NetProF iOS Pronunciation Feedback Demonstration[1]

Tamas Marius[2], Jennifer Melot[3], Gordon Vidaver[3]

DLI Foreign Language Center[2]
542 Rifle Range Road, Monterey, CA
Tamas.Marius@dliflc.edu

MIT Lincoln Laboratory[3]
244 Wood Street, Lexington, MA
Jennifer.Melot@ll.mit.edu, Gordon.Vidaver@ll.mit.edu

One of the greatest challenges for an adult learning a new language is gaining the ability to distinguish and produce foreign sounds. The US Government trains 3,600 enlisted soldiers a year at the Defense Language Institute Foreign Language Center (DLIFLC) in languages critical to national security, most of which are not widely studied in the U.S. Many students struggle to attain speaking fluency and proper pronunciation. Teaching pronunciation is a time-intensive task for teachers that requires them to give individual feedback to students during classroom hours. This limits the time teachers can spend imparting other information, and students may feel embarrassed or inhibited when they practice with their classmates. Given the demand for students educated in foreign languages and the limited number of qualified teachers in languages of interest, there is a growing need for computer-based tools students can use to practice and receive feedback at their own pace and schedule. Most existing tools are limited to listening to pre-recorded audio with limited or nonexistent support for pronunciation feedback. MIT Lincoln Laboratory has developed a new tool, Net Pronunciation Feedback (NetProF), to address these challenges and improve student pronunciation and general language fluency.

NetProF is an iOS App that allows students to practice classroom vocabulary in 11 languages. It has an easy-to-use, intuitive UI that allows the user to listen to native speaker recordings of vocabulary items and practice pronunciation. When a student says a vocabulary item, they receive scores of how well they're speaking at the level of the item overall, per word, and per phoneme within a word. This is made possible by Lincoln Laboratory's automatic speech recognition technology: using in-house audio collections, we have trained neural network-based phoneme recognition models for the low-resource languages NetProF supports. High fidelity custom dictionaries provide precise word pronunciations, and our fast, online, resident decoder provides scores and word and phoneme time alignments in typically less than two seconds. This responsiveness in a hand-held platform allows students to practice a language in a state of educational flow.

NetProF's vocabulary parallels DLIFLC's classroom content, grouping vocabulary items into units and chapters. Items are presented beginning with short, easy to pronounce words, continuing to longer, harder phrases as the chapter progresses. The audio content includes male and female reference recordings at normal and slow speeds. Each item is accompanied by an example sentence with item used in context. Score feedback is provided via intuitive color highlighting. The student can replay their speech and see a "follow the bouncing ball" dynamic display that highlights each word and phoneme as it plays. An overall summary display shows which phonemes the student has trouble pronouncing and examples of words where these phonemes appear. The student can also touch a word and compare their speech against a native speaker's pronunciation.

Early testing with students has shown measurable gains in their pronunciation scores. Our interface provides quick, interactive pronunciation feedback in the palm of your hand.

[1] This work is sponsored by the Defense Language Institute Foreign Language Center under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

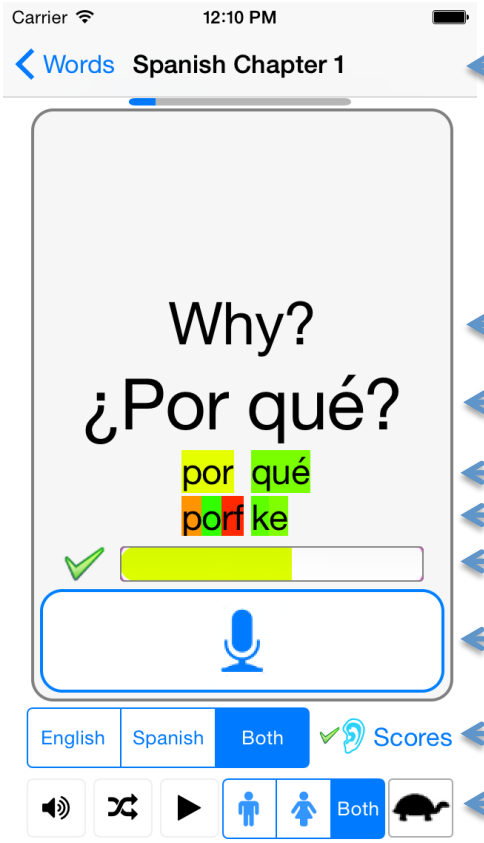

Figure 2: Studying a vocabulary item

Language and Chapter

English Translation, touch to hear

Vocabulary Item, touch to hear

Student speech, scored by color, touch to hear

Student speech phonemes, with scores

Correct/Incorrect and overall score

Press and hold to record yourself

Visibility controls and overall scores

Audio on/off, shuffle, auto advance, and gender and speed control

Figure 1: Summary display of phonemes

Across all items, lowest scoring phonemes first, in word context. Touch to hear student audio contrasted with reference.

In this case the student has trouble saying the Spanish rolled R (rf).