POLLI: a handheld-based aid for non-native student presentations

Elizabeth M Davis, Oscar Saz, Maxine Eskenazi

Carnegie Mellon University, Pittsburgh, PA. USA

emd1@andrew.cmu.edu, osaz@cs.cmu.edu, max@cmu.edu

Abstract

Language learning is finally leaving the classroom for the real world. This usually takes the form of a dedicated implementation on a handheld device. In this paper we describe POLLI, an app that helps non-native students prepare scientific presentations.

Index Terms: handheld applications, pronunciation skills

1. Introduction

In the past few decades, Computer-Aided Language Learning (CALL) has emerged, freeing students from limited time in the classroom to unlimited language exposure and learning opportunities. Students can learn at any time they want, spend as much time as they want and interact with educational material in a variety of ways. While in the past one course had to cover all aspects of the target language (L2), CALL software has gradually become specialized, teaching, for example, just pronunciation, or writing. Students learn what they immediately are in need of. CALL has also become more realistic. Often, rather than consisting of a set of exercises using expressions that the student is unlikely to use ("this is a table and that is a book"), CALL systems are starting to address real student needs. While this is great progress, they still fall short at being permanently present for a learner. Students learn best when they need some knowledge and are going to apply it in their everyday lives. A CALL system that requires a laptop or desktop cannot be out in the street with the student when they want to ask for directions or pay their hotel bill. An omnipresent system makes every new event a learning event.

Within the past three years, people have come to depend on mobile devices, such as tablets, personal data assistants, and smartphones, and ubiquitous internet access, in their everyday lives. These handheld devices, with touchscreens and voice-activated software, allow users to complete important day-to-day tasks from virtually anywhere. Moving away from the mouse-activated software, allow users to complete important day-to-day tasks from virtually anywhere. These handheld devices, with touchscreens and voice-activated software, allow users to complete important day-to-day tasks from virtually anywhere. A CALL system that requires a laptop or desktop cannot be out in the street with the student when they want to ask for directions or pay their hotel bill. An omnipresent system makes every new event a learning event.

2. Background

There are an increasing number of computer-based systems that can teach non-native students the correct pronunciation of an L2 (amongst the many, [6], [7], etc). While these systems give students a basic understanding of how to pronounce phones and give examples students can imitate, they only address the problem in the abstract. They do provide examples that are designed to exercise the new phonetic knowledge in a different contexts. But these contexts are not necessarily related to anything that the student would say in real life. Also, while they do teach many different contexts, most of them are places where a mispronunciation will not cause a misunderstanding. Students then go out in the real world and encounter situations where they are mostly understood, just as long as their pronunciation errors are not in minimal pair contexts situated in sentences where either member of the minimal pair could fit. In these important minimal pair contexts, where meanings can be very different when a pronunciation error occurs, either they get the wrong message across or they confuse the listener as to what they intended to say [8]. The latter case happens when the syntactic context is correct, but no semantic interpretation can be drawn from the utterance ("we take the author’s point" “we make the author’s point”).

In this paper, as in [9], we will call sentences with both types of potential for confusion “confusable contexts”. For example, the phrases “We will adopt your proposal” and “We will adapt your proposal” differ by only one phoneme. Both make sense, but have two distinctly different meanings. For speakers of languages where the phonemes /a/ and /ʌ/ either do not exist or are indistinguishable, it would not be difficult to mistakenly say the phonemes of one sentence while intending the other, and a native English speaker would not necessarily pick up on the mistake.

[12] implemented the BICC algorithm that automatically detects confusable contexts using several measures, including minimal pairs (gleaned from CMUDICT), and bigram part of speech models (Stanford POS tagger). The phonetisation version of words in a given sentence is compared to all of the entries in CMUDICT to find all possible minimal pairs. For each minimal pair where both words in the pair are the same part of speech, a new sentence is generated with the minimal pair word substituted
in it. The new sentence is then evaluated for plausibility by the POS tagger. If it passes this test, the original and the newly-generated sentences are then passed through the POS tagger and compared. If both sentences have the same POS string, and if the immediate context of the confusable word has similar POS tri-grams, and similar word tri-grams, then the sentences are considered to be confusable. This generates some more possibilities than are actually plausible and some heuristics can be used to further prune the number of possible confusable contexts. One way to prune the contexts is to show an individual only the confusions that they are likely to make, given their native language. Thus, a Japanese native speaker could have issues differentiating “we used crowd computing” from “we used cloud computing” in English while this would not be a problem for a native speaker of Spanish or of French. BICC therefore has representations of confusions between L1 and L2 for many L1s.

3. POLLI: The POcket Language Learning Interface

There are many ways that the BICC algorithm described above could be used in language learning systems. For example, a pronunciation trainer could have BICC generate the sentences that students are asked to pronounce and practice, thus replacing the abstract, useless ones used at present. We decided to concentrate on one specific application, keeping in mind that our goal is to address narrow needs, not the broad abstract ones of the past. We believe that in the future students will concur, choosing their learning software according to their immediate needs. Thus POLLI uses the BICC software in the narrowly-focused application of scientific presentations.

POLLI is also mobile, usable anywhere at any time it was needed. Thus for POLLI, since we chose to help non-native students who were preparing to give a talk at a scientific gathering, they could use it to prepare while on a plane, or sitting at a session of the conference. The goal of the system is to show speakers the confusable contexts that they might produce in their talks. For now, the student is left to decide what to do about these contexts – whether to get pronunciation training on how to say them well, or to avoid using them in the talk.

POLLI exists as an application for the Android platform using the BICC algorithm on a piece of English text. Students can either type the text in themselves (Figure 2) or load in a plain text document (Figure 1). They also can specify their native language in the app by selecting it from the list of supported languages, as shown on Figure 3.

The ability to choose to either type in what the person intends to say or to upload the scientific paper they are presenting gives the students more flexibility, depending on whether they create their presentations in a way that is very close to the text or if they author them on a more global semantic basis.
POLLI analyzes the text and displays the output as an expandable list of the original words that, if changed by one phoneme, can change the interpreted meaning of the sentences they appear in (or cause considerable confusion as to what the speaker intended). By selecting an element in the list, the speaker can see the sentences that the words appeared in, as well as the word(s) that listeners might infer from the context if the word was mispronounced. (Figure 4).

Figure 4. POLLI presents the text analysis results

![Figure 4](image)

4. Assessing user acceptance of POLLI

First POLLI was tested for correct functionality: that documents could be uploaded and analyzed, that screen-input text could be analyzed, that correct minimal pairs were found and displayed and that plausible contexts were also found and displayed. With a fully working app, we proceeded to assess its usefulness for foreign undergraduate and graduate students. User tests were carried out with a group of 24 non-native English speakers, students recruited from within Carnegie Mellon University in the Language Technologies Institute and the Computer Science Department. They had many different native languages: Mandarin Chinese, Spanish, Japanese, Russian, Thai, Tamil, Marathi, Portuguese, Greek, and French. Before coming to try POLLI, each subject was asked to submit the text of a familiar piece of writing that they felt comfortable talking about. Upon arrival, subjects recorded a practice presentation summarizing the text they had submitted. After the practice session, they tried POLLI on an Android Nexus 7, opening the app, loading their text, and submitting it for analysis. Subjects were then given ten minutes to review POLLI’s output and prepare another summary presentation. That new presentation was then recorded. During the ten minutes with POLLI, they were free to do whatever they wanted with its output, for example, making notes on their text or asking questions about how to properly pronounce the words POLLI had flagged for them. After they recorded their second presentation, they were given an exit survey. They were asked to say how strongly they agreed or disagreed with a series of statements about the application, their answers ranging from 1 to 5 (Likert scale), where 1 meant strongly disagree and 5 meant strongly agree. The following are a subset of the questions that were asked.

Statements on effectiveness of the feedback were:
I learned something from this app
The information I got from the app was useful

Statements on the usability of the app were:
The system behaved as expected
It was easy for me to load a file into the app
I found the app particularly easy to use.

Statements on the app’s potential for real world use:
There are features I would like to see added to this app that would make me more likely to use it
If the output were presented in a more helpful way, I would be more likely to use this app
I would use this app again if it were available as a paid app

All of the subjects except two were successfully able to run the BICC analysis for their own native language. At the time of this study, phonetic representations of Greek and Russian were not supported for analysis. Thus the Greek speaker chose to use the analysis for Spanish speakers, believing the accent and pronunciation errors to be similar between the two languages. The Russian speaker also chose to use the analysis for Spanish speakers, and answered the questionnaire as if Spanish had been his native language.

5. Results

This section shows subjects' responses to the questionnaire for:
- effectiveness of the feedback
- the usability of the app
- potential for real-world use.

5.1 Effectiveness of the feedback

The results shown in Figures 5, and 6 confirm that the subjects overall found the app to be useful with most answers in the 4-5 range (strongly agreeing with positive statements about the app). We did note that not all inappropriate contexts had been filtered out.

We also gathered verbal comments from the subjects. Several subjects remarked that the “trouble phones” they saw were indeed sounds they have trouble with. Some actually thought that the app had used speech recognition to detect that they had pronounced these phones incorrectly (it did not). Some of the words that the app showed to the users had not been used in their summaries. Subjects seemed to be paying more attention to their speech in the summary they recorded after using POLLI. Some subjects tried to use the flagged words to show that they were paying attention.
5.2 Usability of the app

The results shown in Figures 7, 8 and 9 for usability continue to be positive. The system only crashed the on two texts at the beginning of the study and was always able to give some feedback. Subjects found that the system met their expectations, which not only reflects on the quality of system function, but on the way we chose to present it. The ease of use expressed in Figures 8 and 9 certainly contribute to the overall positive impression.

5.3 Potential for real use

Figures 10, 11 and 12 reflect what subjects think of POLLI’s potential for real use. Subjects were asked if they could think of add-ons that they would like to see in POLLI that would make it more useful. Many of them told us that they would like to be able to speak to it and to get pronunciation feedback. They seemed to appreciate the way BICC’s output is presented on the screen. Some believed, however, that the output could be enhanced with the use of either recorded speech or speech synthesis so that they could hear what a specific context should sound like. Although the subjects did not mention it, they would probably also benefit from hearing their own recordings. Subjects would have liked for the confusable contexts to be ranked in some way (order of appearance, difficulty, probability of making that specific mistake, etc.) and some would have liked to see more text around the confusable contexts.

Presenting subjects with the choice of whether POLLI should be free or paid seemed to be another appropriate way to gauge acceptance. With many apps being offered for free, it was reassuring to find that subjects stated that they would be willing to pay around $1 or $2 for POLLI.
6. Conclusions

We have described POLLI, an Android app that helps non-native students preparing presentations. POLLI embodies both the everpresence that is desirable in language learning and the focus that we believe reflects what students will most want in the future. While we cannot yet conclude that this presentation format affords more learning than a regular laptop interface (and we will want to test this hypothesis in the future), we do observe that the subjects who have tried POLLI have positive feedback and seem very motivated to use it. Since motivation has been shown to have a positive effect on learning [10], we believe that POLLI can make a positive impact on learning.

Acknowledgements

The first author is sponsored by the Semiconductor Research Corporation as part of their Undergraduate Research Opportunities program.

References