Speech Shadowing Method in Language Distant Learning

Voice waveform measurement method

LEE Carson
Ms in Information Science
Hasegawa Lab
A Little Background About Learning Language (Spoken)

- Do you have a young child in your family, a younger sibling, or cousin?
- Take a look at the age where they start to learn more complex speech (Not single words or simple, short sentences)
- If we take a look at how children learn more to speak more complex words and sentences, we can see that a child mimics, or parrots his/her parents’ speech.
Reasons for this research

• There are many software for improving a user’s language, but lack the focus on speech production
• Software for improving speaking skills usually take conversation/memorising route
• Subjects who are able to perform speech shadowing are shown to be able to comprehend and interpret the speech they produce
Research Questions

1. How to improve a user’s spoken language in accordance to a set standard
2. How to detect / analyse what the user said
3. How to deal with different types of voices
   - i.e. An oldman’s hoarse voice vs a little girl’s high pitch voice
4. How to support users in the long term
Research Goal

To develop a software system that is able to capture and analyse a user’s speech and provide input as to where can the user improve his/her speaking skills. The system would also cater to different accents and the primary target language would be English.
Proposed System

1. User would put on a headphone to listen to the speech/conversation playback
2. User then attempt to shadow the speech/conversation as clearly as possible (~0.25s slower)[1]
3. A microphone would be used to pick up and record the user’s voice
4. The captured audio file would be analysed and compared against the original recording (comparison of spectrographs)[2]
5. Data from the analysis would then provide feedback (i.e. score) to the user regarding how close is his/her speech to the original.
6. Inputs to improve performance could also be given

- System would return a score to indicate the learner’s performance in comparison to the native speaker.
- Parts where the learner did below average will be highlighted for repeated practice.
Notes

• System would run on mobile platforms (cloud computing would be used for heavy processing)
• A transcript will be used to provide meaning and clarity to the user
• The user should not attempt to memorise the transcript / recording
• Transcript should be unavailable during speech shadowing
Recap

**Why this method**
- Existing system lacks focus on uniformity
- Lack of speech learning system

**What to learn**
- Pronunciation
- Intonation
- Tempo
- Able to speak clearly and fluently as a native speaker
- Language comprehension improvement

**Goals**
- Speech Shadowing DLS

**Speech Shadowing**
- Learner can comprehend shadowed speech
- Closely repeating after speaker

**Technology**
- Mobile platform
- Cloud computing

**Prerequisites**
- Basic speaking skills
- Good vocabulary
- Good listening skills