location, and hence is realized as an oblique argument. Note that unlike the passive clause (4b), (4a) does not involve any marked macrorole assignment, i.e. the arguments are aligned in the usual linking algorithm.

The present analysis entails that the agent is counts as a location in (4a), and accordingly, makes the prediction that if no semantic relation is established, a pure location surfaces as an oblique argument. The contrast in acceptability shows that the *ni*-marked argument for the intransitive *tukamaru* is not restricted to an agent.

(7) a. Kuma-ga {wana-ni/ryoosi-ni} tukamat-ta.
bear-NOM trap-by/hunter-by be.caught-PST
'The bear was caught {in the trap/by the hunter}.'
b. {Ryoosi-ga/*Wana-ga} kuma-o tukamae-ta.
hunter-NOM/trap-NOM bear-ACC catch-PST
'{The hunter/The trap} caught the bear.'

The fact suggests that in (7a), the *ni*-marked argument should be a realization of the variable *z* in **be-at'** (z_i, y), which indicates that (7a) is an intransitive clause (and not a passive), and carries the meaning of 'the event of the bear's getting caught takes place at some location, which could be identified as the agent.

I also present other empirical data showing that the relevant *ni*-marked argument is a location rather than a genuine agent. The core claim in this paper is that since Japanese has a semantic means of equating the location with the agent, the intransitive clause (4a), which expresses approximately the

same argument realization pattern as the passive clause (4b), allows an 'apparent' agent to be expressed overtly.

References

Haspelmath, Martin (1993). "More on the typology of inchoative/causative verb alternations." In B. Comrie and M. Polinsky (eds.) *Causatives and Transitivity*, 87-120. Amsteradam: John Benjamins.

Kageyama, Taro (1996). *Doosi Imiron* [Verb Semantics]. Tokyo: Kuroshio Publishers.

Van Valin, Robert D Jr. (1993). "A synopsis of Role and Reference Grammar." In Robert D. Van Valin Jr. (ed.) Advances in Role and Reference Grammar, 1-165. Amsterdam: John Benjamins.

Degrees of discourse configurationality and beyond

Anja Latrouite & Robert D. Van Valin, Jr.
Heinrich Heine University Düsseldorf & University at Buffalo

The notion of 'discourse configurational language' was introduced by É. Kiss (1995) to describe languages like Hungarian in which clause-structure is organized in terms of discourse-related notions like topic and focus rather than grammatical relations like subject and object or traditional X-bar constituent structure. The preverbal field in Hungarian has a topic position, followed by a pragmatically more neutral position, followed by a contrastive focus position, while the post-verbal field is not subject to such constraints. Rizzi's (1997) articulation of the left periphery of the clause included TopicP and FocusP functional projections, and while these are part of the grammar of every language, they play a central role in the grammar of discourse configurational languages.

Discourse configurationality is not an all or nothing phenomenon, and languages vary as to their degree of discourse configurationality. In particular, they vary in terms of what grammatical phenomena are sensitive to discourse functions like topic and focus, and to what extent these notions are 'structure-building'. RRG is particularly well-suited to capture these effects, despite not having functionally-specific syntactic structures akin to TopicP and FocusP,as e.g. shown by Bentley (2008). The key to the RRG approach is the enhanced information-structure projection, which has represented the focus-background distinction by means of the contrast between the PFD and the AFD since 1993. Balogh (2019) adds a topic-comment notation to the information units; as illustrated in Fig.1.

Figure 1: Enhanced information-structure projection, following Balogh (2019)

So for Hungarian the first RP in the core would be '[IU]$_{TOP}$', and the remainder of the core would be within '[…]$_{COMM}$'. The PFD would encompass the entire clause, and the AFD would include the pre-nuclear position. The position between the topic and the AFD would be part of both the background (PFD minus the AFD) and the comment, and accordingly it is pragmatically neutral, unlike the positions on either side of it. The information structure projection, thus, an overlay over the constituent projection, providing a rigid bracketing of the constituents that form the Hungarian prefield. This is a clear instance of structure building.

Another example of structure building comes from English, a language not considered to be discourse-configurational. RRG does not have VP as a part of the LSC, and denies that VPs are universal; nevertheless, some languages, e.g. English, clearly have them. Van Valin (2005) shows that one of the sources of VP-like groupings is information structure, where it imposes bracketings on the

constituent projection (cf. the left diagram in Figure 1), yielding units that are involved in topical VP-fronting and VP-ellipsis constructions. This shows that even in a less discourse configurational than Hungarian, there are some traces of IS-driven structure building.

Note that not every instance of marked syntactic structure associated with a special IS-interpretation is viewed as IS-driven or IS-mediated in RRG. The *it*-cleft construction, which comes with a number of different IS-interpretations (cf. Declerck 1984), suggests that we need to assume more than one constructional schema with IS-specification.

The IS notions relevant to grammar and structure building may come with their own language-specific flavor. The topic-prominent languages Japanese and Korean both exhibit topic markers. However, while Korean restricts the marker to continued givenness topics, Japanese employs the topic-marker also for not yet explicitly given aboutness topics that may be assumed as part of the general common ground or as evocable from given linguistic material (cf. Lee & Shimojo). A further important difference is that in Japanese the *wa*-marked aboutness topic may only appear in sentence-initial position (non-initial *wa*-marked constituents are focal), while the *nun*-marked givenness topic may appear in different positions in the sentence. Therefore, the structure-building effect of the topic in Japanese is more pronounced than in Korean. The constraint on *wa* motivates the postulation of an LDP in Japanese, which would be a case of structure-building, while the lack of a comparable constraint in Korean may be indicative of the lack of an LDP.

(1) Vermeulen (2009: 352)

	Hmm,	*ano*	*CD-wa*	*doo-da-ka*	*Siranai*	*kedo…*
	Well	that	CD-WA	how-COP-whether	not.know	But
	„Well, I don't know about that CD, but…"					
a.	*#John-ga*	*Sue-ni*	*ano*	*hon-**wa***	*Kinoo*	*ageteita(yo)*
	John-NOM	Sue-to	that	book-WA	yesterday	gave(PRT)
b.	*?John-ga*	*ano*	*hon-**wa***	*Sue-ni*	*Kinoo*	*ageteita(yo)*
	John-NOM	that	book-WA	Sue-to	yesterday	gave(PRT)
c.	*Ano*	*hon-**wa***	*John-ga*	*Sue-ni*	*Kinoo*	*ageteita(yo)*
	That	book-WA	John-NOM	Sue-to	Yesterday	gave(PRT)

'As for that book, John gave it to Sue yesterday.' [Japanese]

(2) Vermuelen (2009:353)

Hmm, ku CD-**nun** molu-keyss-ko
well this CD-NUN not-know-but
'Well, I don't know about this CD, but

a. John-i Sue-hantey i chayk-**un** ecey cwuesse.
John-NOM Sue-DAT this book-NUN yesterday gave

b. John-i i chayk-**un** Sue-hantey ecey cwuesse.
John-NOM this book-NUN Sue-DAT yesterday gave

c. I chayk-**un** John-i Sue-hantey ecey cwuesse.
this book-NUN John-NOM Sue-DAT yesterday gave
'As for this book, John gave it to Sue yesterday.'[Korean]

The differences with respect to the structure-building function of IS requires us to think about what exactly it means for IS to be an overlay of the constituent structure and how this is captured in the linking procedure. We assume that there are three different ways in which IS can interact with morphosyntax, which can be found to different degrees in all 3 language types: discourse-configurational, semi-discourse-configurational languages and non-configurational languages. For canonical sentences that may receive more than one IS-interpretation, the IS-frame follows from the context and simply mirrors the interpretational calculation after the constituent structure-semantic structure linking. For IS-specific constructions IS is part of the constructional schema that describes the linking. If there is more than one IS-interpretation associated, IS determines construction choice, but not the linking per se. For certain aspects of discourse-configurational languages, IS dictates the linking.

References

Balogh, K. (2019) Focus sensitivity in RRG. *RRG Proceedings 2015*. In preparation.

Bentley, D. (2008) The interplay of focus structure and syntax. In R. Van Valin (ed), *Investigations of the syntax-semantics-pragmatics interface*. Amsterdam: John Benjamins. 263-284.

Declerck, R. (1984) The pragmatics of *it*-clefts and *WH*-clefts. *Lingua* 64: 251-289.

É. Kiss, K., ed. (1995) *Discourse configurational languages*. Oxford: Oxford University Press.

Lee, E. & Shimojo, M. 2016. 'Mismatch of topic between Japanese and Korean.' *Journal of East Asian Linguistics* 25(1), 81-112.

Rizzi, L. (1997) The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, 281-337.

Van Valin, R. (2005) *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Vermeulen, R. (2009) Topic, contrast, contrastive topics. Proceedings 'Formal Altaic Linguistics 5', 361-372.

Operators, associated movement and subject in Mazahua (Otomanguean)

Armando Mora-Bustos & Verónica Fidencio Núñez
Universidad Autonoma Metropolitana-Iztapalapa (México)
& Educación Primaria Indígena SEIEM

The objective of this dissertation is to show that the morphemes that encode associated movement are correlated with the morphological realization of the grammatical subject. The verbal word in Mazahua is illustrated in (1). Knapp (2013) has proposed that this language is quasi-polysynthetic. The ordering of operator-expressing affixes and clitics is very small.

A proclitic TAMP and operators like tense, aspect, modality, location, quantification and associated movement appear on the verb. These grammatical categories are treated as operators modifying different layers of the clause. Each of the clause levels may be modified by one or more operators (Van Valin & LaPolla 1997; Van Valin 2004). The information expressed by these pre-nuclear clitics has discursive implications since they are updated in the communicative interaction; consequently, its informative value may be optional. The verb operators are summarized in (1)

1 NEG/AF=PTL=QUAN2=ADV=INT/CSL=TAMP=QUAN1=VERB-OBJ/OI=LIM=E=PL/DL=PTL
=LOC=DEM

It is only possible to find four clitics anchored to the verb on preverbal position. The proclitic TAMP always appears in an adjacent position to or close to the verb; This proclitic is exhibited, in some contexts as zero morphemes (Ø) and it is preceded by at least one or two clitics. This is illustrated in (2).

2 a. já=k^{h}ó=ʃì=ɾò=ndɨ̀nɾĩ=ni; já=tá=ne=Ø=ndɨ̃nɾĩ
PTL=who=PER=3PST=answer=RPT PTL=LIM=REC=3PST=answer
'Who was going to answer him, they say; they have already answered.' (Txt.af.13)
b. hò=ʃo=né=Ø=sɔ́^{ʔ}jɛ=k'i; nù=ʃùßa ndà=méhɛ
ASE=ADI=REC=3COP=sick=DEM.PROX2 ART.NANF=Juan AU=pozo
'Also, that one was sick. Juan de Pozo Grande.' (Txt.r.12)

Clitics than encode associated movement are illustrated in (3); it refers to grammatical morphemes used primarily with non-moving verbs to express that the action of the verb is associated with a displacement (Koch 1984, Guillaume 2017). The expression of the associated movement is encoded through grammatical morphemes and not lexical units.

3 a. tá=má=Ø=ŋgîtʃ˜i jò=sàn.hòse=k$^{w'}$a
LIM=AND=3.PST=get.in ART.PL.NANF=Saint.José=over.there.PROX2
'They got in over there, the people from San José.' (Txt)
b. já=he=mì=ndʒódɨ sà pègɾo, mì=ndʒódɨ=^{ʔ}ja=^{h}nu
PTL=DEA=3COP=walk Saint Pedro 3COP=walk=PTL=PROX3
'I was already walking in San Pedro, they were walking there.' (Txt)
c. já=βì=Ø=ʧhɨ́ntɨ. já=mì=ɓɨɓɨ mì-màma=go=^{ʔ}ja=mɨ
PTL=VEN=3COP=marry PTL=3COP=be 1POS-mother=1E=PTL=then
'He had already married. My mom was still alive.' (Txt)

In this dissertation, it is proposed that associated motion morphemes are nuclear operators. Nuclear operators have scope over the core; they modify the action, the event or the state itself, without reference to the participants. The morphemes that encode associated movement are located next to the

verbal root; above these is only the quantification morpheme and the TAMP morpheme. The general Mazahua operator schema is summarized in (4).

4 a. $_{FI}$DEC<$_{STA}$RE<$_{NEG}$Ø<$_{MOD}$OBL<$_{ASP}$PERF<$_{DIR}$MA<$_{QUAN}$TOT<EL>>>>>>>>>
b. NEG/AF=STA/MOD=ASP=QUAN$_2$=(REC/GEN/INT)=MA=(TAMP)=CUAN$_1$=VERB-OBJ/OI =LIM=E=PL/DL=PTL=LOC=DEM

References

GUILLAUME, A. (2017). Sistemas complejos de movimiento asociado en las lenguas Takana y Pano: perspectivas descriptivas, tipológica e histórico-comparativa. *Amerindia, 39* (1). 211-261.

KNAPP, M. (2013). *Doctrina y enseñanza en la lengua mazahua. Estudio filológico y edición interlineal del texto bilingüe de Nájera Yanguas.* México D.F.: Instituto Nacional de Lenguas Indígenas.

KOCH, H. (1984). The category of 'Associated Motion' in Kaytej. *Language in Central Australia*, 1, 23-34.

VAN VALIN, R. 2005. *Exploring the Syntax and Semantic Interface*. Cambridge: Cambridge University.

VAN VALIN, R. & R. LAPOLLA. 1997. *Syntax. Structure, meaning, and function*. Cambridge: Cambridge University.

Accommodating Irish Sign Language in an extended Role and Reference Grammar lexicon architecture

Irene Murtagh
Technological University Dublin – Blanchardstown Campus

This paper is concerned with the definition of a Role and Reference Grammar (RRG) lexicon architecture (Nolan 2011a, 2011b) to accommodate the linguistic phenomena unique to sign languages, in particular to this research Irish Sign Language (ISL), in linguistic terms. To date, there is no definition for the architecture of the ISL lexicon in computational terms.

ISL is a visual gestural language articulated in 3D space with no written or aural form. ISL is a linguistically complete, very rich and complex language. Communication across Sign Languages (SLs) occurs using visual-gestural modality, encompassing Manual Features (MFs) and Non-Manual Features (NMFs) (Leeson and Saeed, 2012). MFs include hand shapes, hand locations, hand movements and orientations of the palm of the hands. NMFs include the use of eye gaze, facial expression, mouthing, head and upper body movements. The visual gestural realisation of a word in SL involves the simultaneous and parallel expression of a varied number of MFs and NMFs, each with their own duration, orientation and relative configuration and movement.

In pursuit of defining a lexicon architecture to accommodate ISL, we argue that the theory of qualia structures defined within the theory of the Generative Lexicon (GL) (Pustejovsky, 1991) must to be extended to cater for SLs and their associated linguistic phenomena. We propose that semantic properties, which contribute to the meaning of a sentence, will need to be extended to accommodate ISL within the lexicon.

We motivate a new level of lexical meaning termed *Articulatory Structure Level*, such that the computational phonological parameters associated with this visual gestural language are accommodated. This level of lexical meaning will represent the essential (computational) phonological parameters of the lexical item. These parameters will be used to account for the various linguistic phenomena pertaining to ISL MFs and NMFs, which are necessary to adequately represent ISL within an RRG lexicon architecture (Van Valin and La Polla, 1997; Van Valin, 2005).

We refer to our newly developed linguistically motivated computational framework as the *Sign_A* framework, with the "A" within this term representing Articulatory Structure Level. We leverage our proposed Articulatory Structure Level for lexical meaning to accommodate the linguistic phenomena of ISL and to propose a lexicon architecture capable of accommodating ISL in computational linguistic terms. The Sign_A framework together with Articulatory Structure Level, enables us to provide a definition within RRG for the ISL lexicon in computational linguistic terms. We leverage the Sign_A framework to extend the RRG model to account for lexical entries for ISL verbs, ISL classifiers and ISL nouns within the RRG lexicon.

References

Leeson, L. and J. Saeed 2012. *Irish Sign Language*. Edinburgh, UK: Edinburgh University Press.

Nolan, B. 2011a. *Constructional templates at the morphology-lexicon interface – meaning and the layered structure of the Irish word*. Paper presented at the workshop on Meaning construction at the crossroads of grammar, cognition and communication. Societas Linguistica Europaea, University of La Rioja, Logroño, Spain.

Nolan, B. 2011b. Meaning Construction and Grammatical Inflection in the Layered Structure of the Irish Word: An RRG Account of Morphological Constructions. In: Wataru Nakamura (ed.). *New perspectives in Role and Reference Grammar*. Newcastle upon Tyne: Cambridge Scholars Publishing. 64–103.

Pustejovsky, J. 1991. *The Generative Lexicon*. *Computational Linguistics*, Volume 17(4), pp. 209–441.

Van Valin, R. and R. La Polla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.
Van Valin, R. 2005. Exploring the Syntax-Semantics Interface. Cambridge: Cambridge University Press.

The genitive case in Irish: A functionalist account of complex multifunctional syntax

Eibhlín Ní Fhallamháin
Trinity College Dublin

Most languages have a distinctive word class of adjectives but also use nouns as modifiers. The morphological patterns associated with modifying nouns are generally referred to as the 'genitive case'. Although the prototypical meaning of the genitive case in many languages is generally defined as possession, it marks a wide range of adnominal relations (Carlier & Verstraete, 2013). The structures using the genitive case in Irish are some of the most frequently used, and some of the most complex. This study explores these complexities, and assumes they are motivated by functionality.

Role and Reference Grammar (RRG) and some compatible theoretical work are used as the analytical framework for this study (Van Valin, 2005; Van Valin and LaPolla, 1997). Adjectival and nominal modification are usually defined as providing additional information about a characteristic of a noun. This paper provides specificity to this definition vis-à-vis nominal attribution in a way compatible with RRG. Generative Lexicon Theory, specifically Pustejovsky's qualia analysis (Pustejovsky, 1995), is used to explain the modifying functions of the genitive case in Irish. Rijkhoff's work on nominal aspect markers (Rijkhoff, 2004, p. 120-121) is also used in analysing the expression of quantification in Irish. The concept of syntactic templates in RRG is used throughout. Syntactic patterns are described in RRG as templates stored in a syntactic inventory, a sort of 'constructicon'; language specific features of clause and NP structure are represented in the syntactic templates (Van Valin, 2005, p.13). The syntactic patterns of the NP in Irish are used to represent argument structure, and this is analysed using an RRG framework.

This paper demonstrates that the genitive structure in Irish performs three major functions, to specify a quale, quantification, and to encode relations between an argument and a nominalised form of the verb. There are two different types of modifying noun in Irish, one which follows the head noun in the genitive case and the other before the head noun and in the common case. Both modify the nucleus of the noun phrase, but differently. Modifying genitive nouns in Irish specify a quale and provide additional detail. The syntactic construction used to encode attribution is also used to encode quantification, but the quantifier as an Irish core modifier precedes the modified noun. Unlike the other attributive uses of the genitive, the privileged syntactic argument (PSA) is the noun in the genitive. A noun followed by one of these quantifying terms may be an argument of the main noun, including as the subject of the clause. The quantifying nouns are core operators, and the attributive nouns are part of the nuclear periphery. The Noun-Genitive Noun structure is also used to encode an event as an attribute of the undergoer and is used to form progressive structures. The Verb-Undergoer use of the Noun-Genitive Noun structure is motivated by the parallels between clause and noun phrase structures. The Verb-Undergoer relation is a semantic link that is considered to be a core relation in RRG. Most languages have the same PSA for most syntactic constructions. In Irish generally, the actor is the subject and is the PSA in the common case. The default in Irish is for the undergoer to be the object and follow the subject in the common case. The progressive forms which use the verbal noun followed by the genitive case have a different structure, the structure of a noun phrase, and therefore have a different mapping of roles. This Verbal Noun-Undergoer structure in Irish has the same logical structure as a clause and the syntax of a noun phrase.

Analysis of the Noun-Genitive Noun structure in RRG terms highlights the nature of the differences between two patterns of modification and allows them to be more accurately characterised. The Noun-Genitive Noun template in Irish is shown to be multifunctional; the parallelism in clause and Noun Phrase structure is exploited to extend the range of possible expression.

References

Pustejovsky, J. (1995) *The Generative Lexicon*. Cambridge, MA: MIT Press

Rijkhoff, J. (2004). *The noun phrase* (Oxford studies in typology and linguistic theory). Oxford: Oxford University Press

Van Valin, Jr., R. (2005). *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press

Van Valin, Jr., R. and LaPolla, R. J. (1997). *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press

Different degrees of agentivity in sentience verbs?
On the decomposition of proto-agent features in Polish

Maria Katarzyna Prenner
University of Cologne

In my talk I want to address the question whether agent prominence and flexible feature priorisation can provide a better explanation for agentivity effects than the prototype approach and feature accumulation in Dowty's (1991) sense. I provide new empirical evidence from two rating experiments on active, passive and impersonal constructions in Polish.

Polish has a wide range of impersonal constructions that have no overt subject. One of them is the so called *-no/-to construction* as exemplified in (1):

(1) *Wypi-to* *cał-**ą*** *butelk-**ę**.*
drink-PST.IMPRS whole-F.ACC.SG bottle.F-ACC.SG.
'People drank up the whole bottle.'

Following Malamud (2013), we refer to such constructions as *arbs*, which is short for "constructions with arbitrary interpretations". (The term encompasses several distinct readings, cf. Cabredo-Hofherr 2003). In contrast to canonical passives as in (2), arbs can be formed from all kinds of verbs, including unergative and unaccusative intransitive verbs (Kibort 2008: 265, 271; Krzek 2011: 68-69), and they have the ability of binding a direct object in the accusative case. The implicit subject is obligatorily interpreted as [+human] (e.g. Kibort 2008: 267; Padučeva 2012: 29; Sansò 2006: 255).

(2) *Cał-**a*** *butelk-**a*** *został-a* *wypi-ta*.
whole-F.NOM.SG bottle.F.NOM.SG COP.PST-F.SG drink-PTCP.PASS-F.NOM.SG
'People drank up the whole bottle.'

Recent acceptability judgement studies for Polish (Bunčić to appear) show that in arbs, some verbs perform better than others and some are rather unacceptable, even if the implicit subject is human. This finding leads to the assumption that the availability and acceptability of arb constructions is constrained by an agentivity cline.

Dowty (1991: 572) defines two superordinate proto-roles, the *proto-agent* and the *proto-patient*, only by bundles of entailments generated by the verb's meaning. In Dowty's framework, the proto-agent is defined by five entailments or features: *volition*, *sentience*, *causation*, *movement* and *independent existence*. A 'maximum proto-agent' would exhibit *volition, sentience* and *movement* (Primus 2012: 27). The prototype is considered as the privileged candidate for agent demotion in impersonal constructions and passives. However, Dowty's proposal about feature accumulation being the key indicator for the agent prototype does not hold across linguistic constructions. Evidence from rating experiments for sentience verbs in German (Kretzschmar et. al. to appear) supports this. Kretzschmar et al. propose that the proto-agent features entailed by the verb may be ranked depending on various factors, including the language, the construction itself and its discourse function which meets the criteria of the notion of prominence as elaborated in Himmelmann & Primus (2015). Even for sentience verbs with equal number and identical proto-agent features, there are differences in acceptability in the tested linguistic constructions, both for German and Polish. This, again, leads to the assumption that *sentience* is not an atomic feature und thus has to be decomposed further by differentiating *perception*, *emotion* and *cognition* verbs from each other (e.g. Van Valin, 1999; Viberg, 2001).

In two rating experiments focusing on the same transitive sentience verbs as in the study for German (mainly (i) perception verbs like *widzieć* 'see', (ii) emotion verbs like *nienawidzić* 'hate' and (iii)

cognition verbs like *znać* 'know'), I want to prove the prototype approach and the prominence hypothesis on the example of the *-no/-to construction*, passive and active. The test items for the study are distributed over five different verb groups, consisting of the three above mentioned (i-iii) and two more verb groups that exhibit either three features (iv) (*volition, sentience* and *movement*), e.g. *obserwować* 'watch', or lack any of the agentive features under discussion (v), e.g. *sprawować* 'exhibit'. The active is compared with the passive and *-no/-to impersonal*, respectively. For each of the five verb classes, 4 verbs with uniform semantic and syntactic behavior have been identified. The patient argument in the test items is always inanimate and in the accusative case. For each individual verb, ten different sentences have been constructed, resulting in 40 different sentences per verb class condition. Additionally, 12 ungrammatical control items have been constructed by using dative object-experiencer verbs in the passive voice. A 6-point Likert scale ranging from "very unacceptable" to "very acceptable" is the basis for evaluating the sentences.

The prototype approach predicts that verbs with volitional agents such as selected by the verb group (iv) 'watch' will be rated better than non-volitional experiencers (i-iii), regardless of the linguistic construction. I will discuss the test results and compare to what extent they are in line with the prominence approach and how they differ from the data on German verbs.

References

Bunčić, Daniel (2019), Agent prominence in the Polish *-no/-to*-construction. In: A. Bauer and D. Bunčić (eds.), *Linguistische Beiträge zur Slavistik: XXIV. JungslavistInnentreffen in Köln 17-19, September 2015*. Frankfurt am Main: Peter Lang, 63-76.

Cabredo Hofherr, Patricia (2003), Arbitrary readings of 3pl pronominals. In: M. Weisgerber (ed.), *Proceedings of Sinn und Bedeutung (SuB) 7*, 81–94. Konstanz: Universität Konstanz.

Dowty, David R. (1991), Thematic proto-roles and argument selection, *Language* 67(3), 547-619.

Gast, Volker and van der Auwera, Johan (2013), Towards a distributional typology of human impersonal pronouns, based on data from European languages. In: D. Bakker et al. (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, Berlin: 119–158.

Himmelmann, Nikolaus P. and Primus, Beatrice (2015), Prominence beyond prosody. In: A. De Dominicis (ed.), *Prominences in Linguistics. pS-prominenceS: Prominences in Linguistics. Proceedings of the International Conference,* Viterbo: DISUCOM Press, 38–58.

Kibort, Anna (2008), Impersonals in Polish: An LFG perspective. *Transactions of the Philological Society* 106(2), 246–289.

Kretzschmar, Franziska et. al. (2018), What is a sentient agent? Pre-Proceedings of the International Conference of Linguistic Evidence 2018, Tübingen: 47-49.

Kretzschmar, Franziska et. al. (to appear), An experimental investigation of agent prototypicality and agent prominence in German.

Krzek, Małgorzata (2011), Impersonal *się* constructions in Polish, *Newcastle Working Papers in Linguistics* 17, 67– 93.

Malamud, Sophia A. (2013), (In)definiteness-driven typology of arbitrary items, *Lingua* 126, 1–31.

Siewierska, Anna and Papastathi, Maria (2011), Towards a typology of third person plural impersonals, *Linguistics* 49(3), 575-610.

Padučeva, Elena Viktorovna (2012), Neopredelenno-ličnoe predloženie i ego podrazumevaemyj sub"ekt [The undetermined-personal sentence and its implicit subject], *Voprosy jazykoznanija* 1, 27–41.

Primus, Beatrice (2006), Hierarchy mismatches and the dimensions of role semantics. In: I. Bornkessel, M. Schlesewsky and B. Comrie (eds.), *Semantic Role Universals and Argument Linking. Theoretical, Typological and Psycholinguistic Perspectives*, Berlin: de Gruyter. 53-88.

Primus, Beatrice (2012), Semantische Rollen, Heidelberg: Carl Winter.

Sansò, Andrea (2006), "Agent defocusing" revisited: Passive and impersonal construction in some European languages, *Typological Studies in Language* 68, 232-273.
Viberg, Åke (2001) The verbs of perception, In: M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds.), *Language typology and language universals*, 2nd vol., Berlin: de Gruyter. 1294–1309.
Van Valin, R. D (1999), Generalized Semantic Roles and the Syntax-Semantics Interface. In: F. Corblin, C. Dobrovie-Sorin and J. M. Marandin (eds.), *Empirical Issues in Formal Syntax and Semantics*, The Hague: Thesus, 373-389.

On the interactions with pragmatics in Role and Reference Grammar

Mitsuaki Shimojo
University at Buffalo

Unlike Chomskyan theories, in which syntax is the core component providing input to semantics and pragmatics, parallel architecture theories consist of equally dominant components which interact with each other directly (Jackendoff 2002). Role and Reference Grammar employs parallel architecture, in which syntax, semantics, and pragmatics are independent components but receive input from each other so that they jointly represent a sentence (Van Valin 2014).

With respect to the role of pragmatics as part of the parallel architecture, Van Valin (2005: 182) states that "discourse-pragmatics can influence every aspect of grammar," and many examples have been discussed which show the pragmatic input on syntax, semantics, and the linking algorithm which relates the two. Yet, what has to be demonstrated is the bi-directionality of the interactions with pragmatics, with descriptions of the influence on pragmatics by the other components of grammar. The purpose of this presentation is to show how the discourse representations are influenced by the semantic and syntactic properties of a sentence and to show what input the discourse representations receive from the grammar so that the cognitive model of context can serve as common ground for the subsequent development of discourse.

The discussion will be based primarily on discourse analysis in Japanese. In particular, I will discuss observations which relate to the selection of privileged syntactic argument [PSA] and the use of different argument forms (zero anaphora and post-nominal markings) and post-verbal arguments (in a postcore slot and a right-detached position). Each of these serves as input for the discourse representations because the choice of the forms reflects the speaker's pragmatic intentions about the information conveyed and the relative importance of the information for the purpose of the given discourse. For example, the referents represented by a PSA are more likely to persist in the subsequent discourse (Shimojo 2016); hence, they are intended to be more important. Yet, not all PSA's are presented equally. Among the possible argument forms, ellipsed, topicalized, and case-marked arguments are more likely to persist than those without any post-nominal marking, and those placed in the pre-verbal positions are more likely to persist than those placed post-verbally (Shimojo 2005). These observations suggest that the structural properties serve as "mental processing instructions" (Givón 1993) for the hearer/reader to construct and update their episodic memory as the discourse is processed.

With the discourse-based observations, I propose a mechanism to represent the salience ranking of referents in the discourse representation structure and argue that the ranking is constantly updated on the basis of input from syntax and semantics and that the updated discourse representation serves as input for the construction of subsequent sentences in turn. This new mechanism is necessary because salience is related to continuity of information, and it does not necessarily correlate with presupposition or assertion, which derives the focus structure of sentence.

References

Givón, Talmy. 1993. Coherence in text, coherence in mind. Pragmatics & Cognition 1, 171–227.
Jackendoff, Ray. 2002. The Architecture of the Language Faculty. Cambridge, MA: MIT Press.
Shimojo, Mitsuaki. 2005. Argument Encoding in Japanese Conversation. Hampshire and New York: Palgrave Macmillan.
Shimojo, Mitsuaki. 2016. Saliency in discourse and sentence form: zero anaphora and topicalization in Japanese. In M. M. Jocelyne Fernandez-Vest and Robert Van Valin Jr. (eds.), Information Structure and Spoken Language in a Cross-Linguistic Perspective, 55-75. Berlin & New York: Mouton de Gruyter.
Van Valin, Robert D. Jr. 2005. Exploring the syntax-semantics interface. Cambridge University Press.

Van Valin, Robert D. Jr. 2014. On the place of information structure in a grammar. Comunicación, Cognición, Cibernétic@. Actas del XXXI Congreso de AESLA, 86-106.

Syntax of Japanese dish names on restaurant menus: An emerging structure

Kiyoko Toratani
York University

Japanese dish names on restaurant menus are normally presented in a simple noun phrase, as in *Tenpura soba* '*Soba* [buckwheat noodle] with tempura'. But according to Toratani (2018), some recent dish names take a peculiar structure, where they end in *TE* as in (1).

(1) *kureepu shuzetto banira muusu o soe-**te***
Crêpe Suzette vanilla mousse$_{ACC}$ add-TE
'Crêpe Suzette, adding vanilla mousse-TE' (Toratani 2018: 289)

In narrative, *TE* normally functions as a clause linkage marker requiring a matrix predicate such as *o-dashi-shimasu* 'serve' as in (2).

(2) *banira muusu o soe-**te*** *<u>o-dashi-shimasu</u>*.
vanilla mousse$_{ACC}$ add-TE HONORIFIC-serve-POL
'I will <u>serve</u> (it to you), adding vanilla mousse.'

However, the dish name in (1) cannot overtly realize the matrix predicate, as it will turn the dish name into a sentence, rendering it inappropriate as a dish name (#'I will serve Crêpe Suzette to you, adding vanilla mousse'). Toratani (2018: 290) speculates that the *TE*-marked unit in (1) is a nominal phrase. This paper argues it is a sentential adjunct, drawing on Van Valin (2005).

The structure of (1) must look like Figure 1, where the *TE*-marked unit occurs in the Reference-Phrase[RP]-level periphery, modifying the "head" (the main dish name, 'Crêpe Suzette'). It parallels that of English non-restrictive relative clauses, since the *TE*-marked phrase adds only additional information about the main dish (see Van Valin 2005: 222-223).

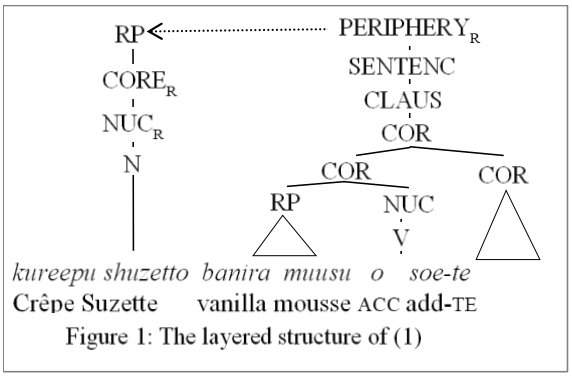


Figure 1: The layered structure of (1)

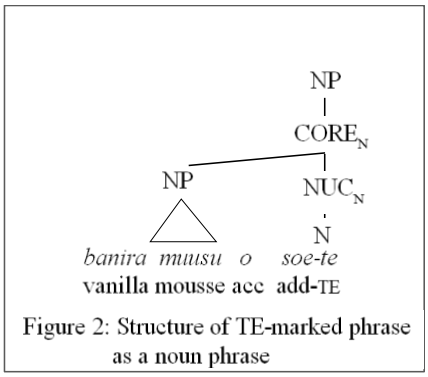


Figure 2: Structure of TE-marked phrase as a noun phrase

Contra Toratani (2018: 290), the unit ending in *TE* cannot be nominal phrase, as it will provide no place for the matrix predicate (see Figure 2), even if it is not overtly realized. In (1), the *TE*-marked phrase is "inference-intensive" (Ohoti 1995: 213), whereby a limited set of matrix predicates is inferred to follow after *TE*, such as *o-dashi-shimasu* '(I will) serve' (cf. (2)). The unit containing the *TE*-marked phrase is posited to be an instance of core cosubordination (Hasegawa 1996: 197), since a core-level modal operator (e.g. *-nakerebanaranai* 'must') would have scope over both cores: one containing *soe-te* 'adding' and the other containing the matrix predicate 'serve', albeit inferred (i.e., what the chef must do is: 'serve by adding', not just 'serve'). It is critical to recognize the presence of the inferred matrix predicate, as it instantiates the proper construal of (1): "Crêpe Suzette, which I [the chef] will serve to you [the diner] by adding vanilla mousse.'.

Following Van Valin (2005: 173), this paper proposes to use discourse representation structure to account for the construal of (1), along the lines represented in Figure 3. It will (i) enable us to directly link event participants and the inferred predicate to semantics, and (ii) provide a place (in "presupposition") to detail the perceived social relations of the chef relative to the diner, where the latter is higher than the former, implying that the matrix predicate must be honorific-marked, if it is ever realized.

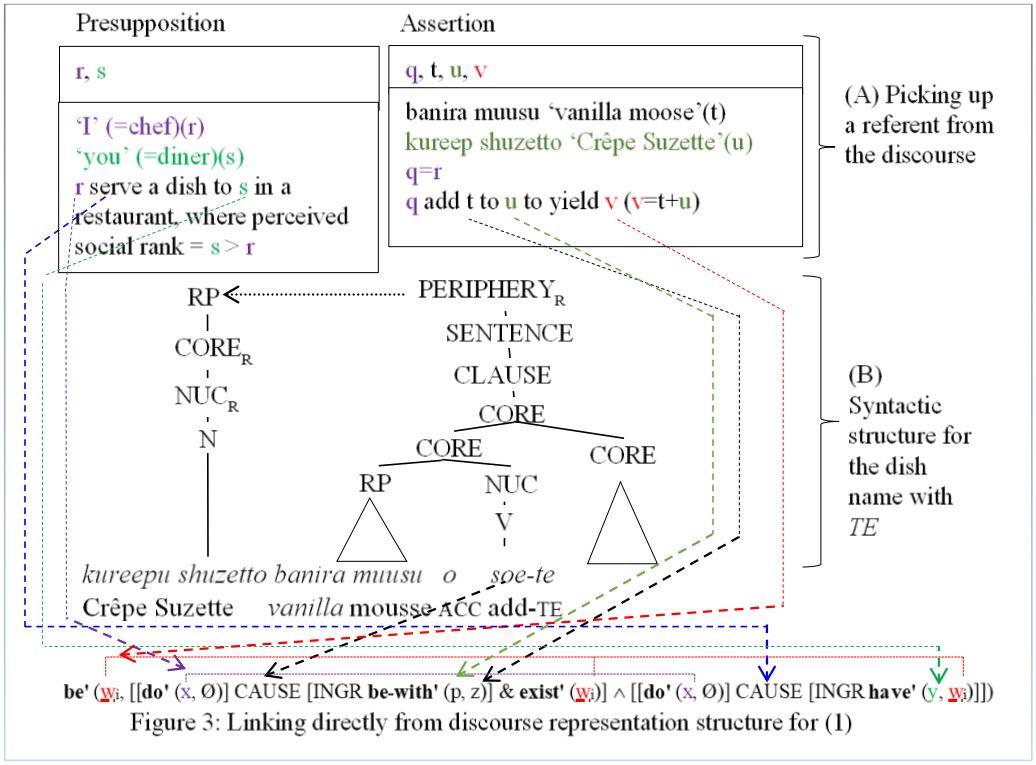


Figure 3: Linking directly from discourse representation structure for (1)

This paper presents evidence from predicate-less phrases, such as (1), on restaurant menus to argue that RRG is architecturally well-equipped to offer an account of the "missing" elements, in agreement with Shimojo (2008).

References

Hasegawa, Y. 1996. *A study of Japanese clause linkage: the connective TE in Japanese*. Stanford: Center for the Study of Language and information.

Ohori, T. 1995. Remarks on suspended clauses: A contribution to Japanese phraseology. Essays in semantics and pragmatics: in honor of Charles J. Fillmore, eds. by M. Shibatani and S. A. Thompson, 201–218. Amsterdam/Philadelphia: John Benjamins.

Shimojo, M. 2008. How missing is the missing verb? The verb-less numeral quantifier construction in Japanese. In Robert Van Valin (ed.), Investigations of the Syntax-Semantics-Pragmatics Interface, 285-304. Amsterdam/Philadelphia: John Benjamins.

Toratani, K. 2018. Japanese -*TE* as an aestheticization marker in the language of menus. *Proceedings of the Forty-Fourth Annual Meeting off The Berkeley Linguistics Society, February 9-11, 2018*, 287-300. Berkeley, CA: Berkeley Linguistics Society.

Van Valin, Robert, D. JR.. 2005. Exploring the syntax-semantics interface. Cambridge: Cambridge University Press.

On the status of lexical categories in RRG

Robert D. Van Valin, Jr.
University at Buffalo & Heinrich Heine University Düsseldorf

Most syntactic theories assume some variety of endocentric syntax, where phrases are projections of lexical or functional/grammatical categories. Moreover, the projecting categories are assumed to be universally valid, i.e. N, V, and A for lexical categories (Baker 2003), and aspect, negation, tense, force, etc. for functional categories. RRG, on the other hand, rejects endocentrism as a fundamental feature of constituent structure. Rather, the two key notions are NUCLEUS, which is a category-neutral unit housing the predicate in the clause, and which, viewed from a cross-linguistic perspective, is not restricted to any particular lexical category or even to being a head (it can be phrasal), and REFERENCE PHRASE [RP], which is a potentially referring expression, which may be headed in principle by a range of lexical categories. The appropriateness of these notions is illustrated in the following well-known example from Tagalog.

(1) a. [CORE [NUC Nagtrabaho] [RP ang babae]].
worked NOM woman
'The woman worked.'
b. [CORE [NUC Babae] [RP ang nagtrabaho]].
woman NOM worked
'The one who worked is/was a woman.'

In (1a) the predicate in the nucleus is a verb, and the head of the RP is a noun, following Himmelmann (2008); in (1b), on the other hand, the predicate in the nucleus is a noun, and the head of the RP is a verb, without derivational morphology or any special morphosyntactic treatment, unlike the corresponding elements in the English translation. Facts like these have led some linguists to argue that Tagalog roots have no inherent lexical category and get assigned one on the basis of how they are used in an utterance (Foley 1998). This is questionable, however, because it confuses lexical category with grammatical function: *nagtrabaho* 'worked' is a predicate in (1a) and an argument in (1b), while *babae* 'woman' is an argument in (1a) and a predicate in (1b), but the category of each word does not change. The point is not that Tagalog possibly lacks lexical categories but rather that there are few restrictions on what can function as the predicate in the nucleus or the head of an RP.

How can RRG approach the issue of lexical categories, given the range of cross-linguistic variation, from languages like Tagalog (Himmelmann 2008), Nootka (Swadesh 1939, Jacobsen 1979) and Lakhota (Boas & Deloria 1942), which seem to make only weak distinctions between nouns and verbs, to languages like Latin, Russian and Dyirbal, which make strict divisions between the two main categories, with English somewhere in the middle? One possibility is to claim that at the most basic level, lexical items fall into one of two classes: they are either REFERRING EXPRESSIONS [REs] or PREDICATES. This is related to the fundamental opposition motivating the layered structure of the clause, namely, the opposition between predicating and non-predicating elements. This distinction derives from the nature of language as a system of communication: communication involves conveying information, which involves propositions, which involve reference and predication, hence REs and predicates. REs and predicates can have one of three grammatical functions: ARGUMENT, PREDICATOR, or MODIFIER. In (1a) a predicate is functioning as the predicator in the nucleus and an RE as the core argument, whereas in (1b) an RE is the predicator in the nucleus and a predicate serves as the core argument. The traditional categories of verb, adjective, adverb and adposition are semantically predicates and grammatically predicators or modifiers, as defaults. In RRG semantic representations it has always been the case that lexical modifiers and meaning-bearing adpositions are represented as predicates.

Defining lexical categories in terms of their morphophonological and morphosyntactic properties leads to the observation that Tagalog, Nootka, Lakhota and other such languages show that languages need not make 'deep' lexical category distinctions, and this seems to follow from ability of lexical items to function as an argument, predicator or modifier rather freely. This ability is an important morphosyntactic property which many or all lexical items share, and it is precisely the morphosyntactic properties that different words do not share that is the basis for assigning them to different lexical categories.

In languages with well-defined lexical categories there are strict constraints on the grammatical function that a given lexical category may have, e.g. nouns (REs) can be arguments, verbs (predicates of certain semantic types) can be predicators, and adjectives and adverbs (predicates of certain semantic types) can be modifiers. These restrictions are among the morphosyntactic properties which define each category. If a category is to have a different grammatical function, e.g. noun or adjective as predicator or verb as argument, special morphosyntactic treatment is necessary, e.g. the use of an auxiliary verb or verbalizing derivational morphology for non-verbal predicators, or nominalizing derivational morphology for non-nominal arguments. These special treatments are also part of the properties defining the different categories. This pattern is found in many European languages, and it has been taken as the norm for human language. This view has led linguists, when confronted with Tagalog-type languages, to posit derivational morphology for verbalization and nominalization which is all marked by zero-morphemes. Thus, on this type of analysis, in (1b) *babae* 'woman' has undergone zero-marked verbalization, yielding 'to be a woman', and *nagtrabaho* 'worked' has undergone zero-marked nominalization, yielding 'the one who worked'. The RRG analysis of (1b) given above does not involve any zero-marked derivational morphology.

What, then, is the status of lexical categories in RRG? The universal semantic distinction is between REs and predicates, which underlies the noun-verb dichotomy. It is analogous to the actor vs. undergoer distinction, which is semantic and universal and which underlies the traditional subject-object dichotomy. The lexical categories beyond noun and verb in a language are differentiations of the functions of predicates as modifiers (i.e. as adjectives and adverbs), and all of the distinctions must be justified morphophonologically and/or morphosyntactically. Adpositional predicates can be arguments, as with verbs like *put*, or adjunct modifiers. Thus, from an RRG perspective, lexical categories are like grammatical relations: language-specific but with a universal semantic foundation. Moreover, they play a rather different role in the non-endocentric syntax of RRG than they play in the endocentric syntax of other theories.

References

Baker, M. 2003. *Lexical categories*. CUP.

Boas, F. & E. Deloria. 1942. *Dakota Grammar*. Washington, DC: US Govt. Printing Office.

Foley, W. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane.

Himmelmann, N. 2008. Lexical Categories and Voice in Tagalog. In *Voice and Grammatical Relations in Austronesian Languages*, 247–293. CSLI Publications, Stanford, CA.

Jacobsen, W. 1979. The noun and verb in Nootkan. *The Victoria Conference on Northwestern Languages*, 83-155. Victoria: British Columbia Provincial Museum.

Swadesh, Morris. 1939. Nootka internal syntax. *IJAL* 9:77-102.

A Role and Reference Grammar of Biblical Hebrew clause linkage

Nicolai Winther-Nielsen
Fjellhaug International University College Denmark; Vrije Universiteit Amsterdam

One of the wexing questions in Biblical Hebrew is the function of *consecutio temporum* 'sequence of tenses'. Traditional primers like Page H. Kelly (1992) will interpret the ubiquitous Hebrew clause linkage marker (CLM) as either a *waw conjunctive* or a *waw consecutive*. The primer will explain that a first verb governs time and mode for subsequent forms, but it will leave to learners to guess at writer intentions - "(t)his literary device gave writers greater flexibility in expressing their thoughts, even though they work with a limited number of verb forms." (p. 210).

A linguistic turn is taking place in the study of Biblical Hebrew. The state of the art linguistic reference grammar of van der Merwe et al (2017) includes an introduction to information structure and offers semantic and pragmatic accounts of some clause relations. But, because Biblical Hebrew does not mark temporal distinctions between past, present, future and modality, translation into tense is still "exclusively determined by the context and the lexical signification of the verb" (p. 154).

While Hebrew information structure is addressed in Winther-Nielsen (MS), this paper will focus on interclausal linkage in order to show how Role and Reference Grammar can provide new insights for Hebrew linguistics. It revisits the earliest work on Hebrew clause linkage (Winther-Nielsen 1995), and it uses the mapping of interclausal relations in RRG (Van Valin and LaPolla 1997; Van Valin 2005). It explores how coordination, subordination and cosubordination nexus at different clause junctures function in a learner corpus used for teaching intermediate level Hebrew grammar and text analysis.

The data for this study consists of a corpus of 2060 clauses or clause fragments selected from a Dutch database that has been developed since 1977 (see the link https://etcbc.github.io/bhsa/ and the RRG-based descriptions of the database in Winther-Nielsen 2008; 2009). Since the Spring of 2015 the database has provided a complete analysis of interclausal connections for the entire Hebrew Bible. In this database clause linkage is coded by numbers consisting of mostly three digits. The first digit X (X**) specifies the connectivity type for a clause and the second digit Y (*Y*) specifies the predicate of this clause, while the Z (**Z) specifies the predicate of the preceding clause, to which it is linked.

Even if this hierarchical structuring of the texts is only a first possible interpretation of the texts, the codes and their occurrence can help linguists select the most frequent types of connections and limit the sample of linkage types to the size of 15 tokens or more for a code (Table 1). It is possible to explore clause linkage for the three major finite conjugations of Biblical Hebrew, the narrative *wayyiqtol*-form (code 7), the perfective *qatal*-form (code 2) and the non-perfective *yiqtol*-form (code 1), and results can be related to linkage with imperatives (code 3), infinitives (code 4) and participles (code 6). This helps an RRG grammarian explore nexus and juncture types in Biblical Hebrew based on frequency and likely function in a given contextual environment.

Table 1. All Codes in Learner Corpus more Frequent than 15

Code	Tense	After	No	Code	Tense	After	No
477	*wayyiqtol*	*wayyiqtol*	181	110	Ø CLM – *yiqtol*	Verbless Cl	21
472	*wayyiqtol*	*qatal*	35	113	Ø CLM – *yiqtol*	imperative	17
427	CLM – *qatal*	*wayyiqtol*	27	130	Ø CLM – imperative	Verbless Cl	18
422	CLM - *qatal*	*qatal*	57	400	CLM – Verbless Cl	Verbless Cl	21
421	CLM – *qatal*	*yiqtol*	40	407	CLM – Verbless Cl	*wayyiqtol*	17
122	Ø CLM – *qatal*	*qatal*	21	103	Ø CLM – Verbless Cl	imperative	23
120	Ø CLM – *qatal*	Verbless Cl	15	107	Ø CLM – Verbless Cl	*wayyiqqtol*	20
527	Reason – *qatal*	*wayyiqtol*	21	101	Ø CLM – Verbless Cl	*yiqtol*	20
522	Reason – *qatal*	*qatal*	17	502	Reason – Verbless Cl	*qatal*	17

411	CLM – *yiqtol*	*yiqtol*	44	64	lᵊ 'to' + infinitive	prep lᵊ	101
413	CLM – *yiqtol*	imperative	21	12	Relative Cl - *qatal*		51
412	CLM – *yiqtol*	*qatal*	19	11	Relative Cl – *yiqtol*		15
111	Ø CLM – *yiqtol*	*yiqtol*	37	16	Relative Cl – Participle		25
112	Ø CLM – *yiqtol*	*qatal*	18				

The case study will use the web-application Bible Online Learner https://bibleol.3bmoodle.dk/ which has been developed since 2008 for corpus driven task-based and persuasive language learning. This application is currently being developed for flexible landing pages, localization and individualization of the interface. A learning designer can then create RRG-defined grammatical terms for the interface, and this can assist researchers who want to explore the use of codes for the study of interclausal connections.

The goal of this project is to explore how we can design a learning environment to train advanced students in Biblical Hebrew RRG. We want to be able to teach nuclear, core and clausal junctures for coordination, cosubordination and subordination through the corpus and explore how interclausal linkage relates to Actionsart and perspectival aspect.

References

van der Merwe, Christo, Naudé, Jacobus, Kroeze, Jan (2017), *A Biblical Hebrew Reference Grammar: Second Edition*, London: Bloomsbury

Kelly, Page H. (1992, *Biblical Hebrew. An Introductory Grammar*, Grand Rapids: Eerdmans

Van Valin Jr., Robert D. (2005), *Exploring the Syntax-Semantics Interface*, Cambridge: Cambridge University Press.

Van Valin Jr., Robert D. and Randall LaPolla (1997), *Syntax: Structure, Meaning and Function*, Cambridge: Cambridge University Press

Winther-Nielsen, Nicolai (1995), *A Functional Discourse Grammar of Joshua: A Computer-assisted Rhetorical Analysis*, Stockholm: Alquist & Wiksell

— (2008), A Role-Lexical Module (RLM) for Biblical Hebrew: A mapping tool for RRG and WordNet. In Robert Van Valin (ed.), *Investigations of the Syntax-Semantics-Pragmatics Interface*, 455-78. Amsterdam and Philadelphia: John Benjamins.

— (2009), Biblical Hebrew Parsing on Display: The Role-Lexical Module (RLM) as a Tool for Role and Reference Grammar. *SEE-J Hiphil* 6:1-51 [http://www.see-j.net/index.php/hiphil/article/view/78).

— (MS), Why Eve Shouldn't Eat the Snake: An Informed Answer from Focus Structure and Reference Tracking in Biblical Hebrew (Revised version of paper presented at RRG Conference on August 2, 2015, at the HH University in Düsseldorf

List of Participants

Name	Affiliation(s)	Email address
Arps, David	Heinrich-Heine-Universität Düsseldorf	david.arps@hhu.de
Balogh, Kata	Heinrich-Heine-Universität Düsseldorf	katalin.balogh@hhu.de
Beck, Hanno	University at Buffalo	hannobec@buffalo.edu
Bellingham, Erika	University at Buffalo	ebelling@buffalo.edu
Belloro, Valeria	Universidad Autónoma de Querétaro	vbelloro@yahoo.com
Bentley, Delia	The University of Manchester	delia.bentley@manchester.ac.uk
Bladier, Tatiana	Heinrich-Heine-Universität Düsseldorf	bladier@phil.hhu.de
Bohnemeyer, Jürgen	University at Buffalo	jb77@buffalo.edu
Bylund, Emanuel	Stockholm University	manne.bylund@biling.su.se
Canu Højgaard, Christian	Vrije Universiteit	chj@dbi.edu
Chiarcos, Christian	Goethe-Universität Frankfurt am Main	chiarcos@informatik.uni-frankfurt.de
Conti, Carmen	University of Jaén	carmenconti@hotmail.com
Evang, Kilian	Heinrich-Heine-Universität Düsseldorf	evang@hhu.de
Everett, Daniel	Bentley University	deverett@bentley.edu
Fäth, Christian	Goethe-Universität Frankfurt am Main	faeth@em.uni-frankfurt.de
Fidencio Núñez, Verónica	Educación Primaria Indígena SEIEM	orevamix@hotmail.com
Fujinaga, Kiyono	University at Buffalo	kiyonofu@buffalo.edu
Generalova, Valeria	Heinrich-Heine-Universität Düsseldorf	generalova.valeria@gmail.com
González, Carlos	Universidad Católica;	cgonzalv@uc.cl
Guerrero, Lilián	Universidad Nacional Autónoma de México	lilianguerrero@yahoo.com
Guerrero, Silvana	Universidad de Chile	silvanagg16@gmail.com
Hafeez, Saima	University at Buffalo	saimahaf@buffalo.edu
Hasegawa, Yoko	University of California, Berkeley	hasegawa@berkeley.edu
Hasler, Felipe	Universidad de Chile	fhasler@uchile.cl
Ibáñez, Sergio	Universidad Nacional Autónoma de México	cecser@prodigy.net.mx
Jaque, Matías	Universidad de Chile	matjaque@uchile.cl
Järnefelt, Pia	Stockholm University	piajarnefelt@gmail.com
Kallmeyer, Laura	Heinrich-Heine-Universität Düsseldorf	kallmeyer@phil.hhu.de
Kawachi, Kazuhiro	National Defense Academy of Japan	kazuhirokawachi@gmail.com

Kishimoto, Hideki	Kobe University	kishimot@lit.kobe-u.ac.jp
Latrouite, Anja	Heinrich-Heine-Universität Düsseldorf	latrouite@phil.uni-duesseldorf.de
Li, Yu	University at Buffalo	yli57@buffalo.edu
Mitchell, Alice	University of Bristol	alice.mitchell@bristol.ac.uk
Möllemann, Robin	Heinrich-Heine-Universität Düsseldorf	moellemann@phil.hhu.de
Montero Melis, Guillermo	Max Planck Institute for Psycholinguistics	montero-melis@biling.su.se
Mora-Bustos, Armando	Universidad Autónoma Metropolitana	lucioamora@gmail.com
Murtagh, Irene	Technological University Dublin - Blanchardstown Campus; Trinity College Dublin	murtagh.irene@gmail.com
Ní Fhallamháin, Eibhlín	Trinity College Dublin	nifhalle@tcd.ie
Orqueda, Verónica	Universidad Católica	vorqueda@uc.cl
Osswald, Rainer	Heinrich-Heine-Universität Düsseldorf	osswald@phil.hhu.de
Park, Sang-Hee	University at Buffalo	sangheep@buffalo.edu
Prenner, Maria Katarzyna	University of Cologne	mprenner@uni-koeln.de
Rind-Pawlowski, Monika	Goethe-Universität Frankfurt am Main	monikapawlowski@live.de
Shimojo, Mitsuaki	University at Buffalo	shimojo@buffalo.edu
Stepanova, Anastasia	University at Buffalo	astepano@buffalo.edu
Toratani, Kiyoko	York University	ktora@yorku.ca
Van Valin Jr., Robert D.	University at Buffalo; Heinrich Heine University of Düsseldorf	vanvalin@phil.uni-duesseldorf.de
Wataru Nakamura	Tohoku University	wnmtohoku@gmail.com
Winther-Nielsen, Nicolai	Fjellhaug International University College Denmark; Vrije Universiteit Amsterdam	nwn@dbi.edu