Why train pronunciation with computer technology?

Many second language (L2) learners want to receive pronunciation training, but typically, language teachers do not spend much class time on pronunciation (e.g., Foote, Holtby, & Derwing, 2011). When used prudently, Computer-Assisted Pronunciation Teaching (CAPT) is a useful solution. It enables learners to work autonomously and at their own pace on those aspects of pronunciation that they need to work on most. CAPT has been found to be as effective as classroom pronunciation training in some studies, especially when overseen by the instructor, and L2 learners who use it often can develop more intelligible pronunciation (e.g., Lee, Jang, & Plonsky, 2015; Martin, 2018; Thomson & Derwing, 2015).

Which aspects of pronunciation can we train with technology?

Because our ability to produce language is related to our ability to perceive it, it is important that CAPT technologies train both perception and production. Including perceptual training can be quite efficient, since research has demonstrated that learners' perception and production can both improve through computerized training that focuses solely on learners' perception (e.g., Lee & Lyster, 2016; Thomson, 2011).

The intelligibility of L2 learners' speech is affected by the production of speech segments and prosodic features including word stress, sentence stress, and intonation. CAPT can be used to train both individual speech sounds and prosody. For example, effective CAPT has been demonstrated on segments that carry a high functional load (e.g., Liakin, Cardoso, & Liakina, 2015), stress assignment (Hincks & Edlund, 2009), and pitch contours (Hirata, 2004; Chun, Jiang, & Ávila, 2013). Training one pronunciation feature (e.g., pitch contours) may have a positive impact on other pronunciation features (e.g., speech segments, as demonstrated in Hardison, 2004). Studies have shown that the imitation of model speakers through shadowing is an effective way to teach a range of pronunciation features, and it may lead to improvements in both comprehensibility and fluency (e.g., Foote & McDonough, 2017). See also the essays on shadowing (Foote, 2018) and mirroring (Meyers, 2018) under Teaching Techniques on this website.

What types of technology are out there?

Two basic kinds of software have been developed to train and evaluate pronunciation: software and apps specifically designed for language learners and those designed for researchers (in general, there is steep learning curve required to make software designed for research, such as PRAAT, useful for L2 learners). Software designed for other purposes can also be used to effectively target pronunciation (e.g., apps for basic sound recordings in Foote & McDonough, 2017; online dictionaries in Levy & Steel, 2015; apps designed to slow down music in Chan, 2017; text-to-speech systems in Liakin, Cardoso, & Liakina, 2017; *Twitter* in Fouz-González, 2017). Another type of software sometimes used to train pronunciation is speech recognition software designed for native speakers. Although some researchers have reported using it successfully in their language classrooms (e.g., McCrocklin, 2016), L2 learners can often

become frustrated with its inability to recognize accented but comprehensible L2 speech and its inability to appropriately identify errors. Some researchers, therefore, recommend against using such software with L2 learners (e.g., Derwing & Munro, 2015).

CAPT is only beginning to make use of advanced automatic speech recognition (ASR) technology that can both effectively recognize and provide appropriate feedback on L2 speech; at this point the best ASR technology available is for learners of Dutch (Strik, Colpaert, van Doremalen, & Cucchiarini, 2012).

What are key features of pedagogically sound CAPT?

Derwing and Munro (2015) remind us that technology should be used judiciously to train pronunciation. A primary reason to exercise caution when choosing software is that although a great deal of advanced pronunciation software has been released in recent years, much of it is not pedagogically sound (Foote & Smith, 2013).

L2 learners often do not notice many pronunciation features unless they are pointed out to them (e.g., Chun, 2013). Therefore, two key features of effective CAPT resources include enhanced input and feedback that is easy to interpret. Enhanced input directs learners' attention to those aspects of pronunciation that are most important. It can involve simple highlighting or underlining of key sounds, easy-to-interpret waveforms or speech spectrograms that show segment durations (e.g., Motohashi-Saigo & Hardison, 2009), intonation contours (e.g., Chun, Kiang, Meyr & Yang, 2015), or auditory enhancement that involves making key features louder than others (e.g., Lee, 2009; Martin & Jackson, 2016). The presentation of sound files together with images or videos of a human head showing the salient articulatory movements may also be effective (Inceoglu, 2015).

One of the biggest problems for L2 learners is that they have a hard time detecting their own pronunciation errors (e.g., Foote, 2010; Menzel et al., 2001). For this reason, CAPT resources should provide high-quality, relevant feedback. That is to say, pedagogically sound pronunciation technology not only indicates learner errors, but it also provides learners with pointers on how to improve their pronunciation to make it more intelligible (e.g., Atwell et al., 1999; Engwall, Wik, Beskow, & Granström, 2004).

When it comes to the software itself, effective CAPT systems allow learners to enter a given program at the point that enables them to focus on their difficulties (Derwing, 2010) and track their improvement over time (Levis, 2007); an example of this is Thomson's (2018) Englishaccentcoach.com. In addition, as Rosell-Aguilar (2017) notes, effective systems should be intuitive and have clear instructions. Finally, Levy and Steel (2015) emphasize that gamified systems may motivate learners and encourage higher levels of engagement.

References
Atwell, E., Herron, D., Howarth, P., Morton, R., & Wick, H. (1999). Pronunciation training: requirements and solutions. *ISLE deliverable.* Available at http://nats-www.informatik.uni-hamburg.de/~isle/public/D14/D14.pdf

Chan, M. J. (2017). *Anytune* slows down sound tracks for language practice. In M. O'Brien & J. Levis (Eds.), *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Calgary, AB, August 2016 (pp. 191–194). Ames, IA: Iowa State University. https://apling.engl.iastate.edu/alt-content/uploads/2017/05/PSLLT_2016_Proceedings_finalB.pdf

Chun, D. M. (2013). Computer-assisted pronunciation teaching. In C. A. Chapelle (Ed.), *Encyclopedia of applied linguistics* (pp. 823-834). Oxford: Wiley-Blackwell. doi: 10.1002/9781405198431.wbeal0172

Chun, D. M., Jiang, Y., & Ávila, N. (2013). Visualization of tone for learning Mandarin Chinese. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*. (pp. 77–89). Ames, IA: Iowa State University. https://apling.engl.iastate.edu/alt-content/uploads/2015/05/PSLLT_4th_Proceedings_2012.pdf

Chun, D. M., Jiang, Y., Meyr, J., & Yang, R. (2015). Acquisition of L2 Mandarin Chinese tones with learner-created tone visualizations. *Journal of Second Language Pronunciation, 1*(1), 86–114.

Derwing, T. M. (2010). Utopian goals for pronunciation teaching. In J. Levis and K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (24–37). Ames, IA, Iowa State University.

Derwing, T. M., and Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam/Philadelphia: John Benjamins.

Engwall, O., Wik, P., Beskow, J., & Granström, B. (2004). Design strategies for a virtual language tutor. Paper presented at the ICSLP 2004 Conference. Available at: http://www.speech.kth.se/ctt/publications/papers04/icslp2004_tutor.pdf

Foote, J. (2010). Second language learners' perceptions of their own recorded speech. *PMC Working Paper Series* W P10-02 (pp. 3–27). Edmonton: Prairie Metropolis Centre. https://sites.ualberta.ca/~pcerii/WorkingPapers/WP1002.pdf

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*(1), 1–22.

Foote, J., & Smith, G. (2013, September). Is there an app for that? Paper presented at the 5th Pronunciation in Second Language Learning and Teaching Conference, Ames, Iowa.

Foote, J., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation, 3*(1), 34–56.

Fouz-González, J. (2017). Pronunciation instruction through Twitter: the case of commonly mispronounced words. *Computer Assisted Language Learning, 30*(7), 631–633.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology, 8*(1), 34–52.

Hincks, R., & Edlund, J. (2009). Promoting increased pitch variation in oral presentations with transient visual feedback. *Language Learning & Technology 13(3)*, 32–50.

Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning, 17*(3-4), 357–376.

Inceoglu, S. (2015). Audiovisual and auditory-only perceptual training: Effects on the pronunciation of French nasal vowels. In J. Levis, R. Mohammed, M. Qian, & Z. Zhou (Eds). *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching*

Conference. Santa Barbara, CA (pp. 104–114). Ames, IA: Iowa State University. https://apling.engl.iastate.edu/alt-content/uploads/2015/05/PSLLT_6th_Proceedings_2014.pdf

Lee, J. (2009). *The effects of pronunciation instruction using duration manipulation on the acquisition of English vowel sounds by pre-service Korean EFL teachers*. Unpublished Ph.D. thesis, University of Kansas.

Lee, J., Jang, J., Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345–366.

Lee, A. H., & Lyster, R. (2016). Can corrective feedback on second language speech perception errors affect production accuracy? *Applied Psycholinguistics, 38*(2), 371–393.

Levis, J. (2007). Computer technology in teaching and researching. *Annual Review of Applied Linguistics, 27*, 184–202.

Levy, M., & Steel, C. (2015). Language learner perspectives on the functionality and use of electronic language dictionaries. *ReCALL : The Journal of EUROCALL, 27*(2), 177–196.

Liakin, D., Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal, 32*(1), 1–25.

Liakin, D., Cardoso, W., & Liakina, N. (2017). The pedagogical use of mobile speech synthesis (TTS): Focus on French liaison. *Computer Assisted Language Learning, 30*(3-4), 348–365.

Martin, I. A. (2018). *Bridging the gap between L2 pronunciation research and teaching: Using iCPRs to improve novice German learners' pronunciation in online and face-to-face classes.* Unpublished PhD dissertation. Penn State University.

Martin, I. A., & Jackson, C. N. (2016). Pronunciation training facilitates the learning and retention of L2 grammatical structures. *Foreign Language Annals, 49*(4), 658–676.

McCrocklin, S. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System, 57,* 25–42.

Menzel, W., Herron, D., Morton, R., Pezzotta, D., Bonaventura, P., & Howarth, P. (2001). Interactive pronunciation training. *ReCALL,* 13(1), 67–78.

Motohashi-Saigo, M., & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology 13(2),* 29–47.

Rosell-Aguilar, F. (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO Journal, 34*(2), 243–258.

Strik, H., Colpaert, J., Van Doremalen, J., & Cucchiarini, C. (2012). The DISCO ASR-based CALL system: practicing L2 oral skills and beyond. *Proceedings of the Conference on International Language Resources and Evaluation (LREC 2012),* Istanbul, May 2012.

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal, 28*(3), 744–765.

Thomson, R. I. (2018). *Englishaccentcoach.com*

Thomson, R. I. & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A Narrative Review. *Applied Linguistics, 36*(3), 326–344.