A Cute Puyfred and a Scary Boppies: How do People and LLMs Interpret Novel Words in Context?

Giovanni Cassani (Tilburg University)

Studies on form-meaning mappings in the lexicon have highlighted many systematic relations between the surface form of words and what they mean. Recent studies further blurred the boundary between words and pseudowords by documenting similar implicit and explicit semantic effects for both. From a language learning perspective, such a blurred boundary makes sense: every word a person knows has been a pseudoword during development, and many valid words are pseudowords to many speakers, who might encounter them and have to quickly shape a semantic representation.

In this talk, I will present ongoing work examining the interplay between semantic connotations conveyed by the word form itself and by the sentence context in which the word form is first introduced, combining behavioral evidence with computational experiments involving Large Language Models. The key manipulation in these experiments concerns the congruence between the valence of the sentence context and the perceived valence of a novel word.

Preliminary findings from a free association task and a self-paced reading task confirm that participants are sensitive to the semantic connotations of isolated pseudowords, previously rated by a different pool of participants: positive pseudowords elicited significantly more positive associates than negative pseudowords. However, even a single sentence context obliterates such connotations in favor of the semantic conveyed by the sentence context: even if *puyfred* elicited negative associates in isolation, a *cute puyfred* elicits positive associates, much like a *cute boppies*. Moreover, the perceived valence of a pseudoword and its congruence with the valence of the sentence context do not influence reading times: participants find it equally hard to integrate *puyfred* and *boppies* in a positive sentence context. Some differences across models notwithstanding, LLMs exhibit a similar pattern.

While preliminary, these findings suggest that systematic form-meaning mappings present in the lexicon do provide a source of information when other linguistic cues are absent, but that speakers learn to disregard such cues when distributional information, albeit quantitatively scarce, is available.