

# Annotating Pronunciation Errors in the PLEC Spoken Learner Corpus

Magdalena ZAJĄC & Piotr PEZIK

Institute of English Studies

University of Lodz

Łódź, Poland

zajac1234@gmail.com; piotr.pezik@gmail.com

Acquiring English pronunciation frequently causes great difficulties for non-native learners. It is a complex process affected by the L1 sound system as well as a number of additional elements such as language aptitude, amount of native input, motivation, various social and biological factors (e.g. Flege & Liu, 2001; Suter, 1976, in Flege, MacKay & Piske, 2001; Tahta *et al.*, 1981, in Flege, MacKay & Piske, 2001; Thompson, 1976, in Dowd, Zuengler & Berkowitz, 1990), etc. As stated by (Szypra-Kozłowska, 2012a), the greatest threat to effective communication is posed not by segmental and prosodic inaccuracies which result mainly from L1 transfer (e.g. *foreign* realized as [fɔʀ'in] by Polish learners), but rather by “idiosyncratically deviant words”, i.e. realizations where a given segment (or segments) in a word is substituted by another which is phonologically and phonetically distant from it and/or in which word stress is incorrectly placed (e.g. *foreign* realized as [fɔ'rejn] by Polish learners). The latter type of mispronunciation may be brought about by several factors. (Szypra-Kozłowska, 2012b) lists the following as some of the major problem areas: spelling-related problems (e.g. pronouncing the silent letter <b> in *lamb*), phonetic ‘false friends’ (e.g. *chaos* realized with a /h/ rather than a /k/ by Polish learners), stress-related issues (e.g. *in'dustry*), pronunciation of words with morphological alternations in related forms (e.g. the pronunciation of *breathe* and *breath*). Given their potentially critical role in successful communication, it seems that frequently mispronounced words should be given a high priority when teaching English pronunciation. At the same time, the full scope of segment substitution as a pronunciation error is difficult to define without sufficient evidence.

This paper reports the results of a study aimed at developing a corpus-based typology of segment substitution mispronunciations. The study is based on the spoken component of the Polish Learner English Corpus (PLEC), which contains time-aligned interviews and other spoken interactions of Polish learners of English. Parallel to the orthographic tier, the PLEC transcriptions contain time-aligned manual annotations of mispronounced words. The explicit error annotation of the corpus makes it possible to investigate the relative frequency of word mispronunciations as well as to look for possible patterns among these mispronunciations. The paper describes the process of annotating such errors and representing them in the central corpus database. Using the corpus evidence available, the authors also compare the frequency of mispronunciations in which one or more segments are replaced with phonetically remote sounds and/or include incorrect stress placement. Issues such as stress placement in word-class pairs, e.g. *'import* (N) and *im'port* (V) are dealt with by examining the context in which they occur. One practical outcome of this study is a corpus-based index of the commonly mispronounced words by Polish learners of English.

## References

- Dowd, J, Zuengler, J & Berkowitz, D. (1999). ‘L2 Social Marking: Research Issues.’ *Applied Linguistics* 11: 16-29.
- Flege, J. & Liu, S. (2001). ‘The Effect of Experience on Adults’ Acquisition of a Second Language.’ *Studies in Second Language Acquisition* 23: 527-552.

- Flege, J., MacKay, I. & Piske T. (2001). 'Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191-215.
- Szpyra-Kozłowska, J. (2012a forthcoming). 'Mispronounced lexical items in Polish English of advanced learners.' *Research in Language*. London: Versita.
- Szpyra-Kozłowska, J. (2012b forthcoming). 'On the irrelevance of sounds and prosody in foreign-accented English.' In: Waniek-Klimczak, E. & Shockey, L. (eds). *Teaching and Researching English Accents in Native and Non-native Speakers*. Heidelberg: Springer.
- Tahta, S., Wood, M. & Loewenthal, K. (1981). 'Foreign Accents: Factors Relating to Transfer of Accent from the First Language to a Second Language.' *Language & Speech* 24: 265-272.
- Thompson, R. (1976). 'Mexican-American English: Social Correlates of Regional Pronunciation.' *American Speech* 50: 18-24.