Rapidly Testing the Interaction Model of a Pronunciation Training System via Wizard-of-Oz

João P. Cabral\(^1\), Mark Kane\(^1\), Zeehan Ahmed\(^1\), Mohamed Abou-Zleikha\(^1\), Éva Székely\(^3\), Amalia Zahra\(^1\)

1School of Computer Science and Informatics, University College Dublin, Ireland.
2School of Computer Science and Statistics, Trinity College Dublin, Ireland.

Aim: Rapidly test and improve an early-stage pronunciation learning system called MySpeech by:
- Testing of human-computer interaction (HCI) model and feedback messages sent to the user.
- Evaluation of overall user satisfaction with the pronunciation learning service.
- Collection of data that can be used to improve the HCI and to personalise the system to the user.

Method: Combine the MySpeech system and the WebWOZ platform (http://www.webwoz.com):

- Web-based interface of MySpeech
- Wizard-of-Oz interface
- Real-time video
- Detected mispronunciations
- MySpeech System
- Input data (user's speech, sentence, etc.)
- Feedback messages

**The MySpeech System**

- Web-based interface of MySpeech
- Detected mispronunciations
- Pronunciation Analysis
  - HMM-based speech recognition
  - Broad phonetic groups to introduce difficulty levels
- Content text
- Reference speech
- Selected content (language, difficulty level, sentence ID)
- User's speech
- Student database

**Wizard-of-Oz platform**

- Platform that permits:
  - Wizard to select predefined sentences grouped into different panels for instructions and feedback.
  - Wizard to type new text or edit predefined sentences.
  - Store all messages sent by the wizard in a database.

**Experiment**

- Test interaction model to guide student in pronunciation training
- Interaction model is an iterative process:
  1. (a) Practice 2 levels: “easy” and “medium”
  2. (b) Select max of 2 different categories
  3. (c) Repeat max of 2 sentences in each category
  4. (d) Repeat same sentence max of 3 times
- Model was simulated by the Wizard:
  - WebWOZ and model explained to the wizard.
  - Sent instructions using the WebWOZ.
  - Interpreted results of pronunciation analysis and sent feedback using the WebWOZ.
- Wizards: 6 members of team who developed the MySpeech system
- Participants: postgrad students (non-native English speakers)

At the end of the experiment, students completed a questionnaire to evaluate the MySpeech performance for English pronunciation training

**Results**

Quantitative analysis of messages used by the wizards:

<table>
<thead>
<tr>
<th>Message Type (panels)</th>
<th>Predefined</th>
<th>Total sent</th>
<th>Rate of New</th>
<th>Rate of Edited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructions</td>
<td>23</td>
<td>276</td>
<td>7.2%</td>
<td>0%</td>
</tr>
<tr>
<td>Corrective feedback</td>
<td>12</td>
<td>125</td>
<td>8.8%</td>
<td>88%</td>
</tr>
<tr>
<td>Positive feedback</td>
<td>8</td>
<td>111</td>
<td>1.8%</td>
<td>0%</td>
</tr>
</tbody>
</table>

> More messages than positive feedback
> More editing by the wizard

Low rates indicate that predefined messages were appropriate.

**Questionnaire about user’s satisfaction evaluation:**

- Instructions were rated good on average.
- Most participants would use MySpeech again (at least 70%).
- Feedback was not always helpful to understand and correct mispronunciations (80% of participants answered “sometimes” and 20% “many times”).
- All participants rated the service as good (80%) or very good (20%).
- Only 20% of participants did not had fun using MySpeech.

**Conclusions and Future Work**

- MySpeech provided good service for non-native speakers to train English pron.
- Test of interaction model showed it performed well
- Plan to use HCI model to automatically generate instructions and feedback
- Data obtained from the wizard will be used to obtain a better set of messages
- In the future other modalities will be developed to show more details about pronunciation errors and suggest ways for correcting them, such as talking head.

This research is supported by the Science Foundation Ireland (Grant 07/CE/1142) as part of the Centre for Next Generation Localisation (www.cnlg.ie)