Operationalising Contextualised Sensorimotor Profiles

Sasha Kenjeeva, Giovanni Cassani, Noortje J. Venhuizen, Afra Alishahi
Research Center for Cognitive Science and Artificial Intelligence, Tilburg University

Extensive evidence suggests that word meaning is partially represented through sensorimotor information, which refers to sensory (e.g., vision) and motor (e.g., foot/leg) interactions with the word's associated referent or action [1]. In everyday linguistic communication, these sensorimotor representations are not static, but modulated by the context in which the word occurs [2]: for example, "cone" in "It was a sugar cone" is more strongly associated to the sensorimotor dimension of taste, compared to "cone" in "It was a traffic cone" [3]. Evidence for this comes from datasets reporting context-dependent sensorimotor ratings, whereby participants rated polysemous words across varying contexts based on how associated they are to a set of sensorimotor dimensions (e.g., taste, smell, vision, hand, head) [3, 4]. While these datasets are limited in scope and sentence complexity (see Table 1 for examples), we here aim to explore how they can be harnessed to derive contextualised sensorimotor profiles (CSPs) for more naturalistic language data for which contextualised ratings are not available. This will allow us to systematically investigate how context affects sensorimotor meaning. For instance, does the CSP of "cone" in "It was a sugar cone" simply inherit the sensorimotor profile of "sugar", or is it better explained by some composition of the sensorimotor profiles of "sugar" and "cone"? Additionally, do different types of contexts – depending on whether they include a modifier or a predicate – modulate CSPs differently? And are all sensorimotor dimensions equally important in modulating a context's representation?

We propose to operationalise CSPs by starting from these questions and drawing on existing theories and resources. Specifically, we will make use of vectors derived from uncontextualised sensorimotor ratings [5] to test how they account for the contextualised ratings from two datasets [3, 4], which can also be vectorised in the same manner. This will allow us to explore simple vector arithmetics to investigate the combinatorial aspect of sensorimotor information, relying on previous work on computational methods of compositionality [6, 7]: a visualisation of this idea can be found in Figure 1. Additionally, we plan to extract contextualised embeddings from visual language models (VLMs), which have been found to partially approximate sensorimotor grounding in humans [8] and can be used to augment human norms [9]. The differences between the two datasets will allow us to investigate how context type affects CSPs, including differentiation between various sensorimotor dimensions: for example, a modifier such as "sugar" might exert stronger effects on "cone" on specific dimensions (e.g., taste and mouth) than the predicate "watch" on "traffic", which might more closely resemble the uncontextualised profile of "traffic". This will integrate sensorimotor information into language-based theories of contextual sense selection [10], an effort consistent with recent work [11].

While prior work has shown that contextual sensorimotor information influences word processing [2], it remains unclear how changes in sensorimotor signatures can be quantified. We therefore aim to explore different operationalisations of contextualised sensorimotor profiles (CSPs) to investigate how sensorimotor meaning is composed in context. Having a tool to obtain CSP beyond short, manually-crafted stimuli will benefit investigations into the role of sensorimotor information for more naturalistic data and tasks, such as reading, and pave the way to systematically examine how we construct meaning in language considering a further piece of the puzzle.

Dataset	Type	Sentence	Noun	Context
Trott [3]	Modifier	*It was a sugar cone.*	*cone*	*sugar*
		It was a traffic cone.	*cone*	*traffic*
Bruera [4]	Predicate	*Watch the traffic.*	*traffic*	*watch*
		Manage the traffic.	*traffic*	*manage*

Table 1: Sentence examples from the human-rated datasets.

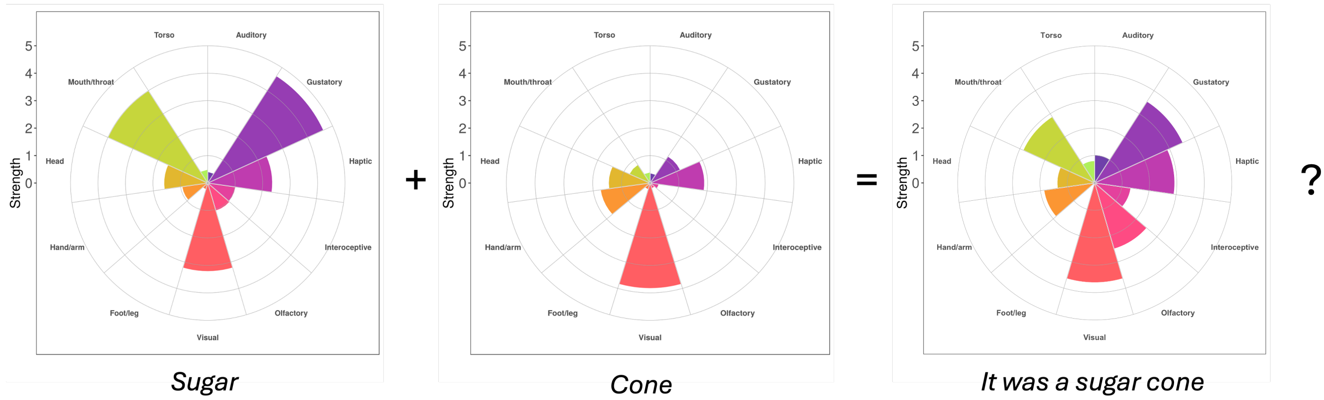

Figure 1: Vector arithmetics hypothesis visualised (using source code from [5]): does the sum of uncontextualised sensorimotor vectors approximate the contextualised sensorimotor profile well?

References

[1] Körner, A., Castillo, M., Drijvers, L., Fischer, M. H., Günther, F., Marelli, M., Platonova, O., Rinaldi, L., Shaki, S., Trujillo, J. P., Tsaregorodtseva, O., & Glenberg, A. M. (2023). Embodied Processing at Six Linguistic Granularity Levels: A Consensus Paper. *Journal of Cognition*, *6*(1), 60. https://doi.org/10.5334/joc.231

[2] Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, *23*(4), 1015–1027. https://doi.org/10.3758/s13423-015-0948-7

[3] Trott, S., & Bergen, B. (2022). Contextualized Sensorimotor Norms: Multi-dimensional measures of sensorimotor strength for ambiguous English words, in context. http://arxiv.org/abs/2203.05648

[4] Bruera, A., Zamani, F., & Poesio, M. (2024). Polysemy through the lens of psycholinguistic variables: A dataset and an evaluation of static and contextualized language models. In D. Bollegala & V. Shwartz (Eds.), *Proceedings of the 13th Joint Conference on Lexical and Computational Semantics (*SEM 2024)* (pp. 35–48). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.starsem-1.3

[5] Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*(3), 1271–1291. https://doi.org/10.3758/s13428-019-01316-z

[6] Marelli, M., Gagné, C. L., & Spalding, T. L. (2017). Compounding as Abstract Operation in Semantic Space: Investigating relational effects through a large-scale, data-driven computational model. *Cognition*, *166*, 207–224. https://doi.org/10.1016/j.cognition.2017.05.026

[7] Mitchell, J., & Lapata, M. (2010). Composition in Distributional Models of Semantics. *Cognitive Science*, *34*(8), 1388–1429. https://doi.org/10.1111/j.1551-6709.2010.01106.x

[8] Jones, C., Bergen, B., & Trott, S. (2024). Do Multimodal Large Language Models and Humans Ground Language Similarly? *Computational Linguistics*, 1–25. https://doi.org/10.1162/coli_a_00531

[9] Trott, S. (2024). Can large language models help augment English psycholinguistic datasets? *Behavior Research Methods*, *56*(6), 6082–6100. https://doi.org/10.3758/s13428-024-02337-z

[10] Klepousniotou, E. (2002). The Processing of Lexical Ambiguity: Homonymy and Polysemy in the Mental Lexicon. *Brain and Language*, *81*(1), 205–223. https://doi.org/10.1006/brln.2001.2518

[11] Božić, S., & Đurđević, D. (2025). The neglected role of lexical ambiguity in embodied cognition models. https://doi.org/10.31234/osf.io/g483a_v1