How to produce vocabulary lists? Issues of definition, selection and pedagogical aims.
A response to Gabriele Stein

Introduction

On the surface level, producing vocabulary lists is easy. The task that once involved an enormous effort and countless hours of meticulous manual counting (cf. West, 1953) can today be achieved in a matter of seconds using current language corpora and corpus software. However, the vocabulary list is only meaningful if several key methodological and conceptual questions are addressed before it is produced. In her contribution, Stein raises several questions about the validity and usefulness of the New General Service List (new-GSL) (Brezina & Gablasova, 2015). Such reflection is genuinely needed because it invites critical rethinking of some of the basic tenets of vocabulary research. This short contribution reviews Stein's argument and replies to the main points raised in her paper. We identify and discuss three fundamental principles of wordlist creation, namely i) construct definition, ii) the target vocabulary identification and iii) the purpose of the wordlist, and show that the paradigm the new-GSL follows differs from the one used by Stein. Our contribution seeks to open a broader discussion about the principles of building wordlists for applied linguistic research and pedagogical purposes. This discussion draws and expands on the principles used in Brezina & Gablasova (2015) for the creation of the new-GSL.

Defining the construct

The first step in the creation of a vocabulary list should involve a definition of the vocabulary construct that the list seeks to represent. Since any type of 'vocabulary' (e.g. general, academic or medical) is merely an abstraction, we need to state where its boundaries lie. For example, the construct of 'core' or 'general' vocabulary has been defined in many different ways in the past; the various definitions differed from each other and were not always compatible. In Brezina & Gablasova (2015), the major aim (RQ1) was to validate the construct of general vocabulary defined as words that recur frequently across different texts, genres and modes of communication. By contrast, Stein employs the term 'core' vocabulary to refer to a very different type of construct, namely, to lexical items that are closely related to 'defining vocabulary' (i.e. a set of words used in dictionary definitions). This core vocabulary construct rests on two major assumptions: 1) the criterion of necessity and 2) semantic atomism. The first criterion assumes that a universal agreement can be reached about which words are necessary in communication; however, as Stein (2002, pp. 3 - 48) herself points out, researchers have disagreed passionately on this subject. Second, Stein's approach implies an atomistic view of linguistic meanings, which assumes that meanings of semantically complex words can be decomposed into simpler units of meaning. This, however, is problematic as a universal assumption (Goddard: 2011: 57-58). In sum, the nature of the two constructs is very different: while the new-GSL follows a quantitative, data-based paradigm, Stein's construct appears to rest on predominately subjective criteria.

Operationalising the construct, finding evidence and replicating the results

Once the vocabulary construct has been defined, the second step in the wordlist creation involves the operationalisation of the construct, i.e. deciding how to identify the words and

what type of evidence to use. Two distinct approaches (and their combinations) have been used in this process: the expert-based and the usage-based approach. The former makes use of authoritative sources of information about language such as dictionaries or (expert) native speakers. The latter relies on empirical evidence from language use (corpora) to select and/or rank words based on frequency and other quantitative criteria. While the new-GSL strictly follows the empirical evidence about word frequencies and distributions in the corpora, which is analysed according to transparent principles, Stein (2002) uses defining vocabularies from three dictionary sources (LDOCE, OALD, CIDE), compiled according to a mixture of (not always fully transparent) criteria. For example, it is not clear how the original source wordlists were compiled, how their authors defined and operationalised lexical core and whether these definitions are mutually consistent.

Arguably, West (1953), and other wordlist compilers after him, used additional criteria to corpus frequency partly to compensate for what they perceived as limitations of the frequency-only approach such as the use of a single corpus and raw frequency to identify the core words. Several of the limitations that led to the use of subjective criteria can now be addressed by advances in corpus techniques. The size of corpora has increased dramatically, providing more robust evidence of language use and driving advances in automated language analysis that allows annotation of linguistic items (e.g. for parts of speech) in large samples of language. Corpus techniques have also evolved beyond recording mere frequency to identifying other meaningful patterns in language use such as dispersion of linguistic items and formulaicity. With advances, new challenges appeared as well; however, there is no doubt in our mind that corpus-based vocabulary research is becoming a strong field with a large potential for innovation in the area of wordlist development as documented by innovative research on formulaic language and meaning disambiguation.

Purpose of the wordlists: Pedagogical potential

A strong motivation for the creation of wordlists is their application in language learning and teaching; wordlists are often evaluated in terms of their pedagogical usefulness. Since language learning and use are complex and dynamic processes, there is no one set of words that will be useful to all L2 learners. Learners' needs will likely differ across contexts, proficiency levels, L1 and cultural backgrounds and personal aims. Any serious claims about pedagogical usefulness should therefore be substantiated by empirical evidence. For example, a major strand in pedagogically-motivated vocabulary research connects usefulness for learners with frequency information (e.g. Gardner & Davies, 2014; Lei & Liu, 2016) since "actual frequency of occurrence is a more reliable indicator of usefulness than pure intuition" (Schmitt & Garnier, 2016, p.5). Frequency information plays an important role in current theories of language acquisition (Ellis, 2014). Frequent words are more often encountered and used by learners in spoken and written communication and their importance in curriculum planning has also been highlighted (Schmitt & Garnier, 2016). The claim that a set of words will be encountered repeatedly in meaning-oriented interactions can be examined through studies of lexical coverage of texts. With respect to the pedagogical uses of defining vocabulary, while we agree with Stein that defining vocabulary is undoubtedly useful for working with dictionaries, evidence is needed for the claim that it is also useful beyond this function (e.g. that the mastery of these words would improve learners' ability to

express meaning). Overall, we believe that more work is needed to link vocabulary lists with systematic, data-based investigation of learners' lexical needs in different contexts.

The new-GSL provides strong and principled evidence about which words are likely to be encountered by learners across a range of language situations. However, while we see a considerable pedagogical potential in the new-GSL, a vocabulary list alone is a long way from effective classroom materials and activities (as a good example see activities by Schmitt et al. (2011) based on the Academic Word List). To facilitate pedagogical use of the new-GSL, a new, interactive website (www.corpora.lancs.ac.uk/vocab) offers tools for further engagement with the list, including automatic calculation of lexical coverage of texts.

Final remarks

Although our approaches differ, we welcome challenges such as Stein's critical contribution and recognise the importance of other initiatives in the field such as The Common Core Vocabulary (Stein, 2002). Likewise we recognise the major achievement represented by West's General Service List, which brought innovation in the field by combining corpus evidence with further compilation principles compensating for the gaps in this evidence. In the development of the new-GSL, we sought to follow this legacy and redefine the 'core' or 'general' vocabulary using state-of-the-art corpus techniques as well as current SLA theories emphasising frequency-based learning. The following guidelines summarise the main principles that we follow in vocabulary list creation and use:

1. Explicitly define the vocabulary construct.
2. Operationalize the vocabulary construct using transparent and replicable criteria.
3. If using corpora, take corpus evidence seriously and avoid cherry-picking.
4. Use multiple sources of evidence to test the validity of the vocabulary construct.
5. Do not rely on your intuition/experience to determine what is useful for learners; collect evidence about learner needs to evaluate the usefulness of the list.
6. Do not present learners with a decontextualized list of lexical items; use/create contextualised materials instead.

REFERENCES

Brezina, V. and **Gablasova, D**. 2015. 'Is there a core general vocabulary? Introducing the New General Service List', *Applied Linguistics* 36/1:1-22.

Ellis, N.C. 2014. 'Frequency-based accounts of second language acquisition' in S. Gass and A. Mackey (eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 193-210). London: Routledge.

Gardner, D. and **Davies, M.** 2014. 'A new academic vocabulary list', *Applied Linguistics*, 35/3: 305-327.

Garnier, M. and **Schmitt, N.** 2015. 'The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses', *Language Teaching Research* 19/6: 645-666.

Goddard, C. 2011. *Semantic analysis: A practical introduction*. Oxford University Press.

Lei, L. and **Liu, D.** 2016. 'A new medical academic word list: A corpus-based study with enhanced methodology', *Journal of English for Academic Purposes 22*: 42-53.

Schmitt, D., **Schmitt, N.**, and **Mann, D.** 2011. *Mastering the academic word list*. Pearson Longman.

Stein, G. 2002. *Developing your English vocabulary*. Stauffenburg.