Cambridge English Centenary
Symposium on Speaking Assessment

Improving intelligent tutoring of pronunciation consonant cluster problems

Gary Pelton, Carnegie Speech

The past decade has seen improvements in the algorithmic detection of errors in non-native pronunciation (Eskenazi 1996, Neumeyer, Franco, Digalakis and Weintraub 2000). This work has led to the development of products that can be used to improve a user's pronunciation. The work of Akahane-Yamada, Kato, Adachi, Watanabe, Komaki, Kubo, Takada, and Ikuma (2004) has been converted to a complete pronunciation training system for Japanese speakers who want to speak English. It has been commercialized by ATR in Japan. NativeAccent™ by Carnegie Speech is another commercial product used to train non-native speakers to improve their pronunciation of English.

NativeAccent started as the Fluency Project (Eskenazi 1996) at Carnegie Mellon University. It has been converted to a complete pronunciation training system, and over the last 10 years thousands of people have used NativeAccent to improve their pronunciation. NativeAccent compares a user's pronunciation to a statistical model of native speakers using a technique called pinpointing. A close match is considered good pronunciation. We have these models for both Midwest American English speakers and British English from speakers whose speech is close to what used to be called BBC English. Rather than doing a detailed analysis of how NativeAccent detects pronunciation issues, this paper documents one of our explorations in how NativeAccent can be improved to give our users a better experience and to better address their specific problems. This exploration is done by examining the logs of user data looking for patterns of problems. Focusing on the student's specific problems is important in the NativeAccent intelligent tutor methodology.

The learning improvement from the use of intelligent tutoring has been shown in many domains. In particular Anderson (1993) shows reduction of one third the training time to achieve a particular level of performance in learning a programming language. The work of Koedinger, Anderson, Hadley and Mark (1997) shows an improvement of 1 standard deviation for students using an Algebra 1 intelligent tutor compared to similar students in a teacher-run Algebra 1 class. Carnegie Speech has published studies (see Eskenazi, Ke, Albornoz and Probst 1998) about how well our users do using this intelligent tutor approach. A more recent unpublished study of 120 people showed the 60 people who only participated in a teacher run pronunciation class (control group) improved 62% in seven hours, and the 60 people in the test group using NativeAccent during half of their pronunciation class time improved 104%. This difference was not quite statistically significant ($r=0.07$), but the time using NativeAccent was relatively small. We see substantial improvement in about 10 hours of use.

Perfect intelligent tutoring relies upon knowing what skill deficiency is the root cause of a mistake and having curriculum that helps the student learn to avoid that mistake. The graphs showing the transformation of errors into learning curves once errors can be assigned to a cause (rule) in Anderson (1993) are very compelling. The empirical results on tutors that

can assign root causes supports that research. Pinpointing provides probabilistic root cause information for substitution, affrication, deletion and other mistakes related to a single phone. Co-articulation problems and those problems associated with a phone in a context, along with errors in automatic speech analysis reduce the effectiveness of the tutoring. However, these problems aren't systematic, and pinpointing does pick up on the systematic issues the user exhibits. Carnegie Speech's experience shows that the small amount of random error doesn't affect the tutoring overmuch.

Pronunciation problems within consonant clusters are the focus of the problems investigated in this paper. Consonant clusters are where the native pronunciation of a word involves a sequence of consonants without intervening vowels. Consonant clusters are a known pronunciation problem in non-native speech (Hultzen 1993). For that reason, NativeAccent's curriculum includes consonant clusters exercises in its lessons. For example, the /t/ lesson might start with exercises on simple words like "tap", "bat" and "later" but the lesson can also include words with "t" in a consonant cluster like "string". The question asked in this paper is whether we should change our detection algorithms and/or our tutoring because our users are exhibiting systematic problems with pronouncing phones within consonant clusters that our current system could handle better. This paper uses data mining on the logs of NativeAccent users to answer both this question and to provide more insight into the various consonant cluster problems our users exhibit. The intelligent tutoring mechanism underlying NativeAccent will be described in detail so that the reader can understand how this data can improve the tutoring process.

References

Akahane-Yamada, R, Kato, H, Adachi, T, Watanabe, H, Komaki, R, Kubo, R, Takada, T and Ikuma, Y (2004) ATR CALL: A speech perception/production training system utilizing speech technology, *Proc. ICA 2004*, 2004, vol. III, 2319–2320.

Altenberg, E (2005) The judgment, perception, and production of consonant clusters in a second language, *IRAL*, 43, 53–80.

Anderson, J R (1993) *Rules of the Mind*, Hillsdale, NJ: Erlbaum.

Celce-Murcia, M, Brinton, D and Goodwin, J (1996) *Teaching Pronunciation*, Cambridge: Cambridge University Press, 1996.

Eskenazi, M (1996) Detection of foreign speakers' pronunciation errors for second language training - preliminary results, *Proc. International Conference on Spoken Language Processing*, Philadelphia.

Eskenazi, M, Ke, Y, Albornoz, J and Probst K. (1998) The Fluency Pronunciation Trainer: Update and user issues, presented at the STiLL Workshop on Speech Technology in Language Learning, Marhollmen.

Hultzen, L (1965) Consonant clusters in English, *American Speech*, 49(1), 5–19.

Koedinger, K R, Anderson, J R, Hadley, W H and Mark, M A (1997) Intelligent tutoring goes to school in the big city, *International Journal of Artificial Intelligence in Education*, 8, 30–43.

Neumeyer, L, Franco, H, Digalakis, V and Weintraub, M (2000) Automatic scoring of pronunciation quality, *Speech Communication*, 30(2), 83–93.