

If you're about distributional semantics, you'll be into this talk: semantic change in the recent history of *into* and *about*

Lauren Fonteyn¹

¹Leiden University Centre for Linguistics

l.fonteyn@hum.leidenuniv.nl

Much like the majority of other English prepositions, *into* and *about* can express a wide array of meanings, including spatial (1), temporal (2), and more abstract senses (3)-(4):

- | | | |
|-----|--|---|
| (1) | Where's all the visitors gone? - Oh, they are about the house somewhere (1885, COHA) | He was into her bed in a flash, stocking and all. (1940, COHA) |
| (2) | The moon will be up about midnight (1911, COHA) | By the time they are into their forties they have enough seniority to chair important committees. (1963, COHA) |
| (3) | Most of the book is about her father (1955, COHA) | She's into erotic movies for the money. What she really is is a musician. (1980, COHA) |
| (4) | I got between her and Dad and demanded a hug. Yes, I was all about the hugs. (2008, COHA) | She'd be up for it. She's into good-looking men. (1995, COHA) |

In Cognitive Semantics, the meanings of prepositions are typically represented as networks of conceptually (i.e. metaphorically, metonymically) linked senses (Cuyckens 1999: 15). However, there is no consensus on the most plausible shape of these representations: the proposed senses in the networks and the distances/connections between them are often derived from subjective intuitions, which are ultimately not falsifiable (Glynn 2014: 17). To attain more credible approximations of such polysemy networks, then, it is key to approach the concept of semantic relatedness in a objective, data-driven way (Gries & Divjak 2005; Sagi et al. 2011). Supported by the sizable Corpus of Historical American English (COHA, 400M words), this study sets out to capture the most plausible diachronic trajectory of the 'fondness' sense attested with *about* and *into* (illustrated in (4)) by means of a state-of-the-art, token-based distributional semantic model called BERT (Devlin et al. 2018). This model creates compressed usage representations of individual contextualized tokens of *into* and *about* (ca. 55,000 tokens) in the form of vectors. By means of these token vectors, we can induce a large-scale and data-driven estimation of the distances or 'semantic relatedness' (cf. Rice 1996) between all senses of *into* and *about* over time.

References: Devlin et al. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805; Cuyckens, H. 1999. Historical evidence in prepositional semantics: The case of English *by*. In Tops et al. (eds.), 15-32. Leuven: Peeters; Glynn, D. 2014. Polysemy and synonymy: Cognitive theory and corpus method. In Glynn, D. & J. Robinson (eds.), 7-38. Amsterdam: Benjamins; Gries, S. Th. & D. Divjak. 2005. Behavioral profiles: A corpus-based approach to cognitive semantic analysis. In Evans, V. & S. Pourcel (eds.), 57-75. Amsterdam: Benjamins; Rice, S. 1996. Prepositional prototypes. In Pütz, M. & R. Dirven (eds.), The construal of space in language and thought, 135-166. Berlin: De Gruyter; Sagi, E., Kaufmann, S. & B. Clark. 2011. Tracing semantic change with Latent Semantic Analysis. In K. Allan & J. Robinson (eds.), 161-183. Berlin: De Gruyter.