Review of: "A Corpus Analysis of Polysemy in CEFR-based English Textbooks"

Barbara Malt[1]

1 Lehigh University

Potential competing interests: No potential competing interests to declare.

The author makes a good point, that learners of English (or other languages) need to learn not just prototypical senses of words but their other senses. And learners need to be aware of the polysemous nature of many common words. It is interesting to see that most commonly used words into English language textbooks are used with only one or two senses.

That said, the advance provided by this paper is fairly small. I didn't understand the first research goal of exploring "the high usage of the most frequently used words". I don't understand what this means, or what it could mean. High frequency words are naturally frequently used. What else about them do we need to find out? This goal needs to be articulated more clearly and more specifically.

As for the result related to this goal, that 65% of the most frequent words in the two textbooks overlap, that seems to be just what is expected. High frequency words in a language are useful in many contexts and sentences. They are anchors around which new vocabulary words can be introduced. Why wouldn't high frequency words tend to overlap in two textbooks? Is there something wrong with that?

The second goal of exploring the extent of polysemy in the two textbooks is more interesting, and the result is more revealing. However, it's not clear what the implication of the result is. Obviously, learners cannot learn polysemous meanings that aren't represented to them. But are they at learning stages in these two textbooks where they are ready to learn multiple meanings of these words? Or is it more beneficial for them to focus on learning the typical and most frequent uses first, with other senses introduced in more advanced instruction?

The author may want to consult literature from Psychology on the development of understanding polysemy by children in their native language. It might help shed some light on how/when language learners are equipped to handle multiple word senses.

Other comments:

Much of the introductory material does not seem necessary for this work, and it is not explained thoroughly enough for reader unfamiliar with the literature to grasp. I'd suggest beginning the paper with the paragraph beginning "Polysemy, the focus of the current study…" on page 3.

I didn't understand the comment about polysemous words being harder to understand, especially compared to a nonsense word. When would a learner encounter a nonsense word? This seems to refer to some experimental paradigm

that does not reflect normal language learning contexts.

Low contextual diversity seems to be an inference by the author based on low polysemy, but I don't think the data actually demonstrate it.

Figures 2, 3, and 4 seem to represent a small subset of the words analyzed - I wasn't sure why these words were the ones graphed.

Figure 1 could be cut - it doesn't provide any further information beyond the 65% overlap stated in the text.