RESEARCH BRIEF SERIES

Teaching and Learning English Pronunciation by Generating the Vocal Tract Shapes from the Frequency Domain Information

Jer-Ming Chen and Dorien Herremans

KEY IMPLICATIONS

- Machine-learning strategies were successfully explored and implemented for (1) phoneme classification, (2) segmentation of variable length phonemes, (3) accent identification, and (4) vowel gesture identification.

- We now know that a sparse acoustic representation of just three speech features, namely fundamental frequency (f_0), first and second formant frequencies (F1 and F2, respectively), are sufficient to offer a reasonable degree of articulatory discrimination and feedback for pronunciation of vowel sounds.

- "Speechie" App is the first smartphone implementation offering real-time off-line visual remedial articulatory cues to a child for pronunciation training, and performs comparably with the top desktop voice/ speech research software.

BACKGROUND

Developing good oral communication skill is an integral part of the Singapore Ministry of Education's English Language 2010 Syllabus. However, learning pronunciation via the traditional way of imitating the teacher's voice or listening to audio examples alone is inherently difficult and not sufficient, because there is no easy or intuitive feedback information for the student to understand whether his/her pronunciation is right or wrong.

The main technical hurdle faced by researchers so far is that automatic (i.e. machine-based) recognition, segmentation and classification of speech elements (phonemes) is inherently very challenging, typically requiring considerable computational power to perform with high accuracy, and thus has not been attempted "in the field" so far.

FOCUS OF STUDY

This study aimed to develop an interactive (game-based) smartphone app to help language learners improve pronunciation by providing real-time visual feedback related to their articulation (lip shape, tongue position), such that learners can intuitively modify their articulation to produce the correct target speech sound. To achieve this, machine-learning algorithms were applied to speech sounds in order to identify which acoustic features of speech offered the most efficient articulatory and visual feedback, which were then implemented as a smartphone app in a lean and efficient manner.

KEY FINDINGS

By innovatively applying machine learning insights gained in this project, trained on appropriate datasets, we successfully implemented speech features in an intuitive and visual manner in a game-based smartphone App to provide real-time visual cues, to be used as a pronunciation training tool to be deployed both in and out of the classroom. These are findings which are reported and published in the references section.

SIGNIFICANCE OF FINDINGS

Classroom Practice

"Speechie" App can be deployed in English Language lessons as (1) a supplementary means for students to practice good pronunciation, and (2) for teachers to monitor the student's pronunciation accuracy, progress, and practice.

Pedagogy Research

"Speechie" App is an effective and accurate tool to assess how well students are picking up correct English pronunciation and the efficacy and progress of students' English pronunciation and pronunciation vs intervention techniques.

Follow-up Work

"Speechie" App is currently developed for training English vowels only.

1. English consonants can be targeted next, although more sophisticated machine-learning algorithms need to be developed, along with more efficient processing architecture to work on the smartphone (or use more powerful smartphones).
2. In addition to English, training for other target languages can also be developed (this requires database training on new language sets).

3. Further, the App's user-interface can be improved, as well as making the user-graphical interface more intuitive, and enhancing the 'gamification' of the training process and objectives, as further pedagogical insight is obtained.

PARTICIPANTS

None (App validation and testing was performed on the research team members).

RESEARCH DESIGN

The machine-learning algorithm was trained in-silica on a speech database, identifying the sparse speech feature representation associated with speech articulation. Accordingly, an App implementing the speech features identified were developed as a game implementing visual articulatory feedback. The App's performance was then validated/benchmarked for accuracy against standard speech research software.

REFERENCES

Khwaja, M. K., Vikash, P., Arulmozhivarman, P., & Lui, S. (2016). Robust phoneme classification for automatic speech recognition using hybrid features and an amalgamated learning model. *International Journal of Speech Technology, 19*(4), 895-905.

Upadhyay, R., & Lui, S. (2018). Foreign English Accent Classification Using Deep Belief Networks. In *ICSC 2018 IEEE 12th International Conference on Semantic Computing* (pp. 290-293). IEEE.

Balamurali, B. T., & Chen, J.-M. (2018). Automated Classification of Vowel-Gesture Parameters using External Broadband Excitation. In *Interspeech 2018* (pp. 2315-2318). India.

Tan, E.Y., Balamurali BT, Lui, S., and Chen J.-M. (2018). *"SpeechieX" – an app for continuous tracking of pronunciation parameters in the field*. Poster session presented at In Current Issues in Child Bilingual Development, Macquarie University, Sydney.

About the authors

Jer-Ming CHEN and Dorien HERREMANS are with the Singapore University of Technology & Design, Singapore.

Contact Jer-Ming Chen at jerming_chen@sutd.edu.sg for more information about the project.

This brief was based on the project AFD 05/15 SL: Teaching and Learning English Pronunciation by Generating the Vocal Tract Shapes from the Frequency Domain Information.

How to cite this publication

Chen, J. M., & Herremans, D. (2019). *Teaching and Learning English Pronunciation by Generating the Vocal Tract Shapes from the Frequency Domain Information* (NIE Research Brief Series No. 19-008). Singapore: National Institute of Education.

Request for more details

Please approach the Office of Education Research, National Institute of Education, Singapore to obtain a copy of the final report.

>> More information about our research centres and publications can be found at: http://www.nie.edu.sg