

Morphological complexity in semantic access: an investigation of novel complex word processing

Ludovica Monte¹, Marco Marelli¹

¹ Department of Psychology, University of Milano-Bicocca, Italy

Background: Every time we come across a new word, we naturally tend to infer its meaning [1, 2]. What happens when the new word we are exposed to is morphologically complex? Are a combination of a stem and a suffix processed as a single unit conveying a composed meaning? Distributional semantics helps us addressing this issue by providing quantitative predictions about word meanings, even unattested ones [3, 4]. To understand whether our mind generates meanings according to automatic combinatorial processes, this study examines novel morphologically complex nonwords and test related predictions from a compositional distributional semantic model in a semantic priming paradigm [5].

Method: In the CAOSS model (Compounding as Abstract Operation in Semantic Space) [6], as adapted to derivation [5], the semantics of stems and affixes are characterized as distinct semantic vectors that can be combined to form new meanings. We applied this model to obtain semantic representations for novel derived words. These were generated by moving from a set of stems belonging to three grammatical categories: nouns, verbs, and adjectives. The frequency of the stems was verified using the *CoLFIS* corpus and the *SubTlex-IT* database. Only low-frequency words were selected - consistently to observations highlighted by previous studies [7–9] according to which the decision latencies of a lexeme are facilitated by low-frequency stems. Based on such existing stems, 170 pseudowords were generated as stimuli, characterized by a legal combination (therefore responding to grammatical rules in Italian but not existing in that language) of a stem and a suffix, e.g. **eludimento*, eng. **eludement*. Each of these new pseudowords, which acts as a prime word in our experimental condition, was represented via CAOSS as a vector in a semantic space where other existing words were also present, linked to it at different levels of semantic proximity. Within such neighbourhoods, we selected the closest existing word of each novel pseudowords (e.g. *sistema*, eng. *system*) and designated it as the target word, which was kept constant across the experiment. Indeed, in two control conditions, the same target was preceded by either the original stem (e.g. *eludere*, eng. *to elude*) or an unrelated word (e.g. *chitarra*, eng. *guitar*), following the condition that the vector distance between the target and the pseudoword prime had to be greater than the distance between the target and the stem prime.

The experiment, conducted on Italian-speaking adult participants, adopts an unmasked semantic priming paradigm. After a fixation cross, the prime (e.g. *eludimento* vs *eludere* vs *chitarra*) is presented with a SOA of 400 ms, followed by the target word (e.g. *sistema*). Participants are asked to evaluate whether the target is an existing Italian word or not. RTs and accuracy are collected as dependent variables.

Results and Discussion: The study is currently in progress. The expectations are as follows: since the semantic proximity between the pseudoword and the target word is higher than the proximity between the target and the control primes, a more pronounced priming effect is expected, with a significant reduction in reaction times compared to the control conditions. If *eludimento* is more effective than *eludere* in evoking *sistema*, it can be inferred that the combined meaning is more easily recoverable than its stem, even if the word morpho-logically complex and completely unfamiliar, suggesting a routine and automatic computation of new meanings in presence of novel complex forms [10].

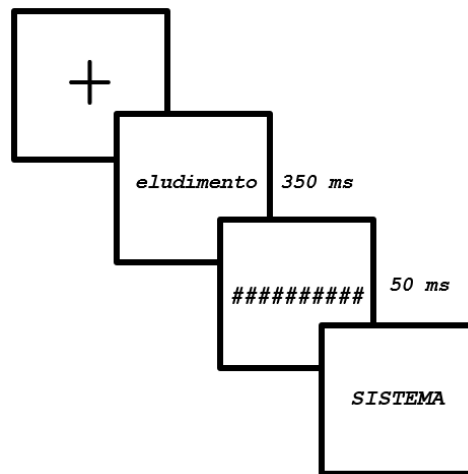


Figure 1: Graphical illustration of the experimental procedure.

	Prime	Target
Pseudoword / Target word	eludimento (eng. eludement)	SISTEMA (eng. system)
Stem / Target word	eludere (eng. to elude)	SISTEMA (eng. system)
Unrelated word / Target word	chitarra (eng. guitar)	SISTEMA (eng. system)

Table 1: Examples of the Experimental Conditions.

References

- [1] Wang, S., Roller, S., & Erk, K. (2017). Distributional modeling on a diet: One-shot word learning from text only.
- [2] Paller, K., Kutas, M., & Mayes, A. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology*, 67(4), 360–371. [https://doi.org/10.1016/0013-4694\(87\)90124-6](https://doi.org/10.1016/0013-4694(87)90124-6)
- [3] Amenta, S., Crepaldi, D., & Marelli, M. (2020). Consistency measures individuate dissociating semantic modulations in priming paradigms: A new look on semantics in the processing of (complex) words. *Quarterly Journal of Experimental Psychology*, 73(10), 1546–1563. <https://doi.org/10.1177/1747021820927663>
- [4] Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review*, 122(3), 485–515. <https://doi.org/10.1037/a00%2039267>
- [5] Bonandrini, R., Amenta, S., Sulpizio, S., Tettamanti, M., Mazzucchelli, A., & Marelli, M. (2023). Form to meaning mapping and the impact of explicit morpheme combination in novel word processing. <https://doi.org/10.1016/j.cogpsych.2023.101594>
- [6] Marelli, M., Gagné, C., & Spalding, T. (2017). Compounding as Abstract Operation in Semantic Space: Investigating relational effects through a large-scale, datadriven computational model. <https://doi.org/10.1016/j.cognition.2017.05.026>
- [7] Amenta, S., Günther, F., & Marelli, M. (2020). A (distributional) semantic perspective on the processing of morphologically complex words. *The Mental Lexicon*, 15(1), 62–78. <https://doi.org/10.1075/ml.00014.ame>
- [8] Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, 3(232), 232. <https://doi.org/10.3389/fpsyg.2012.00232>
- [9] Burani, C., Dovetto, F., Spuntarelli, A., & Thornton, A. (1999). Morpholexical access and naming: The semantic interpretability of new root-suffix combinations. *Brain and Language*, 68(1-2), 333–339. <https://doi.org/10.1006/brln.1999.2073>
- [10] Günther, F., & Marelli, M. (2020). Trying to make it work: Compositional effects in the processing of compound “non-words”. *Quarterly Journal of Experimental Psychology*, 73(7), 1082–1091. <https://doi.org/10.1177/1747021820902019>