Fluidity: Real-time feedback for speaking fluency development

Ralph ROSE
Center for English Language Education (CELESE)
Faculty of Science and Engineering
Waseda University

Acknowledgments
Hiroaki Suzuki, Junichi Inagaki, Masayuki Motoori, Yukikatsu Fukuda, Tatsuhiro Nomaguchi, Aiko Ooe, Maiko Serizawa

This research is partially supported by a Waseda University Grant for Special Research Projects (#2011B-152) and a Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) Grant-in-Aid (#24520661)
Overview

• Background
  – Automated assessment and feedback
  – Fluency in L2 speech
• Fluidity overview
• Usability evaluation
• Summary
Automated assessment of L2 speech

• Pronunciation (with visual feedback*)
  – Segmental: Cucchiarini et al 2009; Patten and Edmonds 2013*

• Fluency
  – ETS SpeechRater (Zechner et al 2009)
  – Versant (Pearson, Ordinate; Bernstein 1999)
  – CASEC (Hayashi et al 2004)

Useful overviews: Eskenazi 2009; Gamper and Knapp 2010
Feedback to learner

• Eskenazi (1999) - “Learners must receive pertinent corrective feedback”
• Most systems provide rapid feedback.
• In human-human communication, some feedback is in real-time
  – Back-channeling (uh-huh)
  – Head movements (nodding, shaking)
  – Facial expressions
• Is it possible to provide real-time feedback on fluency-related matters in human-computer interaction?
Fluency in L2 speech

• Scope
  – Broad: speak a language proficiently
  – Narrow: speak smoothly with minimal but natural hesitation

• Segalowitz (2010): levels of fluency
  – Cognitive fluency: ease of mental preparation
  – Utterance fluency: smoothness of articulation
  – Perceptual fluency: hearer's view of smoothness

Focus of the present work
Fluidity in L2 speech

- Crosslinguistic Corpus of Hesitation Phenomena (Rose 2013: http://filledpause.com/chp/cchp)
  - Speech recordings of Japanese (L1) and English (L2)
  - Parallel elicitation tasks: read/spontaneous speech
  - Intra-speaker comparisons possible

- Fluency factors that correlate with L2 proficiency
  - Speech rate (Correlated with L1 speech)
  - Silent pause rate
  - Silent pause duration (Independent of L1 speech)
  - Filled pause rate
Fluidity: fundamental aims

- Measure various utterance fluency characteristics and update them in real-time.
- Provide real-time feedback to learner on utterance fluency measures.
- Provide opportunity for learner to review their production together with visual representation of fluency measures.
- Provide feedback in a manner that emulates human-human communication.

A work in progress!
Fluidity: fluency measures

- Phonation time
- Silence time
- Syllable count: energy peaks
  (cf., Bhat et al 2010)
- Silent pause count: silence > 300ms
  (cf., De Jong and Bosker 2013)
- Filled pause count: stable formants and pitch
  (cf., Audhkhasi et al 2009)
Fluidity: Real-time fluency feedback

R. Rose, Waseda Univ.

Fluidity: main window

Requires Java SE 6 or greater

Detection settings

Audio input settings

Audio meter

Fluency measure indicators

Requires Java SE 6 or greater

Detection settings

Audio input settings

Uses TarsosDSP (Joren Six) and AudioInfo.java (Jonathan Simon) libraries
Fluidity: playback window

- Playback controls
- Waveform representation
- Fluency visualization
  - Filled pauses
  - Speech
  - Silent pauses
Fluidity: usability testing

- Participants (n=14, so far)
- Procedure
  - Practice speaking with Fluidity.
  - Adjust settings to fit their production.
  - Respond to survey questions about the experience.
- Still in progress...
Fluidity: user response

- Fluidity was easy to use.
- I enjoyed using Fluidity.
- The indicators gave accurate measurements without any adjustment.

- I could adjust my speech based on feedback from the indicators.
- Fluidity made me think about how I can improve my English fluency.
- Fluidity would make a good commercial application.
Fluidity: desirable features

- Capability to save recordings
- Capability to import sample recordings of native speakers
- Background noise through headphones to simulate different environments
- Animated face/head that blinks its eyes in a realistic manner
- Animated face/head that nods in a realistic manner
- Animated face/head that interrupts me if I've been silent too long
Fluidity: technical considerations

- Silence threshold often needs to be adjusted.
  - Solution: Do a better job of auto-detecting the sound configuration and microphone settings.

- Filled pause detection is difficult. Even after adjusting sensitivity, many participants did not see their filled pauses detected accurately (or at all).
  - Solution: Try other algorithms for detecting stable formants and pitch.
Fluidity: user comments

- 「語学を専攻していましたが、発音や文法にとらわれることが多く、流暢さを考えることがあまりなかったので、勉強になりました。」
  - Although I majored in languages, I have mostly studied about pronunciation and grammar and have not studied much about fluency. So, this was very educational.

- 「具体的にどうすれば良いかは分かりませんが、このFluidityを基板としたゲーム形式のアプリを使えば、すごく楽しく使えるかと思います。」
  - I wasn't really sure how to make use of Fluidity objectively, but if I could use it like a game application, I think it would be very enjoyable to use.
Summary

- Fluidity is designed to give real-time feedback to L2 English speakers on their utterance fluency.
- The application is still under development, but is capable of providing real-time feedback on most fluency measures, plus visualizations for review.
- Users find the application interesting, fun, and they are motivated to think about how to improve fluency.
- Users noted that the interface is difficult to use, and filled pauses are not accurately detected.
- Future development will include improving accuracy of fluency measurements and creating a user interface that better matches human-human interaction.
References


