

Linking across the syntax-lexicon *continuum*: horizontal links between denominal psych-verbs and light verbs constructions in Italian

Flavio Pesciotta

Even though the syntax-lexicon *continuum* represents one of the core assumptions of CxG, research has rarely addressed the topic of the relationship between semantically similar morphological and syntactic constructions (Masini 2019). The investigation of such competing constructions and how they can be possibly linked together can be fruitful to expand our knowledge regarding the *continuum*, and to better understand the nature of constructional links in light of the different behaviour of lexical and syntactic items.

We explore this topic by studying the relationship between Light Verb Constructions (LVCs) (e.g., Ježek 2004; Pompei 2017) and synthetic verbs (SVs) in Italian. As a matter of fact, LVCs (1b) often have a semantically similar SV counterpart (1a) in Italian. This violation of the principle of isomorphism generates a competition between two different strategies of predicate formation from nouns.

- 1) **coraggio** ‘courage’
 - a. **incoraggiare**
 - b. **dare coraggio** (lit. ‘to give courage’) ‘to encourage’

We aim to analyze this relationship by focusing on SVs and LVCs derived from psych nouns, in order to understand how it can be modeled in the constructional network. Since it would be impossible to posit any common “mother” construction, their relationship constitutes a privileged viewpoint to study the nature of horizontal links, both paradigmatic and allostructional (cf. Cappelle 2006; Van de Velde 2014; Jackendoff & Audring 2020).

We extracted 87 non-derived psych nouns from the ItEm database (Passaro et al. 2015). Successively, we retrieved the derived SVs from the GRADIT dictionary (2007) and the stative, inchoative and causative LVCs created from such nouns from the itWaC corpus (Baroni et al. 2009).

The composition of our database shows that the probability of creating SVs and LVCs increases with the frequency of the nouns; however, we do not find more than one SV per noun, while this is not the case for LVCs. This results in a constructional division of labour between LVCs and SVs: for instance, the presence of LVCs often allows to express the aspect-causative types that are not lexicalized by the SV derived from the corresponding noun. Nonetheless, LVCs do not always occupy “empty” semantic niches: we found 37 cases in which an SV and one or more LVCs express the same meaning.

To analyze which variables affect the choice between these semantically overlapping constructions, we selected 10 nouns from the database, balancing the list according to frequency factors and to the composition of the database, and we extracted a sample of 500 occurrences from corpora of written and spoken Italian. 190 of them were collected from all the available corpora of spoken Italian: LIP (De Mauro et al. 1993), KIParla (Mauri et al. 2019), and Radiocast-IT¹ (Masini & Combei, in press). The remaining 310 occurrences were extracted as a random sample from the written Italian corpus CORIS (Rossini Favretti et al. 2002). We will code them by a number of parameters, including: noun type; text genre; argument structure features; adjectival/adverbial modification. Mixed-effects regression will be used to better understand the impact of these factors.

¹ The corpus is still under construction at the University of Bologna (link to the project: <https://site.unibo.it/radiocast/it>).

We expect LVCs to be favoured by the presence of animate subjects and by the possibility of noun modification (Bonial & Pollard 2020; Wiskandt & Turus 2022), and more generally, to find some form of division of labour also in the case of constructions apparently occupying the same semantic niche.

References

- Bonial, Claire Nicole & Kimberly A. Pollard. 2020. Choosing an event description: What a PropBank study reveals about the contrast between light verb constructions and counterpart synthetic verbs. *Journal of Linguistics* 56(3). 577–600.
- Cappelle, Bert. 2006. Particle placement and the case for “allostructions”. *Constructions* SV 1(7).
- Jackendoff, Ray & Jenny Audring. 2020. Relational Morphology: A Cousin of Construction Grammar. *Frontiers in Psychology* 11. 1–12.
- Ježek, Elisabetta. 2004. Types et degrés de verbes supports en italien. *Lingvisticae Investigationes* 27(2). 185–201.
- Masini, Francesca. 2019. Competition Between Morphological Words and Multiword Expressions. In Franz Rainer, Francesco Gardani, Wolfgang U. Dressler & Hans Christian Luschützky (eds.), *Competition in Inflection and Word-Formation*, 281–305. Cham: Springer.
- Pompei, Anna. 2017. Verbi complessi con struttura [V + SP] e verbi supporto. Proprietà e test. *Studi e Saggi Linguistici* 55(2). 109–136.
- Van de Velde, Freek. 2014. Degeneracy: The maintenance of constructional networks. In Ronny Boogaart, Timothy Coleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*, 141–180. Berlin/Boston: De Gruyter.
- Wiskandt, Niklas & Dila Turus. 2022. Systematic semantic differences between object-experiencer LVCs and corresponding simplex verbs in German. Presented at the *9th International Conference on Grammar and Corpora 2022*, 30 June - 2 July 2022, Ghent, Belgium.

Corpora and lexicographic resources

- CORIS = Rossini Favretti, Rema, Fabio Tamburini & Cristiana De Santis. 2002. CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model. In Andrew Wilson, Paul Rayson & Tony McEnery (eds.), *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*, 27–38. München: Lincom-Europa.
- GRADIT = *Grande dizionario italiano dell'uso*, directed by Tullio De Mauro. 2007. Turin: UTET.
- ItEm = Passaro, Lucia C., Laura Pollacci & Alessandro Lenci. 2015. ItEM: A Vector Space Model to Bootstrap an Italian Emotive Lexicon. In Cristina Bosco, Sara Tonelli & Fabio Massimo Zanzotto (eds.), *Proceedings of the Second Italian Conference on Computational Linguistics CLiC-it 2015*, Trento, 3-4 December 2015, 215–220. Torino: Accademia University Press.
- itWaC = Baroni, Marco, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky Wide Web: A collection of very large linguistically processed Web-crawled corpora. *Language Resources and Evaluation* 43 (3). 209–226.
- LIP = De Mauro, Tullio, Federico Mancini, Massimo Vedovelli & Miriam Voghera. 1993. *Lessico di frequenza dell'italiano parlato*. Milano: Etaslibri.
- KIParla = Mauri, Caterina, Silvia Ballarè, Eugenio Gorla, Massimo Cerruti & Francesco Suriano. 2019. KIParla corpus: a new resource for spoken Italian. In Raffaella Bernardi, Roberto Navigli & Giovanni Semeraro (eds.), *Proceedings of the Sixth Italian Conference on Computational Linguistics (CLIC-it 2019)*. CEUR-WS, vol. 2481.
- Radiocast = Masini, Francesca & Claudia Roberta Combei. In press. RadioCast-It. Corpus di italiano della radio e dei podcast. In Silvia Ballarè, Ilaria Fiorentini & Emanuele Miola (eds.), *Le varietà dell'italiano contemporaneo*, 196–198. Roma: Carocci.