

Critical Commentary

SELLING THE (WORD) FAMILY SILVER? A RESPONSE TO WEBB'S "LEMMA DILEMMA"

Benjamin Kremmel 

University of Innsbruck

In his commentary “The lemma dilemma,” Webb reflects on a discussion that has occupied some vocabulary researchers for a while now, and that has enjoyed increased attention, discussion, and publication interest in the past five years in particular. The question at the heart of it is whether the traditional lexical counting unit of the word family, often implicitly understood as including inflected and derived forms of a base word up to level six in Bauer and Nation’s (1993) taxonomy, should be operationalized in vocabulary research and teaching. In light of the title of Brown et al.’s (2020) recent paper published on this issue, which ambitiously suggests having settled the discussion by having found “*the* most appropriate lexical unit for L2 vocabulary research and pedagogy” (p. 1, emphasis added), I would agree with Webb’s call for more measured claims in principle, and echo that we still, as a field, seem some way away from stating confidently or even implying that we have discovered a definitive answer to this issue. Nor is the discussion settled whether there would or should even be *one* definitive answer to this. This short commentary does not set out to settle that either, of course, but merely aims to add further musings, reflections, and questions to ponder in our search for useful answers.

The advantage of settling on one standardized unit for all contexts and purposes appears evident. Particularly in the world of research and assessment we would certainly gain comparability between studies and interpretability of results. However, some would probably argue that the importance of comparability may be overestimated, given that there are often so many further factors that render vocabulary studies difficult to compare (e.g., different frequency lists, different operationalizations of vocabulary knowledge in the measurements). Webb, in his commentary, distinguishes between at least two contexts: research and pedagogy. He argues that smaller units, such as *lemmas* or *flemmas*, may be more valuable and appropriate in research, and larger lexical units more feasible and adequate for pedagogy. However, I would argue that it should be carefully considered whether to draw the line between exactly these two areas of application, as it further promotes the notion that these two are only marginally related. Instead, one could argue that for vocabulary research to be more relevant for pedagogical practice, the two worlds should speak the same “language” and have a similar focus when it comes to lexical counting units. This does not mean that I do not agree with Webb when he rightly

questions a general one-size-fits-all approach. It simply means that there may be more to be gained by also examining more closely different contexts and purposes within the two areas. When we are looking for “the most appropriate lexical unit for L2 vocabulary research and pedagogy,” should this not also mean we should research the question in L2s other than English, for example? At the very least, I would agree with Webb, does appropriacy often depend on factors such as proficiency level, type of use, and others listed in his commentary’s conclusion.

Apart from the issue of keeping the areas of research and pedagogy separate, Webb’s claims about the value of larger lexical units for pedagogy, and thus their increased ecological validity in research, also seem worth discussing. It appears unclear whether “presenting headwords together with their inflections and derivations” does provide a meaningful shortcut in lexical learning, or whether this is a practice that we would really find to be widespread in language classrooms around the world. Webb admits that there is a lack of research to date that has examined how the lexical unit influences L2 vocabulary learning. We similarly know very little about the actual use and perceived usefulness of word lists beyond lexical profiling for vocabulary learning, despite knowing a great deal of word lists at this point. Again, I would agree with Webb’s call for more research in this area before making grand assertions.

It is important to note at this point, however, that there is research evidence to suggest that the word family is not only questionable for the measurement of productive vocabulary knowledge, as Webb’s commentary might make it seem. Ward and Chenjundaeng (2009) and Kremmel and Schmitt (2016) have both demonstrated that learners struggle not only with the production of derivatives but also with identifying links between headwords and derivatives, receptively. Kremmel (2016) has put this forward as an argument for at least considering lemmas as counting units, and Brown et al. (2020) also elaborate on this point in their argument for lemmas being the most appropriate unit, citing multiple studies with different foci. However, what should not be ignored in this search for maximum measurement precision is that learners in these studies—despite not being able to establish all or a “sufficient” number of links between word family members—were always able to show a rather substantial knowledge of word family relationships or affixations. This appears to suggest some facilitation in learning that, albeit yet to be determined precisely, should not be neglected altogether.

Indeed, one of the driving forces behind the word family versus lemma versus flemma debate has been the area of vocabulary assessment. Where we are after precise measurement and valid score interpretations, it is at the very least problematic if we are giving learners too much credit for knowing entire word families when the evidence suggests that they don’t know them fully. Webb points out that smaller lexical units would result in less feasible tests, at least for more proficient learners. Testing 1,000 word families by means of 30 items does indeed seem more manageable than testing 3,281 lemmas with 98 items. A key chink in the argumentative armor, however, is that what the field seems to be currently doing is test 3,281 lemmas with 30 items, at the cost of score interpretability. Webb’s argument would also seem somewhat at odds with Laufer and Cobb’s (2020) recent findings that “lexical coverage is unlikely to be affected by the use of word families as the lexical unit” (Webb, 2021, p. 945). If that were the case, then why would we as a field not opt for the more precise unit? Perhaps Cobb and Laufer’s (2021) idea of nuclear word families could be a step toward finding an appropriate middle ground. Related to

this, it seems worth bringing issues of practicality into the appropriacy discussion as well, where Schmitt (2010) has argued that lemmas may be preferable to word families as their use involves less discussion as to what inflections and derivations should be included in a lexical unit.

Finally, it seems noteworthy that the discussion so far has largely neglected evidence relating to the psycholinguistic realities of word families or lemmas. While some studies have shown processing to happen somewhat along the lines of the notion of a word family (e.g., Bertram et al., 2000), others have argued based on their findings that “L2 learners rely more on lexical storage and less on combinatorial processing of morphologically complex words” (Silva & Clahsen, 2008, p. 245). Van de Vijver and Baer-Henney (2019) also recently demonstrated that the structure of the representation of a word family is far from a straightforward and conclusively addressed issue. Schmitt’s (2010) conclusion that “the psycholinguistic status of the word family is still undetermined” (p. 190) thus seems to hold true even now.

Questioning and arguing for or against the appropriacy of particular lexical units is what we should be doing as vocabulary researchers. Calls for having found the most appropriate for all contexts and uses may, however, be premature at this point. A great deal more research needs to be conducted and replicated regarding this issue that has only recently received the scholarly attention in the discussion that it deserves. But we don’t need to sell the word family silver just yet. We just need to keep exploring for which contexts and purposes it might be worth parting with this heirloom of the world of L2 vocabulary research and what concept could bring added value if we used it as a substitute.

REFERENCES

- Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6, 253–279.
- Bertram, R., Baayen, R., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, 42, 390–405.
- Brown, D., Stoeckel, T., McLean, S., & Stewart, J. (2020). The most appropriate lexical unit for L2 vocabulary research and pedagogy: A brief review of the evidence. *Applied Linguistics*. Advance online publication. <https://doi.org/10.1093/applin/amaa061>
- Cobb, T., & Laufer, B. (2021). The nuclear word family list: A list of the most frequent family members, including base and affixed words. *Language Learning*, 71, 834–871. <https://doi.org/10.1111/lang.12452>
- Kremmel, B. (2016). Word families and frequency bands in vocabulary tests: Challenging conventions. *TESOL Quarterly*, 50, 976–987.
- Kremmel, B., & Schmitt, N. (2016). Interpreting vocabulary test scores: What do various item formats tell us about learners’ ability to employ words? *Language Assessment Quarterly*, 13, 377–392.
- Laufer, B., & Cobb, T. (2020). How much knowledge of derived words is needed for reading? *Applied Linguistics*, 41, 971–998.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.
- Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11, 245–260.
- Van de Vijver, R., & Baer-Henney, D. (2019). Paradigms in the mental lexicon: Evidence from German. *Frontiers in Communication*. <https://doi.org/10.3389/fcomm.2018.00065>
- Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, 37, 461–469. <http://doi.org/10.1016/j.system.2009.01.004>
- Webb, S. (2021) The lemma dilemma: How should words be operationalized in research and pedagogy? *Studies in Second Language Acquisition*, 43, 941–949. <https://doi.org/10.1017/S0272263121000784>