Neiriz, R. (2019). The Phonetics 3D [Review]. In J. Levis, C. Nagle, & E. Todey (Eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Ames, IA, September 2018 (pp. 467-472). Ames, IA: Iowa State University.

TECHNOLOGY REVIEW

The Phonetics 3D

Reza Neiriz, Iowa State University

INTRODUCTION

Developed by Fuminori Homma and Yasushi Okumura and edited by Masashi Negishi, a professor in Graduate School of International Studies and Institute of Global Studies at Tokyo University of Foreign Studies, *The Phonetics 3D* is a mobile app designed to teach pronunciation to English language learners (ELL). The second version of this app, published in 2018, is available for only iOS users at $7.99. Below, I will first explain how the app works, and then I will present a critical review of the app.

How it works

The first run of the app launches a step-by-step tutorial of how to use it, which can be skipped and accessed later. The homepage has a navigation bar on the top of the screen dividing the content of the app into basic and advanced levels. Under the Basic tab, there is a list of phonemes written in IPA (Figure 1). These phonemes are grouped into stops, fricatives, nasals, laterals, semi-vowels, short vowels, long vowels, and diphthongs, and each group has a short description of its characteristics. The advanced tab contains instruction on phonological features of different phonemes in different environments (Figure 1). It is comprised of consonant clusters, devoicing, unreleased stops, nasal release, lateral release, assimilation, elision, reduction, and aspiration. Each section has a short description of the phonological feature and a list of buttons each representing the target phoneme(s).

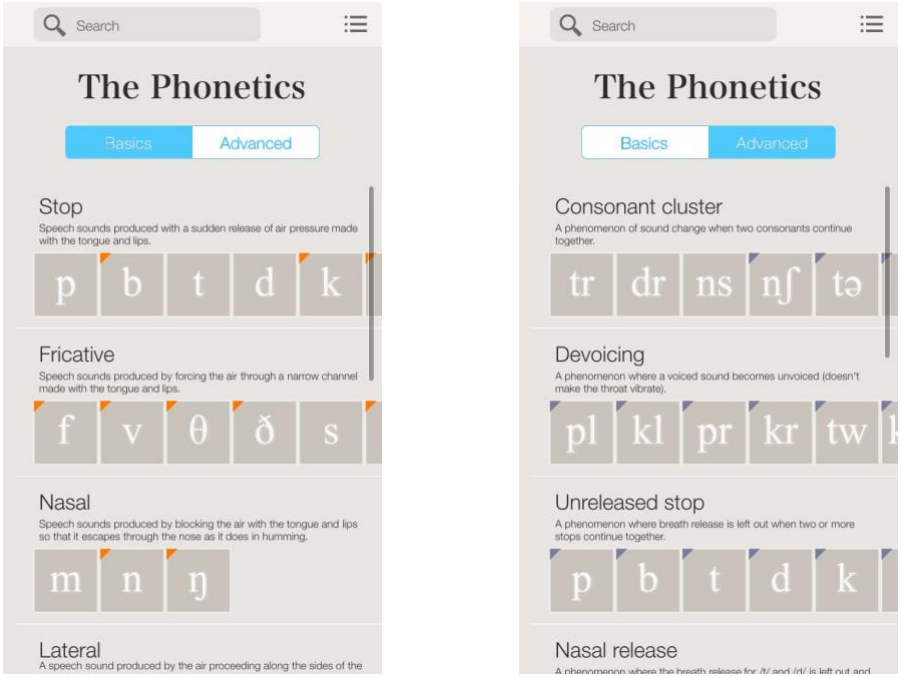

Figure 1. Basic (left) and advanced (right) menus.

Selecting each of the phonemes under the Basic tab takes the user to a new page which has an augmented reality head (Figure 2). This is simply a 3D image of a head visible from the tip of the nose to the throat. It is transparent and unicolor, but different tones of grey highlight the articulators. Users can swipe right or left to have a frontal or side view of this head. To the bottom-left of the head, there are two buttons representing the names of two native speakers whose voices could be used for pronunciation, i.e. Mike and Judy (Matthew K. Miller and Victoria Pate). Below the 3D image box, there are steps which explain articulation mechanisms accompanied by still image icons of side-view cross-section of a head in which blue areas represent the shape of the vocal tract at the respective stage of articulation. The 3D head on top of the screen reinforces this by showing an animated version of this explanation. When the user taps on the 3D image, the entire articulation is illustrated in animation, and tapping on each step illustrates the articulation mechanism of only that step. The animated head uses a white color for the airflow in voiceless phonemes and a blue color for their voiced counterparts. At the bottom of this page, there are two more tabs, one of which has example words containing the target phoneme both spelled and transcribed using IPA. Tapping on these words plays the word along with the 3D illustration of articulation through the transparent head. The last tab is similar to the second tab, but it contains words with phonemes that are similar to the target phoneme in terms of the place of articulation. For instance, for /m/, the software lists words with /p/ and /b/ as they are bilabial like /m/.

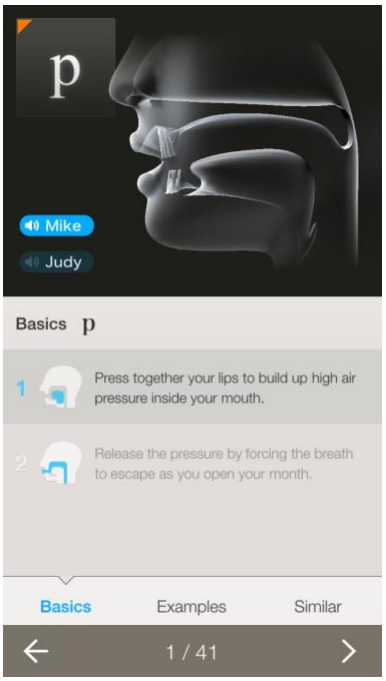

Figure 2. Instructions and modeling page of the Basic tab items.

The instruction page of the phonetic features under the Advanced tab is extremely similar to that of the phonemes (Figure 3). There are, however, two differences. First, there is an added written explanation of the phonological feature, such as how /t/ becomes more like /ʧ/ in *try*. Second, it contains only the examples tab, which lists words containing this feature. The user can tap on each word to both listen to and see the animated articulation of the word. Users can also choose to play the word slowly, and the app resynthesizes the original pronunciation sound file to produce a low-speed playback.

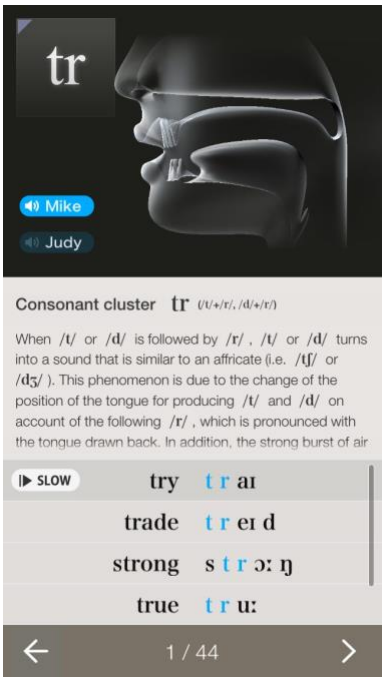

Figure 3. Instructions and modeling page of the Advanced tab items.

One last feature of the app is its search function. Users can type in any word and see both the IPA transcription and use the audio-visual representation of it through the head and either the voices of the speakers of the app or Siri if the word is not already recorded in the app (Figure 4).

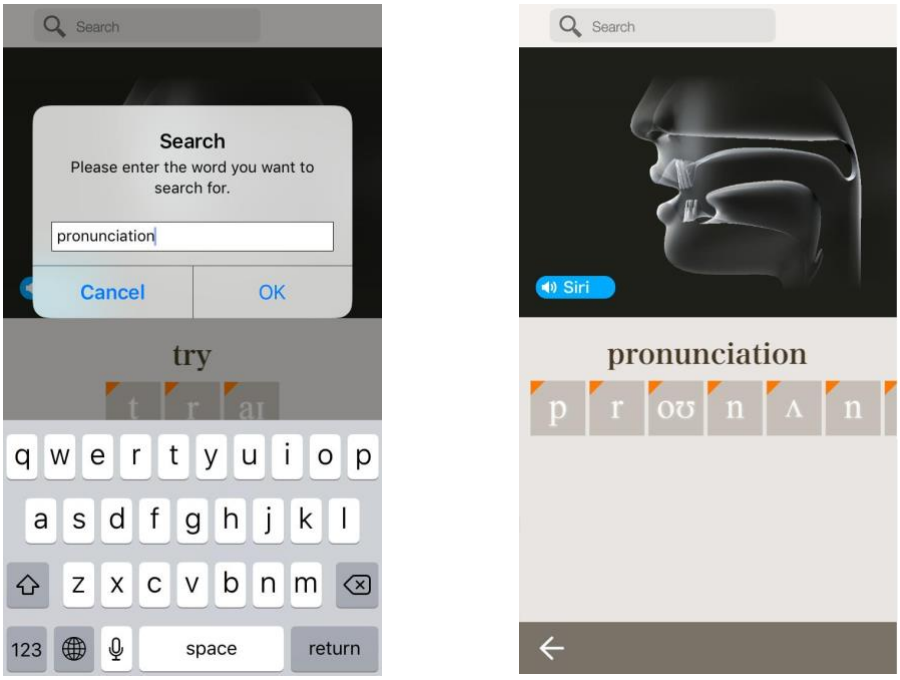

Figure 4. Word search page (left) and 3D articulation simulation of the search result (right).

CRITICAL REVIEW

According to The Phonetics 3D website, this app is "The most beautiful learning tool designed for those who are learning the pronunciations of the English language." While the aesthetics are admirably well-designed, the audience mentioned by the developers is too broad. The use of IPA phonetic symbols along with reference to articulators and articulatory mechanisms can certainly render this app less than useful for beginner and intermediate language learners, and even for advanced learners of English without formal training in phonetics and articulatory anatomy and mechanisms. Perhaps these features make this app more suitable for both native and non-native English speaking teachers. For the former, while most of the articulation involves tacit knowledge, this app can help them turn this knowledge into a procedural form so that they can transfer it to their students. As to non-native teachers, they can refine their pronunciation, both at declarative (knowing how to pronounce) and procedural (being able to actually perform the pronunciation) levels to provide a better role model and instruction of pronunciation for their students.

The goal of a 3D-animated representation of articulatory mechanisms through a transparent virtual head might have been to support auditory instruction. However, the literature is not conclusive in this regard. Similar virtual head representations to teach pronunciation have been used before such as ARTUR (Engwall & Balter, 2007), Baldi (Massaro & Light, 2003), and MASSY (Fagel & Madany, 2008) with inconclusive results. Massaro and Light (2003) compared the effect of using Baldi's face with that of another virtual head whose vocal tract and articulators were visible in improving Japanese speaker's pronunciation of /r/ and /l/. The results showed no difference. In another study, Massaro, Bigler, Chen, Perlman, and Oui (2008) used a virtual head with a visible vocal tract to teach Arabic consonants. The difference between virtual head users and the group presented with only auditory input was negligible. The results, however, have not always been negative. Massaro and Light (2004) found improvements in hearing-impaired American children through audio-visual training of consonant clusters, fricative-affricate, and voicing distinctions. Fagel and Madany (2008) also found improvements in children's lisping after using an augmented reality head to teach the articulation of /s/ and /z/. In short, Engwall (2012) argues that using audio-visual articulation training is not necessarily conducive to improvement in pronunciation if it is not accompanied by feedback. All of these studies show that using *The Phonetics 3D* might only be useful if it is used in combination with feedback from teachers who are familiar with phonetics and phonology. In other words, it can be a useful teaching aid in pronunciation classes where teachers can focus on only the problematic areas while all students spend time on the targeted practice of pronunciation instead of following classroom-wide drills which might not benefit everyone equally.

While *The Phonetics 3D* is a promising tool in aiding teaching pronunciation, it is lacking a major component of pronunciation, the suprasegmental features. Although this app has included phonological features, such as co-articulation, students interested in improving their pronunciation to approximate native-like speech could have benefitted from the inclusion of suprasegmental features in the app. These features have been shown to be strong indicators of accent. For instance, Van Els and de Bot (1987) found that even with low-pass filtered speech (retaining suprasegmental features) and monotonized speech (retaining segmental features), the participants could detect the foreign accent in speakers. Therefore, the inclusion of suprasegmental training material in *The Phonetics 3D* would be a welcome addition.

CONCLUSION

The Phonetics 3D comes with useful features, such as transparent and animated representation of vocal tract and the use of different colors for voiced and devoiced features, and it can be a useful pronunciation learning tool when accompanied by instruction and feedback. Yet, an addition of suprasegmental features to this app can better respond to the needs of users who wish to achieve a more native-like pronunciation.

REFERENCES

Engwall, O. (2012). Analysis of and feedback on phonetic features in pronunciation training with a virtual teacher. *Computer Assisted Language Learning, 25*(1), 37-64.

Engwall, O., & Balter, O. (2007). Pronunciation feedback from real and virtual language teachers. *Computer Assisted Language Learning, 20*, 235-262.

Fagel, S., & Madany, K. (2008). A 3-D virtual head as a tool for speech therapy for children. In International Speech Communication Association (Ed.), *9th annual conference of the International Speech Communication Association 2008 (Interspeech 2008)* (pp. 2643-2646). Red Hook, NY: Curran Associates, Inc.

Massaro, D. W., & Light, J. (2003). Read my tongue movements: Bimodal learning to perceive and produce non-native speech /r/ and /l/. *Eurospeech* (pp. 2249-2252). Brisbane, Australia: European Speech Communication Association (ESCA).

Massaro, D., & Light, J. (2004). Using visible speech for training perception and production of speech for hard of hearing individuals. *Journal of Speech, Language, and Hearing Research, 47*, 304-320.

Massaro, D., Bigler, S., Chen, T., Perlman, M., & Oui, S. (2008). Pronunciation training: The role of eye and ear. *Interspeech* (pp. 2623-2626). Geneva, Switzerland: International Speech Communication Association (ISCA).

van Els, T., & de Bot, K. (1987). The role of intonation in a foreign accent. *The Modern Langauge Journal, 71*(2), 174-155