Corpora and ELT: the Frequency Lists Issue, Quantitative and Cognitive Factors

Aquilino SÁNCHEZ & Raquel CRIADO
Department of English Studies
University of Murcia
Murcia, Spain
asanchez@um.es; rcriado@um.es

The amount of words a learner knows has often been taken as an indicator of her/his language proficiency. Laufer (2010: 26) states that "vocabulary may be the major factor in reading comprehension". In this sense, corpora and the frequency lists derived from them may be taken as highly useful sources for the selection and elaboration of language teaching materials. This potential was already perceived in the first part of the 20th century (Thorndike & Lorge, 1944; West, 1953) and became evident in the second part of the same century, when electronic corpora increased facilities for the elaboration of more reliable frequency lists (Kucera & Francis, 1967).

Scholars usually emphasize the importance of the most frequent words because speakers need them for the most basic communicative situations (Laufer, 2010; Nation, 2001; Waring, 2003). Consequently, the most frequently occurring words should be learnt first. This is what textbook authors, at least since the seventies, have aimed at when deciding on the lexical component to be included in each one of the predefined linguistic levels, or when manipulating or reshaping the texts selected.

Besides, studies in SLA and cognition highlight the neural limitations in the acquisition of knowledge, heavily conditioned by the capacity of human working memory (Anderson, 2010; Schmidt, 1994; Ullman, 2004; among others). The rules governing language acquisition in the brain and information processing and consolidation in long-term memory do not allow for the quick and "easy" learning of all the lexical items theoretically needed for the development of each linguistic level. This paper addresses the issue of whether teaching materials have adapted or not, and to what extent, to a) authors' current stance in connection to frequency lists and their use in vocabulary teaching; and b) current cognitive and SLA research showing students' limitations for vocabulary learning and their real learning potential in specific time frames (Waring, 2003).

The research is based on the vocabulary analysis of an internationally-known ELT textbook series published by Oxford University Press: *New English File* (Elementary (2004), Pre-intermediate (2005), Intermediate (2006)), covering A1, A2 and B1 levels. Firstly, we compiled an *ad hoc* corpus with the words in the textbooks. Secondly, we quantified the vocabulary included in each one of them and compared it to the lexical expectations per level. Thirdly, we correlated the vocabulary of each textbook with the BNC-based frequency ranges (Nation, 2001). Finally, we analyzed the amount of lexical learning that students should attain as expected by textbook authors against the students' real learning potential as uncovered by the research previously indicated.

Data obtained will supply us with a reliable picture on how each textbook from the series progressively adapts to the claims regarding linguistic levels, and whether they adjust to frequency lists derived from corpora studies and take into consideration the students' cognitive potential as revealed by research in Cognitive Psychology. Pedagogical recommendations for materials design and adaptation and vocabulary teaching in general will be drawn as framed within the insights from Corpus Linguistics, SLA and Cognitive Psychology studies.

References

Anderson, J. R. (2010). *Cognitive Psychology and its Implications* (7th ed.). New York: Worth Publishers.

Kucera, H. & Francis, W. N. (1967). *Computational Analysis of Present Day American English*. Providence: Brown University Press.

Laufer, B. (2010). 'Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension.' *Reading in a Foreign Language* 22: 15-30.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Oxenden, C. & Latham-Koenig, C. (2004, 2005, 2006), *New English File Elementary, Pre-intermediate, Intermediate Student's Books*. Oxford: Oxford University Press.

Schmidt, R. (1994). 'Implicit learning and the cognitive unconscious: Of artificial grammars and SLA'. In Ellis, N.C. (ed.). *Implicit and Explicit Learning of Languages*. London: Academic Press. 165-209.

Thorndike, E.L. & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. Teachers College: Columbia University.

Ullman, M. T. (2004). 'Contributions of Memory Circuits to Language: the Declarative / Procedural Model.' *Cognition* 92: 231-270.

Waring, R. (2003). 'At what rate do learners learn and retain new vocabulary from reading a graded reader?' *Reading in a Foreign Language* 15: 130-163.

West, M. (1953). *A General Service List of English Words*. London: Longman: Green & Co.