

Proposal for a Master's thesis

Quantitative typology of valency classes

Siyu Tao

Draft as of November 24, 2022

Abstract

1 Introduction

Argument structure is widely studied and debated in linguistics. The nature of the relationship between verbs and their arguments as well as the syntactic expression of underlying lexical semantics.

2 Theoretical Framework

2.1 Valency and valency phenomena

In chemistry, *valency*, or *valence*, refers to the combining power of an atom or radical. The valency of any atom can be measured by the number of hydrogen atoms that it can combine with or displace in a chemical compound (Law and Rennie, 2020). This same term has been used in linguistics to similar effect and refers to the combining power of a word, primarily a verb or other predicate, with other words or elements of the sentence.

Lucien Tesnière is generally credited with introducing the term valency to linguistics with his syntactic theory of valency and dependence, as presented in the posthumously published *Éléments de syntaxe structurale* (1959; English translation 2015).¹ In another of Tesnière's metaphors, each verbal node, being the center of

¹It should be noted that while Tesnière is rightly credited with the introduction of a theory of linguistic valency, the metaphor of valency itself has made appearances as early as in Peirce (1897), among others (Przepiórkowski, 2018).

sentence structure, is not unlike a “theatrical performance” with the verb expressing the process and the nouns being the *actants* (what we would now call *arguments*) in this performance. Just like how atoms of different elements allow for a greater or lesser number of bonds, different verbs can combine with a greater or lesser number of actants, i.e., their valency.

While the term valency is borrowed into linguistics from chemistry, the study of the phenomena which are covered by or otherwise overlap with valency has a much longer tradition, dating to the early beginnings of linguistics from the *kāraka* concept of semantic relation between verb and noun (Ganeri, 2011) in Pāṇinian grammar to modern case grammar (Fillmore, 1968).

Most linguistic theories assert the centrality of the verb in determining either or both the syntactic and semantic structure of a sentence, corroborated also by psycholinguistic evidence (Healy and Miller, 1970). This places valency and the issues of *argument encoding* squarely at the center of the inquiry into the interface between lexical semantics and syntax.

In generative grammar, the syntactic valency of a verb is treated under a similar notion of *subcategorization* (Chomsky, 1965). As an example, a transitive verb must be followed by a direct object, whereas intransitive verb cannot, as such transitive and intransitive verbs form subcategories of the category verb. Verbs are therefore assigned to *subcategorization frames* which are considered part of the lexical entry of the verb, which specifies the number and type of complements (objects and obliques), as well as of the subject in later theories, that the verb can be subcategorized for. Note that the subcategorization here is primarily syntactically driven. Jackendoff (1972, 1987, 1992), following Katz and Fodor (1963) and Gruber (1962), further develops a theory of thematic relations and posits that argument structure serves as the interface between syntactic and thematic structures.

As compared to the broad distinctions such as those made between transitive and intransitive verbs, the verb classes in Levin (1993) provide a vastly more fine-grained categorization of verbs based on their syntactic behavior. Guided by the assumption that the syntactic behavior of verbs are determined semantically, Levin reasons that patterning together classes of verbs based on their diathesis alternations should result in semantically coherent verb classes. Levin’s work has been highly influential both in the development of valency theory and in computational approaches to lexical semantics. VerbNet Kipper et al., 2006, 2008; Kipper-Schuler, 2005 is a prominent example of projects, combining WordNet Fellbaum, 1998; Miller, 1995 with Levin-style verb classes.

A different line of research stems from Charles Fillmore’s frame semantics Fillmore, 1977a,b, 1982 as developed from his earlier case grammar (Fillmore, 1968, 1970) FrameNet (Fillmore and C. Baker, 2015)

further dev in construction grammar and corresponding approaches to construc-

tion Goldberg (1992, 1995)

CxG would consider valency frame as a level of construction. Whether this construction is autonomous will depend on whether the unpredictability condition is satisfied - in so far that the properties of valency frame cannot be predicted from other grammatical units. Croft (2012) takes a CxG approach.

2.2 Dependency grammars

de Marneffe and Nivre (2019)

2.3 Typological perspectives on valency and dependency

Typological interest in valency is primarily focused on cross-linguistic mismatches.

Already Tesnière (1959) describes the process of *metataxis*, by which syntactic structures of one language is translated to those of another.

Levin and Rappaport Hovav (2005) identifies five major questions that are necessary for a complete theory of argument realization.

One of the key question is also whether this valency frame are syntactically defined or semantically so. Since in the study of argument structure, we're dealing with the syntactic expression of lexical and non-lexical semantics, this question is particularly difficult to answer.

C. F. Baker and Lorenzi, 2020; Ellsworth et al., 2021 FrameNet and typology

Croft et al., 2017 Croft's typology and dependence grammar Tsunoda (1981, 1985) proposes a hierarchy of verbs

3 Data and methodology

3.1 Data sources

UD Treebanks Zeman et al., 2022

ValPaL (Hartmann et al., 2013) is a database

3.2 Methodology

Clustering non trivial problem to distinguish between complement vs adjuncts

Alignment

Information Theory

4 Work plan

5 Conclusion

Bibliography

- Baker, Collin F. and Arthur Lorenzi (May 2020). “Exploring Crosslinguistic Frame Alignment”. In: *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet*. Marseille, France: European Language Resources Association, pp. 77–84. ISBN: 979-10-95546-58-0. URL: <https://aclanthology.org/2020.framenet-1.11> (visited on 09/25/2022).
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.
- Croft, William (2012). *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press. ISBN: 978-0-19-924859-9.
- Croft, William, Dawn Nordquist, Katherine Looney, and Michael Regan (2017). “Linguistic Typology Meets Universal Dependencies”. In: *Proceedings of the 15th International Workshop on Treebanks and Linguistic Theories (TLT15), Bloomington, IN, USA, January 20-21, 2017*. Ed. by Markus Dickinson, Jan Hajic, Sandra Kübler, and Adam Przepiórkowski. Vol. 1779. CEUR Workshop Proceedings. CEUR-WS.org, pp. 63–75. URL: <http://ceur-ws.org/Vol-1779/05croft.pdf> (visited on 08/22/2022).
- De Marneffe, Marie-Catherine and Joakim Nivre (2019). “Dependency Grammar”. In: *Annual Review of Linguistics* 5.1, pp. 197–218. DOI: 10.1146/annurev-linguistics-011718-011842. eprint: <https://doi.org/10.1146/annurev-linguistics-011718-011842>. URL: <https://doi.org/10.1146/annurev-linguistics-011718-011842>.
- Ellsworth, Michael, Collin Baker, and Miriam R. L. Petruck (2021). “FrameNet and Typology”. In: *Proceedings of the Third Workshop on Computational Typology and Multilingual NLP*. Proceedings of the Third Workshop on Computational Typology and Multilingual NLP. Online: Association for Computational Linguistics, pp. 61–66. DOI: 10.18653/v1/2021.sigtyp-1.6. URL: <https://www.aclweb.org/anthology/2021.sigtyp-1.6> (visited on 11/01/2021).
- Fellbaum, Christiane, ed. (1998). *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. Cambridge, Mass: MIT Press. 423 pp. ISBN: 978-0-262-06197-1.
- Fillmore, Charles J. (1968). “The Case for Case”. In: *Universals in Linguistic Theory*. Ed. by Emmon Bach and Robert T. Harms. New York: Rinehart and Winston, pp. 21–119.
- (1970). “The Grammar of HITTING and BREAKING”. In: URL: https://www1.icsi.berkeley.edu/pubs/ai/ICSI_grammarofhitting12.pdf (visited on 09/18/2022).
- Fillmore, Charles J. (1977a). “Scenes-and-Frames Semantics”. In: *Linguistic Structure Processing*. Ed. by Antonio Zampolli. Fundamental Studies in Computer Science.

- Amsterdam: North Holland Publishing Company, pp. 55–82. ISBN: 978-0-444-85017-1.
- (Dec. 20, 1977b). “The Case for Case Reopened”. In: *Grammatical Relations*. Ed. by Peter Cole and Jerrold M. Sadock. BRILL, pp. 59–81. ISBN: 978-90-04-36886-6 978-90-04-36852-1. DOI: 10 . 1163 / 9789004368866 _ 005. URL: <https://brill.com/view/book/edcoll/9789004368866/BP000005.xml> (visited on 11/24/2022).
 - (1982). “Frame Semantics”. In: *Linguistics in the Morning Calm*. Ed. by Linguistic Society of Korea. Seoul, Korea: Hanshin Publishing Company, pp. 111–137.
- Fillmore, Charles J. and Collin Baker (Jan. 1, 2015). “A Frames Approach to Semantic Analysis”. In: *The Oxford Handbook of Linguistic Analysis*. Ed. by Bernd Heine and Heiko Narrog. Oxford University Press. ISBN: 978-0-19-967707-8. DOI: 10 . 1093 / oxfordhb / 9780199677078 . 013 . 0013. URL: <https://academic.oup.com/edited-volume/28050/chapter/211991006> (visited on 11/23/2022).
- Ganeri, Jonardon (July 21, 2011). “Kāraka: Meanings in Composition”. In: *Artha: Meaning*. Oxford University Press. ISBN: 978-0-19-807413-7. DOI: 10 . 1093 / acprof : oso / 9780198074137 . 001 . 0001. URL: <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198074137.001.0001/acprof-9780198074137> (visited on 11/22/2022).
- Goldberg, Adele E. (Jan. 1, 1992). “The Inherent Semantics of Argument Structure: The Case of the English Ditransitive Construction”. In: 3.1, pp. 37–74. ISSN: 1613-3641. DOI: 10 . 1515 / cogl . 1992 . 3 . 1 . 37. URL: <https://www.degruyter.com/document/doi/10.1515/cogl.1992.3.1.37/html> (visited on 11/24/2022).
- (Mar. 1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Cognitive Theory of Language and Culture Series. Chicago, IL: University of Chicago Press. 271 pp. ISBN: 978-0-226-30086-3. URL: <https://press.uchicago.edu/ucp/books/book/chicago/C/bo3683810.html> (visited on 11/23/2022).
- Gruber, Jeffrey S. (1962). “Studies in Lexical Relations”. PhD thesis. Massachusetts Institute of Technology.
- Hartmann, Iren, Martin Haspelmath, and Bradley Taylor (2013). *The Valency Patterns Leipzig Online Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: <https://valpal.info/>.
- Healy, Alice F. and George A. Miller (June 1, 1970). “The Verb as the Main Determinant of Sentence Meaning”. In: *Psychonomic Science* 20.6, pp. 372–372. ISSN: 2197-9952. DOI: 10 . 3758 / BF03335697. URL: <https://doi.org/10.3758/BF03335697> (visited on 11/17/2022).
- Jackendoff, Ray S. (1972). *Semantic Interpretation in Generative Grammar*. Studies in Linguistic Series. Cambridge, MA: The MIT Press. 400 pp. ISBN: 978-0-262-10013-

7. URL: <https://bac-lac.on.worldcat.org/oclc/323868> (visited on 11/22/2022).
- (1987). “The Status of Thematic Relations in Linguistic Theory”. In: *Linguistic Inquiry* 18.3, pp. 369–411. ISSN: 0024-3892. JSTOR: 4178548.
 - (Apr. 22, 1992). *Semantic Structures*. MIT Press. 340 pp. ISBN: 978-0-262-60020-0. Google Books: 7wbYlHis6OEC.
- Katz, Jerrold J. and Jerry A. Fodor (1963). “The Structure of a Semantic Theory”. In: *Language* 39.2, pp. 170–210. ISSN: 0097-8507. DOI: 10.2307/411200. JSTOR: 411200.
- Kipper, Karin, Anna Korhonen, Neville Ryant, and Martha Palmer (May 2006). “Extending VerbNet with Novel Verb Classes”. In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC’06)*. LREC 2006. Genoa, Italy: European Language Resources Association (ELRA). URL: http://www.lrec-conf.org/proceedings/lrec2006/pdf/468_pdf.pdf (visited on 11/23/2022).
- (Mar. 1, 2008). “A Large-Scale Classification of English Verbs”. In: *Language Resources and Evaluation* 42.1, pp. 21–40. ISSN: 1572-8412. DOI: 10.1007/s10579-007-9048-2. URL: <https://doi.org/10.1007/s10579-007-9048-2> (visited on 09/18/2019).
- Kipper-Schuler, Karin (2005). “VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon”. PhD thesis. USA: University of Pennsylvania.
- Valency (Valence) (Mar. 19, 2020). In: *A Dictionary of Chemistry*. Ed. by Jonathan Law and Richard Rennie. Oxford University Press. ISBN: 978-0-19-884122-7. URL: <https://www.oxfordreference.com/view/10.1093/acref/9780198841227.001.0001/acref-9780198841227> (visited on 11/22/2022).
- Levin, Beth (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, Ill.: Univ. of Chicago Press. ISBN: 978-0-226-47532-5 978-0-226-47533-2.
- Levin, Beth and Malka Rappaport Hovav (2005). *Argument Realization*. Research Surveys in Linguistics. Cambridge: Cambridge University Press. ISBN: 978-0-521-66331-1. DOI: 10.1017/CBO9780511610479. URL: <https://www.cambridge.org/core/books/argument-realization/8B1F52537D0D9EF4EB63B6AE30DDB6DE> (visited on 11/08/2022).
- Miller, George A. (Nov. 1, 1995). “WordNet: A Lexical Database for English”. In: *Communications of the ACM* 38.11, pp. 39–41. ISSN: 0001-0782. DOI: 10.1145/219717.219748. URL: <https://doi.org/10.1145/219717.219748> (visited on 11/23/2022).
- Peirce, Charles S. (1897). “The Logic of Relatives”. In: *The Monist* 7.2, pp. 161–217. ISSN: 0026-9662. URL: <https://doi.org/10.5840/monist18977231> (visited on 09/18/2022).

- Przepiórkowski, Adam (Jan. 1, 2018). "The Origin of the Valency Metaphor in Linguistics". In: *Lingvisticæ Investigationes* 41.1, pp. 152–159. ISSN: 0378-4169, 1569-9927. DOI: 10.1075/li.00017.prz. URL: <https://www.jbe-platform.com/content/journals/10.1075/li.00017.prz> (visited on 09/18/2022).
- Tesnière, Lucien (1959). *Éléments de syntaxe structurale*. Paris: C. Klincksieck.
- (2015). *Elements of Structural Syntax*. John Benjamins Publishing Company. ISBN: 978-90-272-1212-2. DOI: 10.1075/z.185. URL: <https://library.oapen.org/handle/20.500.12657/30722> (visited on 08/22/2022).
- Tsunoda, Tasaku (Jan. 1, 1981). "Split Case-Marking Patterns in Verb-Types and Tense/Aspect/Mood". In: 19.5-6, pp. 389–438. ISSN: 1613-396X. DOI: 10.1515/ling.1981.19.5-6.389. URL: <https://www.degruyter.com/document/doi/10.1515/ling.1981.19.5-6.389/html?lang=en> (visited on 09/25/2022).
- (1985). "Remarks on Transitivity". In: *Journal of Linguistics* 21.2, pp. 385–396. ISSN: 0022-2267. JSTOR: 4175793.
- Zeman, Daniel et al. (2022). *Universal Dependencies 2.10*. URL: <http://hdl.handle.net/11234/1-4758>.