Chapter 16

Computer-Assisted Pronunciation Training and Assessment (CAPTA) Programs: Requirements, the Current State of Affairs, and Challenges for the Future

Chiharu Tsurutani
Griffith University, Australia

ABSTRACT

Pedagogical support for pronunciation tends to fall behind other areas of applied linguistics and CALL (Computer-Assisted Language Learning) due to technological difficulty in speech recognition and the lack of knowledge in phonetics of both language teachers and learners. This chapter discusses the gap between the need for pronunciation training and the capacity of CAPTA programs in terms of phonetic and phonological development of second-language (L2) learners. Pronunciation difficulties experienced by L2 learners will be explained cross-linguistically, and the most recent developments in the production of CAPTA programs will be discussed in relation to the type of pronunciation errors dealt with by these programs. Considering that native-like pronunciation is no longer required in the current multi-lingual society, the author proposes achievable and pedagogically sound goals for the development of CAPTA programs as well as for L2 learners.

INSTRUCTION

Since computer-based exercises and assessments became available in language classrooms, numerous programs have been developed to enhance the four language skills: reading, listening, writing, and speaking. Speaking is the only area that was not fully addressed in the range of CALL programs and marking is not automated in officially recognized proficiency tests, such as those conducted under the International English Language Testing System (IELTS) and the American Council on the Testing of Foreign Languages (ACTFL) (Jamieson, 2005;
Computer-Assisted Pronunciation Training and Assessment (CAPTA) Programs

Bernstein, Van Moere, & Cheng, 2010). While other skills can be trained and assessed using existing technology for CALL programs, speaking is hard to assess since L2 speech contains various phonetic deviation. The technological difficulty of assessing the speech produced by second-language learners has been a pressing issue for speech engineers (see the section, Recent Development in CAPTA Programs); we can only wait for the advancement of speech technology. However, achieving the intelligible level of pronunciation is crucial for communication. As globalization and internationalization progress, the need to communicate in languages other than one’s own has become almost inevitable, not only in travelling, but also in business and education, which requires a reasonable level of intelligibility. In order to respond to the on-going needs of L2 learners, we need to continue our work for developing new methods and ideas. The training and assessment of pronunciation are the areas that benefit most from the nature of CALL programs, which involve independent study, unlimited repetition and immediate response and feedback. By focusing on the crucial factor for the improvement of pronunciation, we should be able to provide language learners with the guidance and tools to acquire the adequate level of pronunciation efficiently rather than waiting for slow and gradual improvement through exposure to the target language. This chapter starts with an illustration of the phonetic and phonological challenge L2 learners face. Following that, the capacity of current CAPTA programs will be discussed in relation to the types of L2 speech errors. Finally, the target of pronunciation teaching in the 21st century—speech that is intelligible to native listeners—will be defined and discussed in relation to its incorporation into the assessment criteria of CAPTA programs.

TYPES OF L2 PRONUNCIATION ERRORS

Before discussing the factors that are crucial in the improvement of pronunciation, we need to know the mechanism behind L2 errors. What kind of hurdles do L2 learners have to overcome? Pronunciation errors are not simply incorrect production of vowels or consonants. Spoken language consists of three components: segmental features, suprasegmental features and prosodic organization.

Spoken language:

- Segmental features (consonants, vowels).
- Suprasegmental features (pitch, duration, loudness).
- Prosodic organization (accent, intonation).

Segmental errors are the incorrect production of consonants and vowels, and what most people associate with pronunciation errors. Suprasegmental errors refer to errors in pitch, duration, and loudness of sounds, which are larger units than segmentals. Prosodic organization is phonological organization of language, such as accent, tone, and intonation. While segmental features are expressed by orthography and easy to point out, incorrect suprasegmental features and prosodic organization are difficult to capture and explain. Nevertheless, prosodic features (suprasegmental and prosodic organization) are known to have more impact on learners’ intelligibility, as judged by native speakers, than do segmental features (Anderson-Hsieh, Johnson, & Koehler, 1992; Hahn, 2004; Tsurutani, 2009; Warren, Elgort, & Crabbe, 2009). This leads to the question of how prosodic differences between languages affect L2 production.

Figure 1 outlines Archibald’s (1998) accent (prosodic) typology.

First, languages are divided into two groups, one that uses accent and one that does not. Accent relates to suprasegmental features, and makes one