Computational modelling of polysemy sense relatedness

Ksenija Mišić[1,2], Dušica Filipović Đurđević[1,2,3]

[1] *Laboratory for Experimental Psychology, University of Belgrade – Faculty of Philosophy, Serbia*
[2] *Department of Psychology, University of Belgrade – Faculty of Philosophy, Serbia*
[3] *Laboratory for Experimental Psychology, University of Novi Sad – Faculty of Philosophy, Serbia*

Background: Polysemy is a semantic phenomenon where words can have multiple related senses (e.g., paper as writing material or a scientific article) [1]. Previous studies investigating whether the polysemy effect (more senses lead to faster processing) can be simulated within the error-driven learning (EDL) framework showed that, in addition to cue competition, strong outcome competition is necessary for learning multiple related senses [2]. The outcome competition was achieved by adding context words at the outcome level. The assumption was that more similar senses share more context words, which boosts learning at the outcome level [3]. However, this was demonstrated within naïve discriminative learning approach. This model treats outcomes independently of each other, limiting the ability to conclude whether the sense relatedness was indeed captured.

In this study, we aimed to test more explicitly whether EDL can capture the relatedness between senses. We used linear discriminative learning model that allows for continuous instead of discrete outcomes, that take outcome relatedness into account. By varying the role of the context (cues, outcomes, or both) we explored whether sense relatedness can arise from learning. We employed three types of networks: Gc2V, with polysemous and context words as orthographic cues and static embeddings as outcomes, focusing on cue competition; G2Vc, with polysemous words as cues and contextualised embeddings as outcomes, focusing on outcome competition; and Gc2Vc, in which context was present at both the cue and outcome levels.

Method: We took polysemous nouns, verbs, and adjectives from an existing database [] and extracted occurrences of each word and immediate context words from SrWaC [4] corpus. Due to large computational demands of calculating contextualised embeddings (BERT [5]) we took a subsample of occurrences of each word preserving the differences in frequency between words. For static embeddings we used fastText embeddings [6]. Cues consisted of orthographic trigraphs derived either from the polysemous word alone or from the polysemous word together with its context. For each part of speech and each network type, we calculated average cosine similarity between predicted outcomes for each word. Such measures were also calculated for contextualised embeddings, as they represented a reference point.

Results: G2Vc network failed to capture any variation within polysemous words. For the other two networks, we calculated similarity measures for predicted outcomes and found no significant correlation with the average similarity of contextualised embeddings. To test whether similarity measures of contextual embeddings capture sense relatedness effect, we modelled lexical decision latencies and found a significant facilitatory effect on processing.

Discussion: Heavy reliance on outcome competition (G2Vc) did not produce sense-relatedness effects, nor did the cue-competition model (Gc2V). Introducing context at both cue and outcome level also failed to produce any relatedness effects. The inability of sense relatedness effect to arise from the EDL, may not be due to the network setup, but rather due to using a subsample of contexts that polysemous words appeared in. This led to the reduction of variation in the context and may have affected the structure of the senses as they appear in the contexts. Additionally, cosine similarity may have been affected by high collinearity among contextualised embeddings and predicted outcomes, or it may simply not be sufficiently sensitive.

References

[1] Rodd, J. (2020). Settling Into Semantic Space: An Ambiguity-Focused Account of Word-Meaning Access. *Perspectives on Psychological Science*, *15*(2), 411–427. https://doi.org/10.1177/1745691619885860
[2] Mišić, K., & Đurđević, D. F. (2024). Can polysemy effects arise from cue competition-only error-driven learning?
[3] Đurđević, D. F., & Kostić, A. (2023). We probably sense sense probabilities. *Language, Cognition and Neuroscience*, *38*(4), 471–498. https://doi.org/10.1080/23273798.2021.1909083
[4] Ljubešić, N., Klubička, F., Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2016). *{Bs,hr,sr}wac -web corpora of bosnian, croatian and serbian*. Association for Computational Linguistics. https://doi.org/10.48550/arXiv.1810.04805
[5] Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information.