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*
  - Supra-segmental: Anderson-Hsieh 1992*; Taniguchi and Abberton 1999*; de Wet et al 2009
- 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

Fluency 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.

Current focus of development

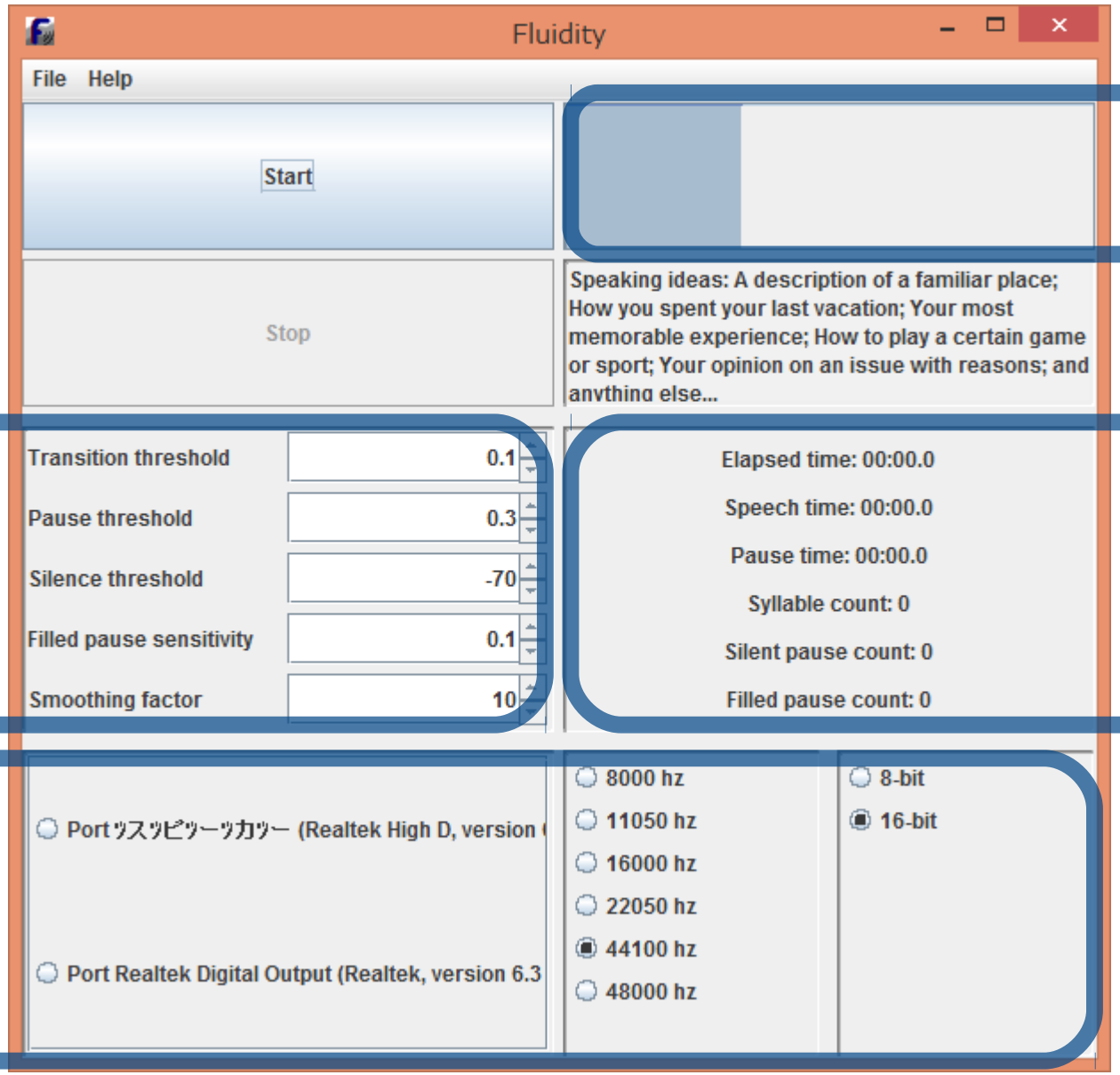
A work in progress!

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: main window

Requires Java SE 6 or greater



audio level meter

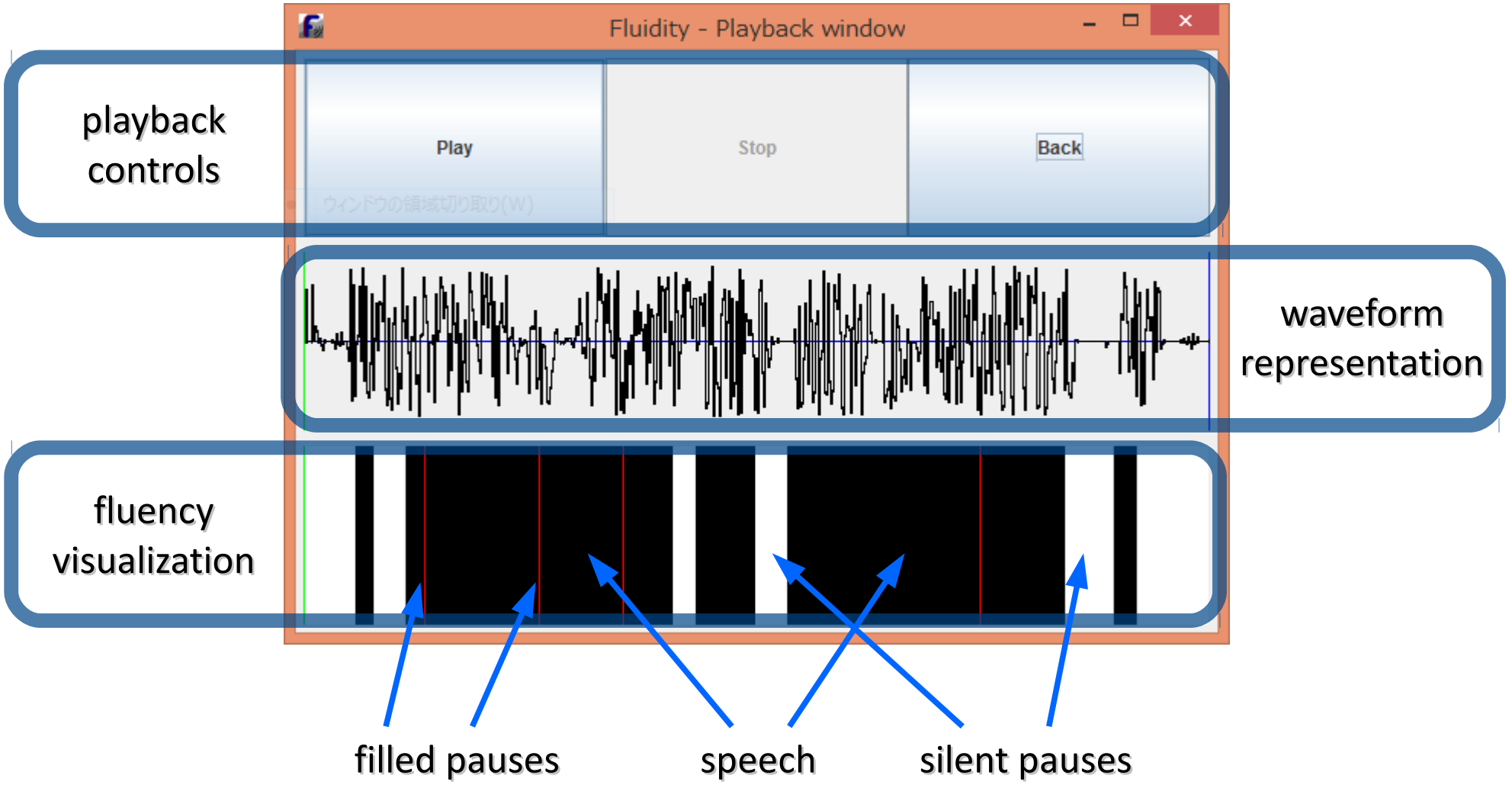
detection settings

fluency measure indicators

audio input settings

Uses TarsosDSP (Joren Six) and AudioInfo.java (Jonathan Simon) libraries

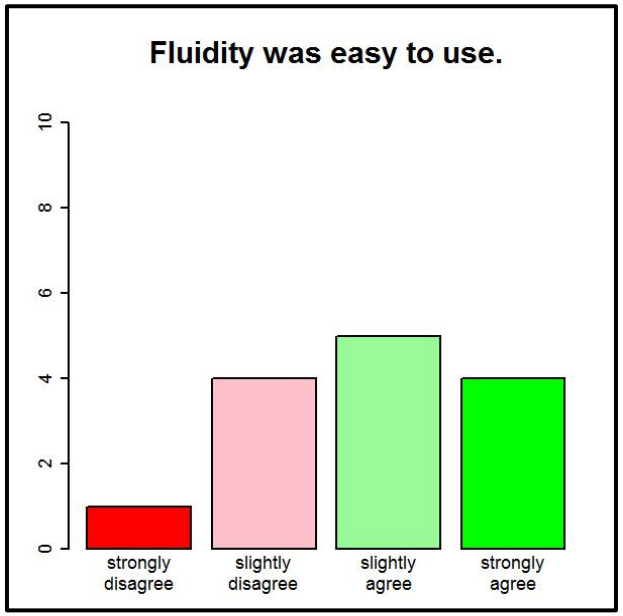
Fluidity: playback window

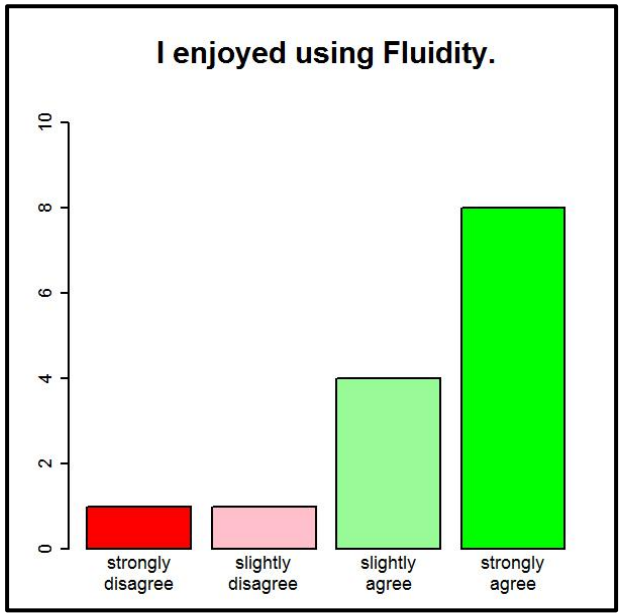


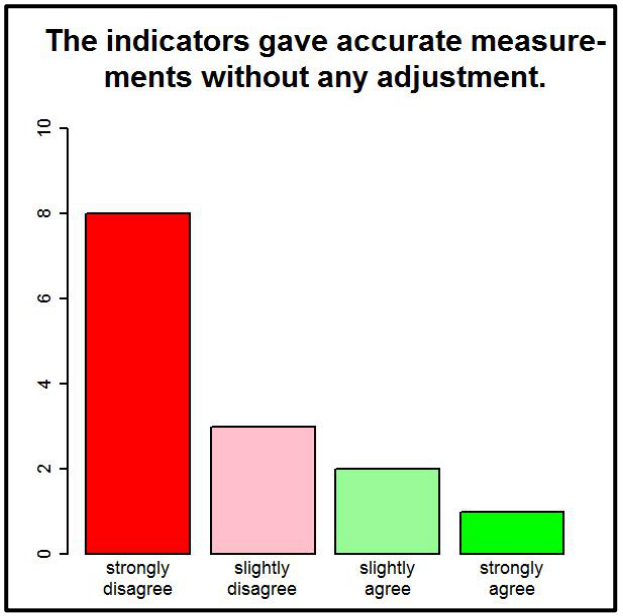
Fluidity: usability testing

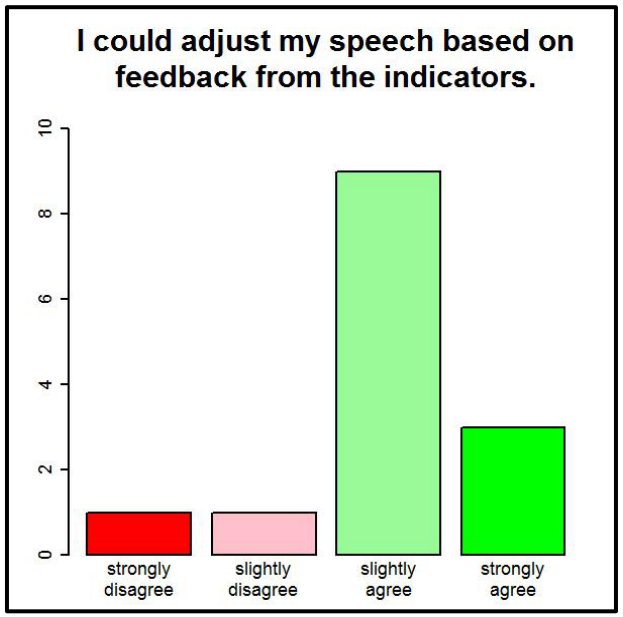
- Participants (n=14, so far)
- Procedure
  - Practice speaking with Fluidity.
  - Adjust settings to fit their production.
  - Respond to survey questions about the experience.
- Still in progress...

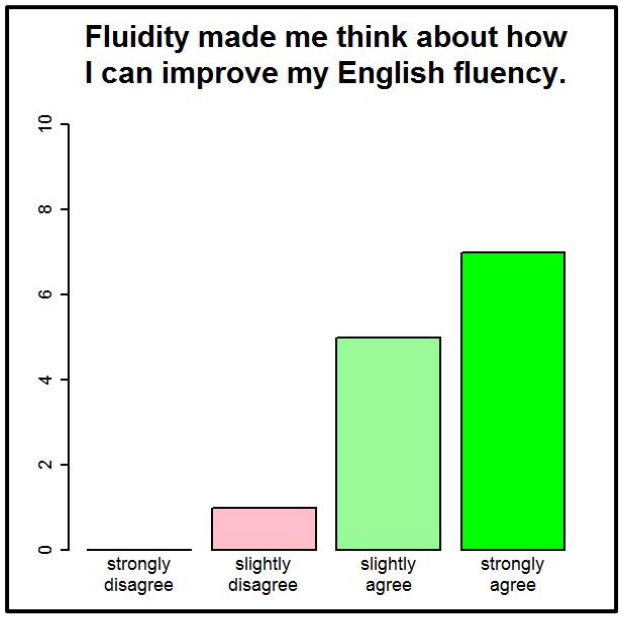
Fluidity: user response

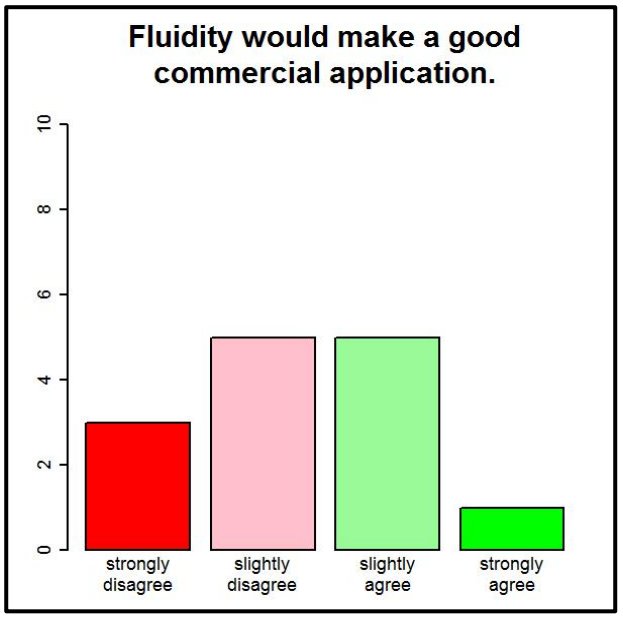




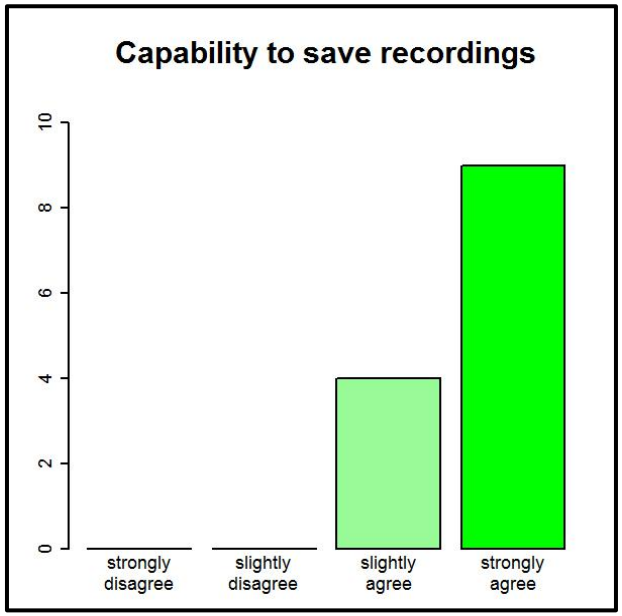


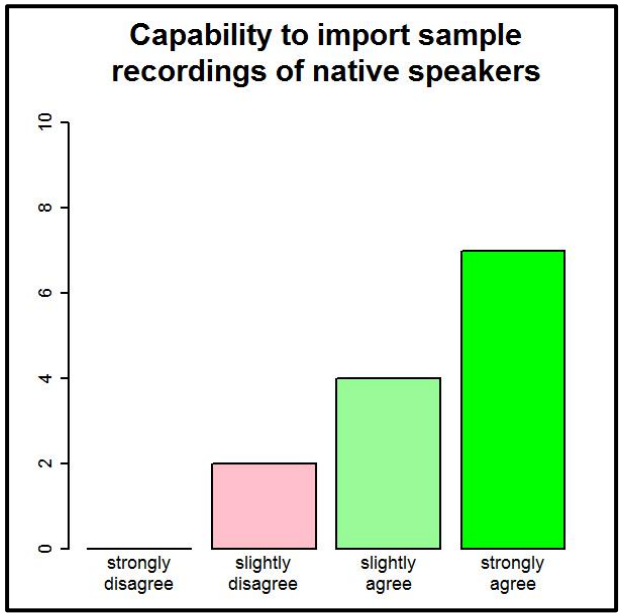


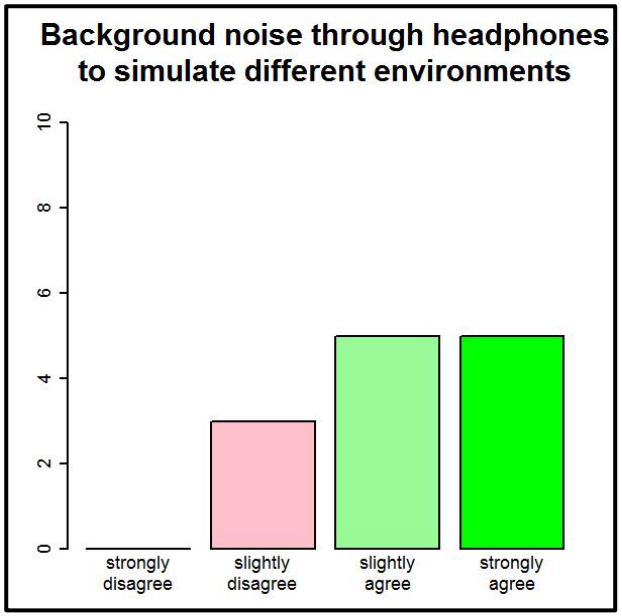


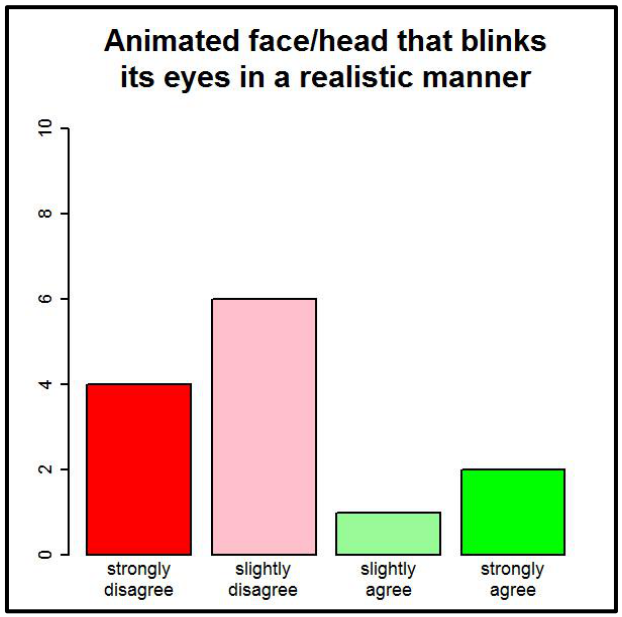


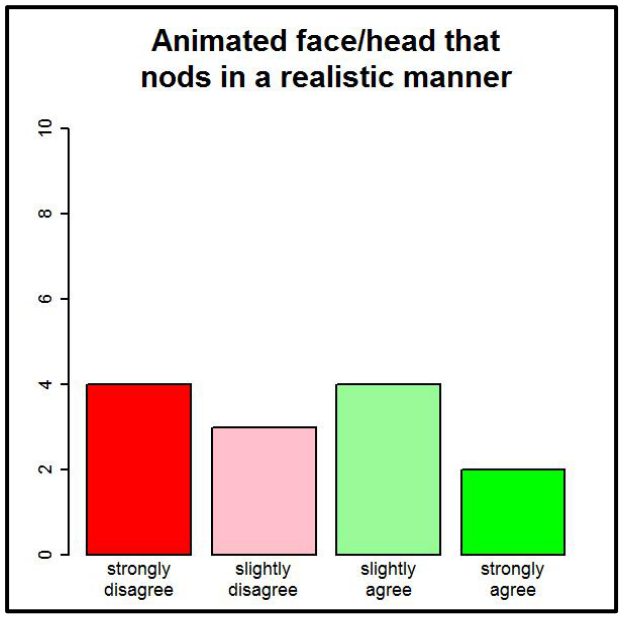
Fluidity: desirable features

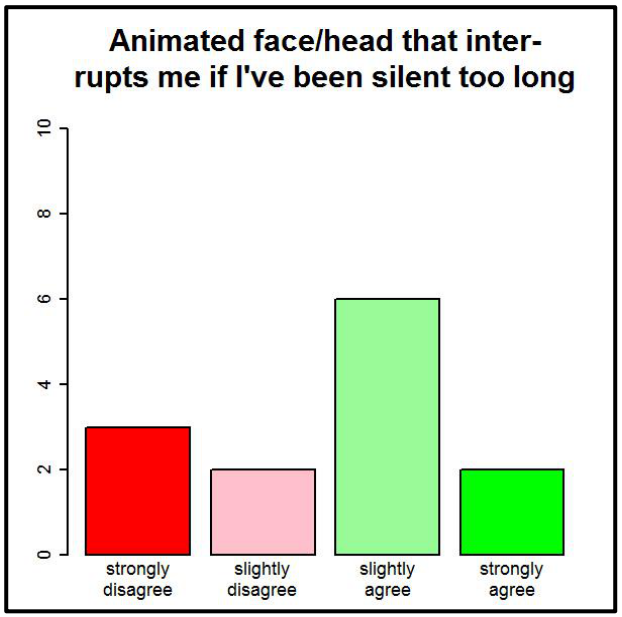












Fluidity: technical considerations

- 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れば，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

Anderson-Hsieh, J. Using electronic visual feedback to teach suprasegmentals. *System*, Vol. 20, pp. 51-62, 1992.

Audhkhasi, K.; Kandhway, K.; Deshmukh, O.D.; Verma, A., (2009) Formant-based technique for automatic filled-pause detection in spontaneous spoken English. *IEEE International Conference on Acoustics, Speech and Signal Processing, 2009*, pp. 19-24 April 2009.

Bernstein J (1999). PhonePass™ testing: Structure and construct. Menlo Park, CA: Ordinate.

Bhat, S., Hasegawa-Johnson, M., Sproat, R. Automatic Fluency Assessment by Signal-Level Measurement of Spontaneous Speech. *2010 INTERSPEECH Satellite Workshop on Second Language Studies: Acquisition, Learning, Education and Technology*, Makuhari, Japan, 2010.

Cucchiarini, C., Neri, A., Strik, H. (2009) Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback, *Speech Communication*, 51(10):853-863.

Eskénazi, M. (1999) Using Automatic Speech Processing for Foreign Language Pronunciation Tutoring: Some Issues and a Prototype. *Language Learning & Technology* 2(2):62-76.

Hayashi, N., Nogami, Y., Maeda, K., and Ikeda, H. (2004) Practical use of computerized adaptive testing in Japan -Development and operation of “CASEC”. ICP-2004 at Beijing, China. August 13th.

De Jong, N.H., Bosker, h.R. (2013) Choosing a threshold for silent pauses to measure second language fluency. *The 6th Workshop on Disfluency in Spontaneous Speech* (DiSS), pp. 17-20.

Iomi Patten and Lisa A. Edmonds (2013) Effect of training Japanese L1 speakers in the production of American English /r/ using spectrographic visual feedback. *Computer Assisted Language Learning*.

Rose, R.L. (2013) Crosslinguistic Corpus of Hesitation Phenomena: A corpus for investigating first and second language speech performance. In: *Proceedings of the 14th Annual Conference of the International Speech Communication Association* (Interspeech 2013), pp. 992-996.

Segalowitz, N. (2010) *Cognitive bases of second language fluency*. Routledge.

Simon, J. AudioInfo.java, audio info library. Retrieved from http://www.javafaq.nu/java-example-code-479.html on 25 December 2014.

Six, J. TarsosDSP, Java library. Retrieved from https://github.com/JorenSix/TarsosDSP on 25 December 2014.

Taniguchi, M. and Abberton, E. (1999) Effect of Interactive Visual Feedback on the Improvement of English Intonation of Japanese EFL Learners. *Speech, Hearing and Language: Work in Progress* No. 11 (pp.76-89) Department of Phonetics and Linguistics, University College London.

de Wet, F., Van der Walt, C., and Niesler, T.R. (2009) Automatic assessment of oral language proficiency and listening comprehension, *Speech Communication*, 51(10):864-874.

Zechner, K., Higgins, D., Xi, X., and Williamson, D.M. (2009) Automatic scoring of non-native spontaneous speech in tests of spoken English, *Speech Communication*, 51(10)883-895.