

A distributional approach to usage-based exemplar representation of polysemous and productive derivational meaning: the case of Italian -ATA suffixation

Chiara Paolini – KU Leuven (PhD student – corresponding author)

Alessandro Lenci – University of Pisa

Denis Paperno – Utrecht Universiteit

This study is interested in the Italian denominal suffix -ATA, whose derivational meaning is characterized by an extremely diverse and fragmented (Grossmann & Rainer, 2004), yet productive polysemy. Derived from the female past participle, the suffix maintains its most prototypical actional meaning when attached to verbal bases, describing different realizations of an action, such as a prolonged activity (1a) or a single act (1b). However, when attached to nominal and adjectival bases, the semantic complexity and diversity of -ATA drastically increase. For example, expression of quantity and events (1c, where two meanings overlap), or typicality of behaviour or action (1d), among others.

- (1) a. *passeggiata* (stroll): the act of strolling
- b. *martellata* (hammer blow): the single act of hammering / a stroke with a hammer
- c. *spaghetтата* (spaghetti dinner): a group of people sharing a spaghetti meal / a big quantity of spaghetti
- d. *berlusconata*: a typical behaviour or affirmation, usually funny or ridiculous, of the former Italian PM Silvio Berlusconi

It has been a main challenge for morphologists (Gatti & Togni 1991; Gaeta 2002; Acquaviva 2005; Fiorentini 2010) to provide a comprehensive and exhaustive characterization of the heterogeneous semantic production of -ATA. The traditional solution has consisted of adopting highly general semantic labels such as “the single act of V”, “the typical behaviour of N”, or “a quantity of N”. These paraphrases, however, do not cover all different derived meanings in use, especially the ones appearing in nonce formations (Bauer, 1983) and neologisms. Furthermore, there is a high degree of disagreement about the semantic structure of -ATA, indicative of the limitations imposed by traditional high-level semantic features, with regard to semantic plasticity and predictability of the usage of this suffix. Therefore, new formations tend to be underrepresented and underestimated within the traditional paradigms, even if they are widely produced, recognized and used by speakers. For this reason, we aim to investigate the structured polysemy of -ATA starting from the local similarities of meaning to conventionalized morphological constructions (Bybee 2010).

This study proposes a usage-based exemplar representation of the polysemous derivational meaning (Bybee and Eddington 2006), with analogy accounting for productive use, within the framework of Construction Morphology (Booij 2010). We adopted a two-pronged approach: (1) the organization of 410 -ATA formations elicited from the itTenTen16 corpus (Jakubíček et al 2013) into groupings of semantically similar nouns related to a high-frequency exemplar (2) an empirical distributional semantic analysis (Lenci 2018). With this analysis we first evaluate the internal and external consistency and goodness of exemplar clusters (Bonami & Paperno 2018), then visualize the findings with the t-SNE algorithm (van der Maaten & Hinton 2008).

Our results show that the derivational polysemy of -ATA is indeed solidly structured rather than a loose grouping of exceptions to the rule. Overall, this study provides further support for including exemplar representations in the study of lexeme formation processes, as the inclusion of analogy supplies new and exciting clues for our understanding of semantic fuzziness in derivation morphology.

References

- Acquaviva, P. (2005). I significati delle nominalizzazioni in-ATA e i loro correlati morfologici. In: Thornton, Anna Maria and Grossmann, Maria (eds.). *La formazione delle parole: atti del XXXVII convegno della Società Linguistica Italiana*. Roma: Bulzoni Editore.
- Bauer, L. (1983). *English word-formation*. Cambridge: Cambridge University Press.
- Bonami, O., & Paperno, D. (2018). Inflection vs. derivation in a distributional vector space. *Lingue e linguaggio*, 17(2), 173-196.
- Booij, G. (2010). Construction Morphology. In: *Language and Linguistics Compass*, 4: 543-555. doi:10.1111/j.1749-818X.2010.00213.x
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Bybee, J. & Eddington, D. (2006). A usage-based approach to Spanish verbs of 'becoming'. *Language*, 323-355.
- Fiorentini, I. (2010). *Pedalata, Carrambata, Spaghetтата. Funzioni e produttività del suffisso –ATA nell'italiano d'oggi*. MA thesis.
- Gaeta, L. (2002). *Quando i verbi compaiono come nomi: Un saggio di Morfologia Naturale*. Milano: Franco Angeli.
- Gatti, T. & Togni, L. (1991). A proposito dell'interpretazione dei derivati in-ata e in s-. In: *Arbeitspapier Nr. 30, Fachgruppe Sprachwissenschaft*. Konstanz: University of Konstanz.
- Grossmann, M. & Rainer, F. (2004). *La formazione delle parole in italiano*. Berlin, Boston: Max Niemeyer Verlag. <https://doi.org/10.1515/9783110934410>
- Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen corpus family. In: *7th International Corpus Linguistics Conference – CL*, 125-127.
- Lenci, A. (2018). Distributional models of word meaning. *Annual review of Linguistics*, 4, 151-171.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).