Native language effects on pronunciation accuracy in L2 English

Jeroen Geertzen, Theodora Alexopoulou, Brechtje Post, and Anna Korhonen
{jg532, ta259, bmbp2, alk23}@cam.ac.uk
DTAL -- University of Cambridge

We investigated longitudinal aspects of word pronunciation acquisition for learners of English as a second language (L2). By using automatic speech recognition (ASR), we obtained an accuracy measure of how closely the learners' pronunciation resembles that of first language (L1) speakers, and used this measure to look for differences between four L1's: Russian, German, Italian and French.

Speech was sampled from the spoken/speech part of the EF-Cambridge Open Language Database (Post et al., 2012), which is developped by the international educational organization EF Education First and the University of Cambridge. These speech data are being collected from learners enrolled on ``Englishtown'', EF's e-learning platform, which provides 16 CEFR-aligned skill/proficiency levels.

The spoken part of EFCAMDAT contains scripted speech in reading-aloud exercises, and elicited speech in listen-and-repeat exercises in which the student is asked to repeat a word or phrase and attempt to mimic its pronunciation and intonation. Speech data of these two types of exercises were collected during a three-month period for the L1s of interest. All production of each learner who enrolled on Englishtown during this period was recorded, resulting in a dataset containing word samples of single word and small sentence recordings at each of the 16 levels from Russians (N=291), Italians (N=197), Germans (N=170), and French (N=156) learners, totalling 509,001 word samples involving 5,386 unique words.

An ASR system (Dragon ASR) was trained on native speech for each exercise separately, assuring low word error rates. The word recognition confidence scores, estimated by the ASR by comparing the learner's speech to the trained native speech, provided our pronunciation accuracy measure. When averaging word scores first by learner, and then by level, we find that pronunciation accuracy shows a steady upward trajectory as learners get more exposed to pronunciation exercises.

Word scores averaged by learner and then by L1, show that whereas French, Russian, and Italians increase their pronunciation accuracy considerably across increasing levels, this is distinctly less so for Germans. The most likely explanation lies in the phonological similarity between English and German as opposed to English and the other two L1s.

We will further provide an analysis on individual words that show most divergence amongst L1s and review their phonetic properties, and future work will involve the use of phoneme lattices from ASR output to analyse phonetic L1 effects. Most importantly, we believe this approach will be particularly useful to explore large learner data.

Reference:
Post, B., D. Alexopoulou, J. Geertzen, and A. Korhonen (2012). The EF Cambridge Corpus of Learner English: L2 speech data. Abstract presented at the 2012 IVACS Symposium. Cambridge, UK.