How can we determine, detect, and develop fluency in second language speech?

Ralph L. Rose

<rose@waseda.jp>

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

Introduction

Well, in my own life I'd break it up in stages, when I had a difficult youth. My father wasn't in the house. I've written about this. You know there were times where I've experimented with drugs, and I drank in my teenage years. And what I trace this to is a certain selfishness on my part. I was so obsessed with me, and the reasons that I might be dissatisfied, that I couldn't focus on other people. And I think the process for me of growing up was to recognize that it's not about me, it's about ...

Absolutely. But look, you know, when I find myself taking the wrong step, I think a lot of the times it's because I'm trying to protect myself, instead of trying to do God's work. And so that I think is my own failure

Barack Obama (August, 2008); Saddleback Presidential Forum

Introduction

Well, i- i- i- in my own life I'd- I'd break it up in stages, when uh I had a difficult youth. Uh my father wasn't in the house, uh I've written about this, uh there- uh uh you know there were times where uh I've experimented with drugs, and I drank, uh yeah in my teenage years, a-nd wh- what I trace this to is uh a certain selfishness on my part, I- I was so obsessed with me, and you know the- the reasons that I might be dissatisfied, that I- I- w- I couldn't focus on other people. And uh y- you know I think the process for me of growing up was to recognize that it's not about me, it's about …

it's about- absolutely, so- so- but- but look, you know, th- the uh wh- when I uh wh- when I find myself um taking the wrong step, I think a lot of the times it's because I'm trying to protect myself, instead of trying to do God's work. And- and- an- and so that I think is- is my own failure

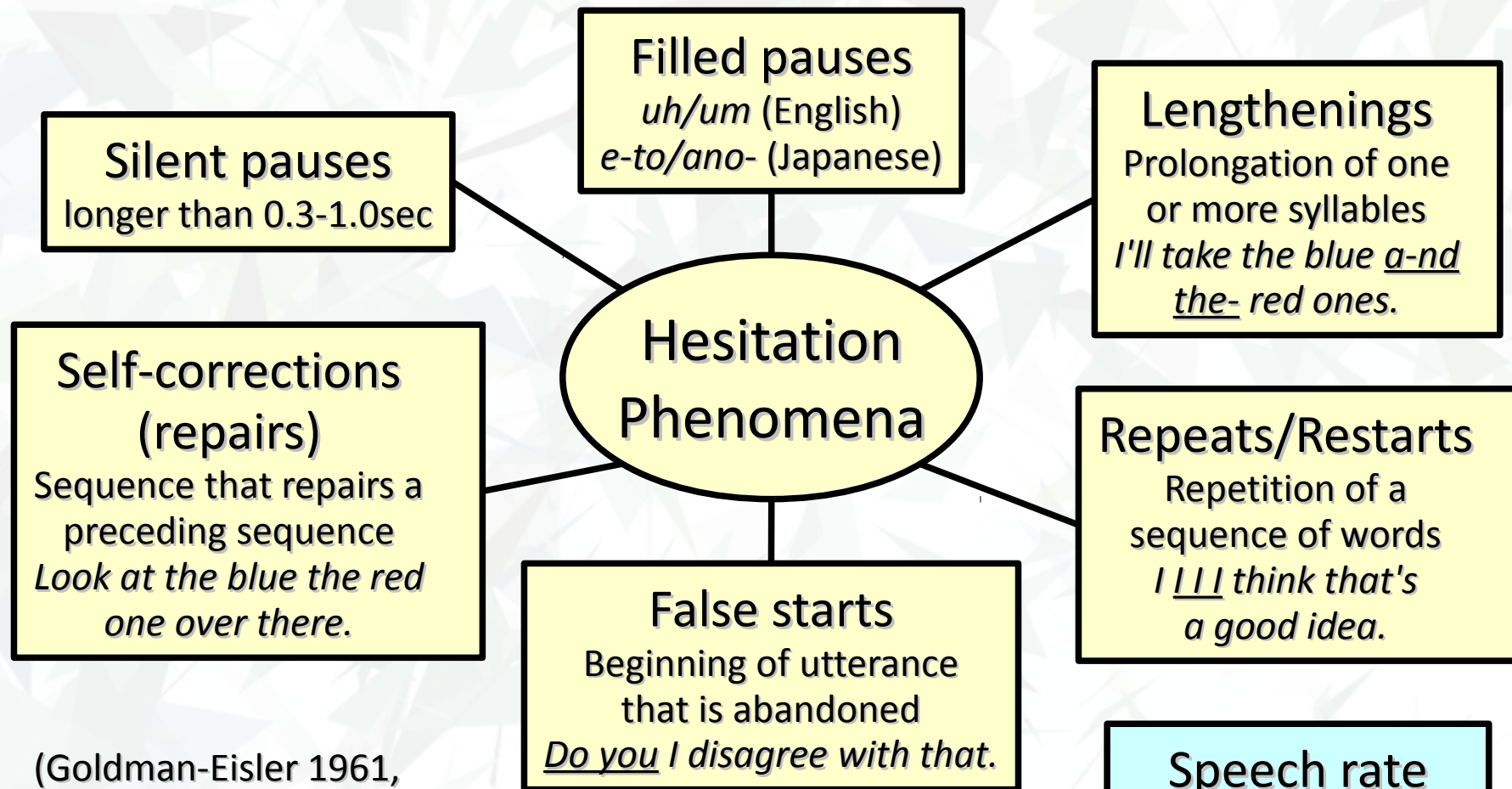Barack Obama (August, 2008); Saddleback Presidential Forum

Overview

- Determining L2 fluency
  - Hesitation phenomena in speech
  - Crosslinguistic Corpus of Hesitation Phenomena
  - Hearer perception of fluency
- Detecting and developing L2 fluency
  - Feedback in L2 development
  - Fluidity application
  - ...and complexity, too

Fluency

- Scope of fluency
  - 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
- Many researchers looking at cognitive-utterance dimension
  - De Jong et al (2012) – articulation rate strongly related to fluency
  - Don and Lickley (2015) – relationship between working memory and fluency

Observations of fluency

Filled pauses
uh/um (English)
e-to/ano- (Japanese)

Lengthenings
Prolongation of one
or more syllables
*I'll take the blue <u>a-nd</u>
<u>the-</u> red ones.*

Silent pauses
longer than 0.3-1.0sec

Hesitation Phenomena

Self-corrections (repairs)
Sequence that repairs a
preceding sequence
*Look at the blue the red
one over there.*

Repeats/Restarts
Repetition of a
sequence of words
*I <u>I I I</u> think that's
a good idea.*

False starts
Beginning of utterance
that is abandoned
<u>Do you</u> I disagree with that.

Speech rate
by word, by syllable,
with/without pauses

(Cucchiarini et al 2010)

(Goldman-Eisler 1961,
Levelt 1983, 1989,
Maclay and Osgood 1959,
Rochester 1973, inter alia)

Hesitation phenomena in L2 production
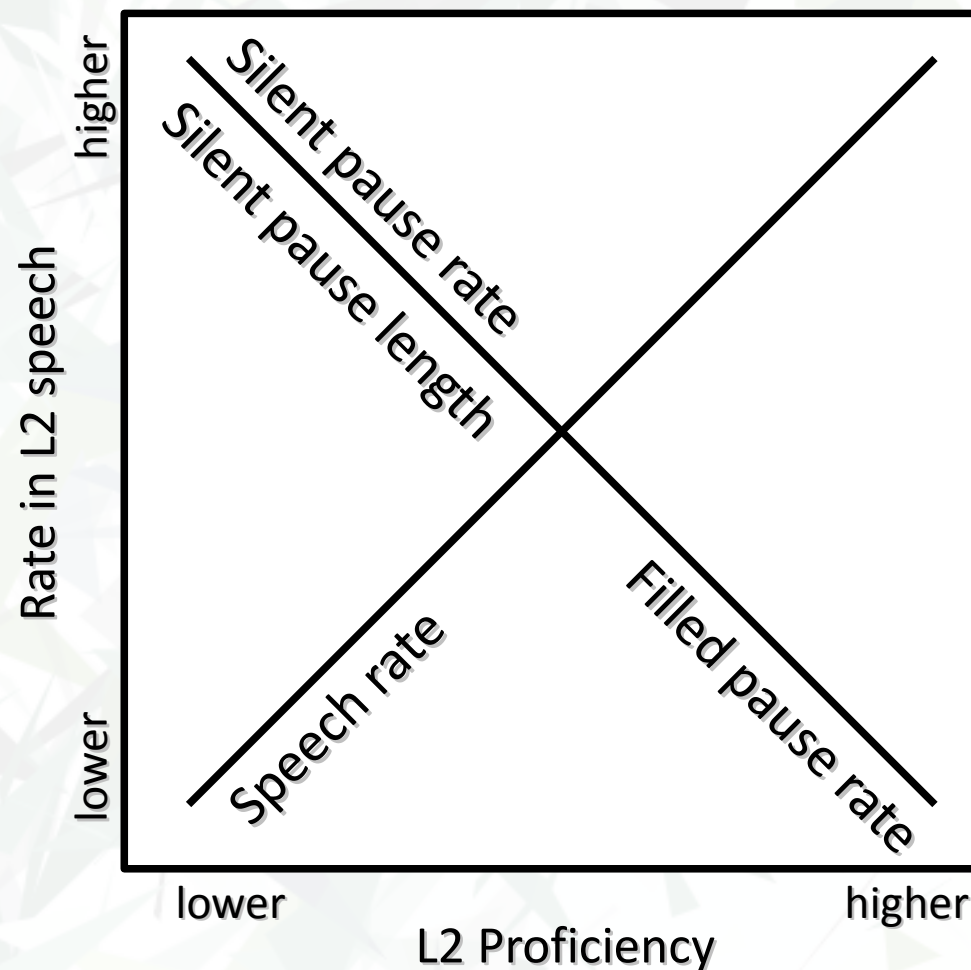
(Cucchiarini et al 2010, Kormos and Dénes 2004, Riazantseva 2001, Rieger 2003, Tavakoli 2011, Trofimovich and Baker 2006, 2007, Wu 2008)

Hesitation phenomena in L2 production

- Gradually, more studies are including L1 observations.
    - Derwing et al (2009) and Cox and Baker-Smemoe (2012) observed that both speech rate and pause rate in L1 and L2 production are correlated.
    - De Jong et al (2015) found measures of L2 articulation rate were more meaningful when "corrected" for L1 speech patterns.
- The current research is part of a project designed to contribute to greater understanding of the relationship between L1 hesitation patterns and L2 hesitation patterns.

Research questions

- What is the relationship between hesitation patterns in L1 and L2 speech?
- What is the developmental trajectory of the use of hesitation phenomena in L2?
- What is the relationship between hesitation patterns in L2 speech and hearer's perception of fluency?

Crosslinguistic Corpus of HP (CCHP)

- Participants: L2 learners of varying proficiency levels
- Elicitation tasks
  - Spontaneous speech: picture description, topic narrative
  - Reading aloud
  - Performed in both L1 and L2
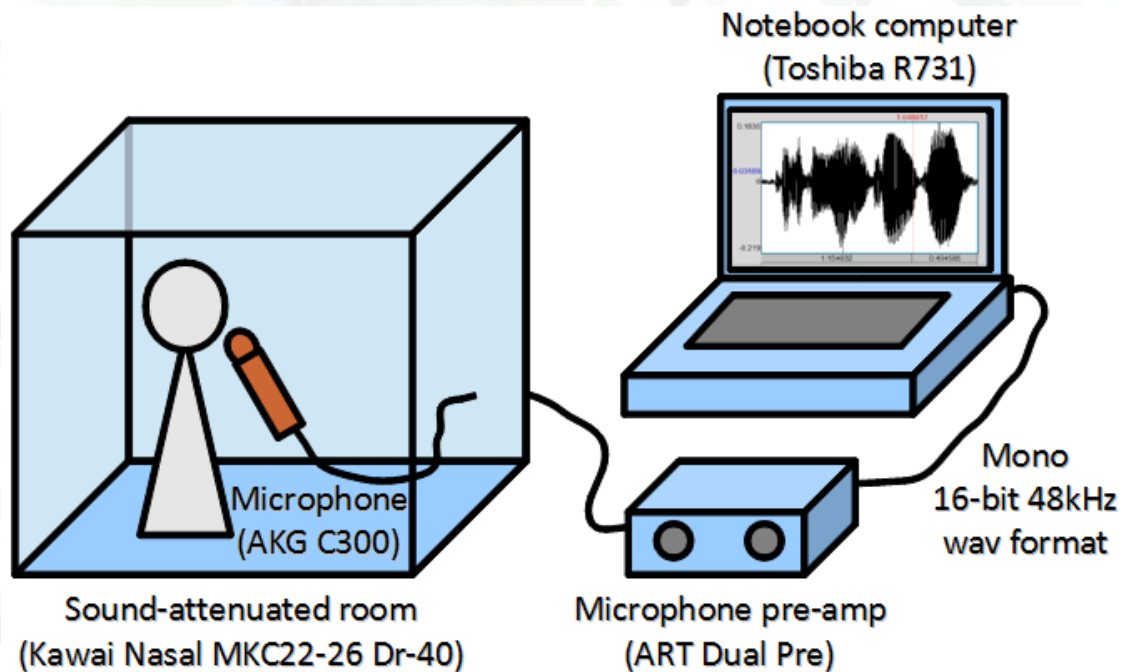
Crosslinguistic Corpus of HP (CCHP)

- Demographic information: age, gender, L2 proficiency (standardized test scores, experience abroad, self-assessment)

- Annotation
  - Transcripts, HP, word & pause intervals
  - Two annotators, one checker

```
<UTTERANCE>
 <T>in</T>
 <T>America</T>
 <T FILLED-PAUSE="yes">uh</T>
 <T>there's</T>
 <T>a</T>
 <T FILLED-PAUSE="yes">uh</T>
 <T>very</T>
 <T>famous</T>
 <T FILLED-PAUSE="yes">uh</T>
 <T>and</T>
 <T>loved</T>
 <T FILLED-PAUSE="yes">uh</T>
 <T>basketball</T>
 <RP>
  <O>
   <T>cl#</T>
  </O>
  <T FILLED-PAUSE="yes">uh</T>
  <E>
   <T>association</T>
  </E>
 </RP>
 <T>which</T>
 <T>is</T>
 <T>called</T>
 <T>NBA</T>
 <T>National</T>
 <T>Basketball</T>
 <T>Association</T>
 <T>I</T>
 <T>think</T>
</UTTERANCE>
```

CCHP Results: Basic Statistics

- Participants: 36 Japanese L1 / English L2 speakers
- Full corpus
  - 62,632 words
  - 11 hrs, 31 min
- Spontaneous speech
  - 40,296 words
  - 8 hrs, 43 min
- Read speech
  - 22,336 words
  - 2 hr, 48 min
- Transcriber agreement
  - 91.5%

- 15,837 silent pauses
- 3,516 filled pauses
- 1,689 self-corrections
- 518 repeats

# L1-L2 Utterance Flu. vs. Perceptual Flu.

- Aim:
  - Compare L1 vs. L2 utterance fluency characteristics.
  - Examine which utterance fluency characteristics correlate with perceptions of fluency by hearers.
- L1-L2 utterance fluency factors measured with praat script (Quené et al 2011)
- Perceptual fluency
  - Fluency ratings (1=low … 9=high) obtained via Amazon Mechanical Turk
  - Obtained fluency ratings on 7 30-second clips of L2 speech from all corpus participants.
  - Used attention checks and background monitoring of audio player activity to check that instructions were followed.

L1-L2 Utterance Fluency
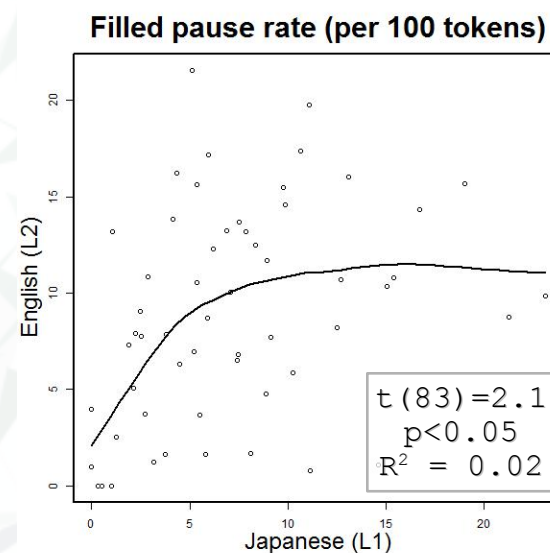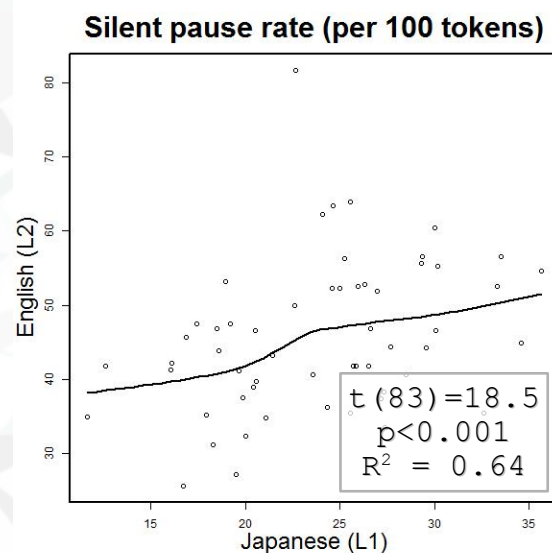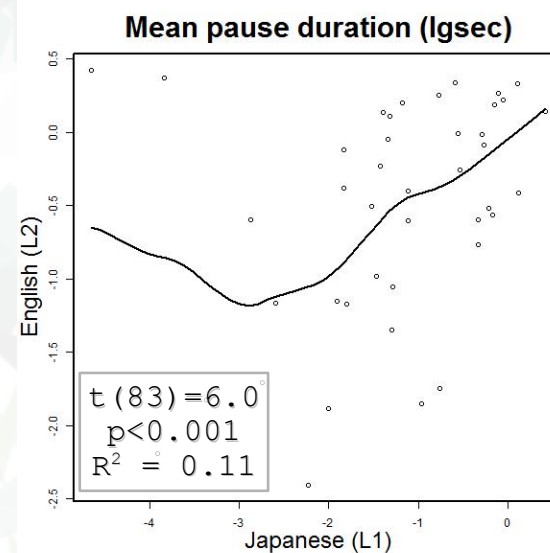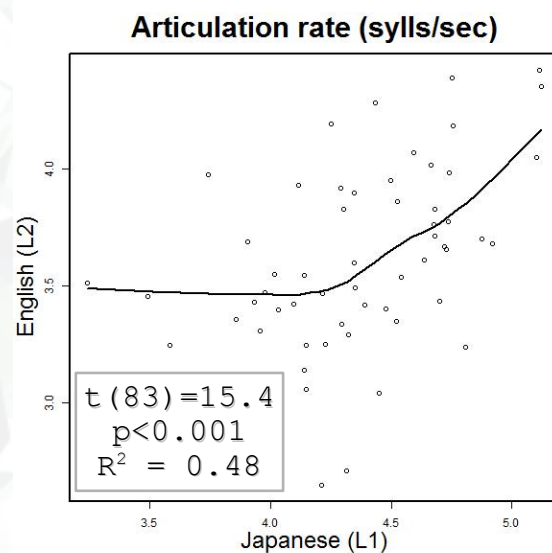
Utterance Fluency vs. Perceptual Fluency
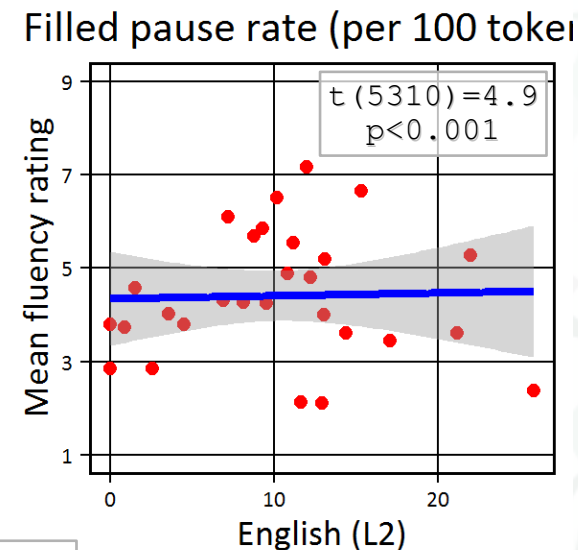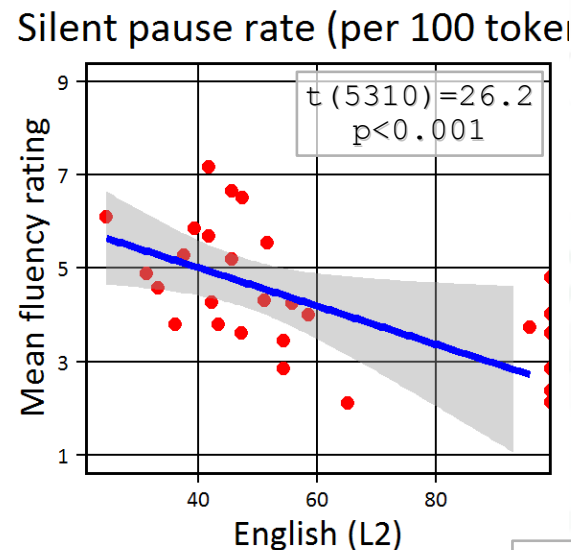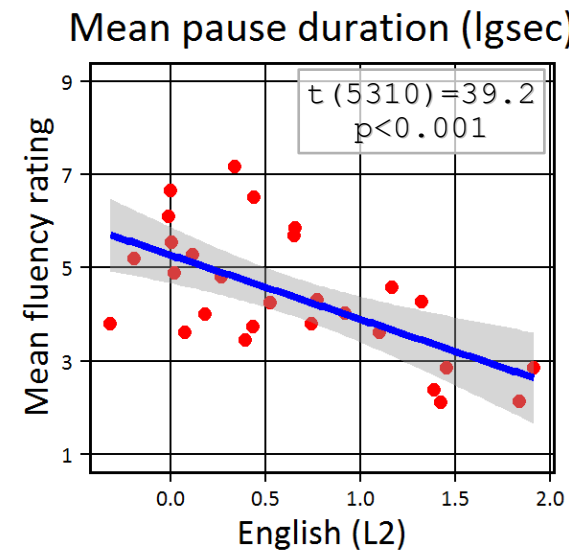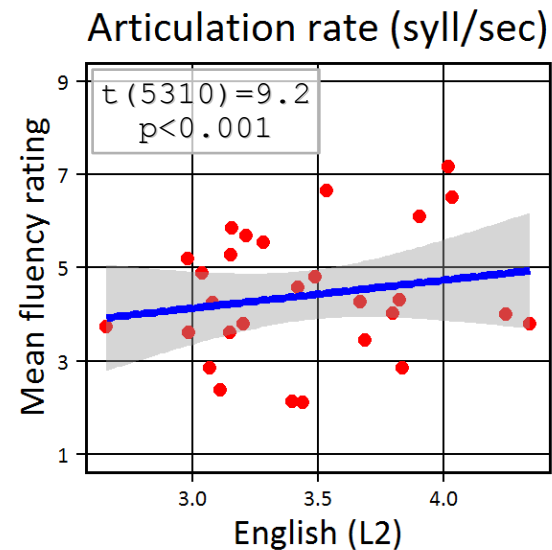
CCHP results: Summary

- Results show that for utterance fluency, silent pause duration and filled pause rate indicate learners' L2 proficiency.
  - Other L2 hesitation phenomena correlate with those of L1.
- Fluency raters, however, seem to rely on silent pause duration and rate, and less so on articulation rate and filled pause rate.

Future Work with CCHP

- Deeper annotation
  - Syntactic structure
  - Part-of-speech information
  - Syllable and phoneme intervals
  - (F1,F2) measurements
- More speakers
- More L1-L2 combinations
  - Taiwan Chinese L1 – English L2
  - English L1 – French L2
  - English L1 – Spanish L2
  - Korean L1 – English L2

CCHP Public Corpus

- Assembling a public version of the Crosslinguistic Corpus of Hesitation Phenomena is ongoing.
- When complete, audio files and annotated transcripts will be available for free download.
- Some files are already available for download: http://www.filledpause.com/chp/cchp
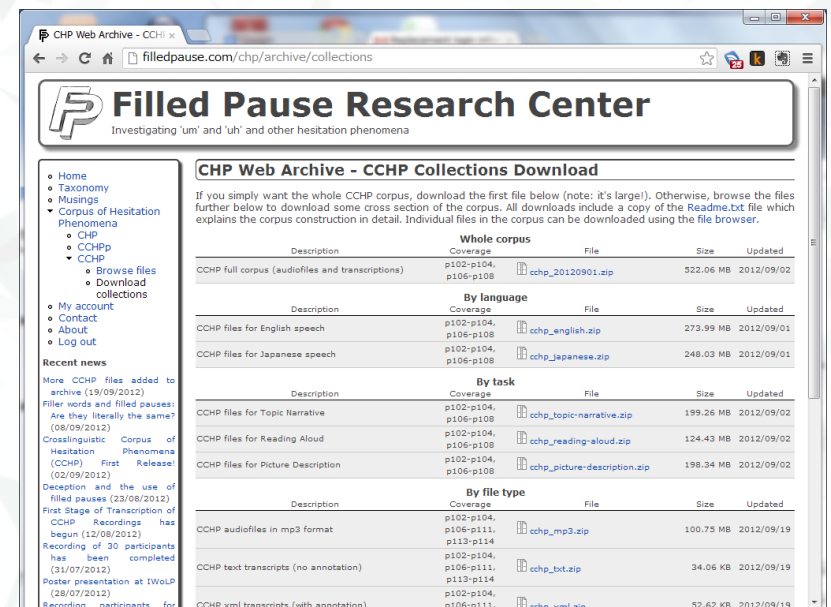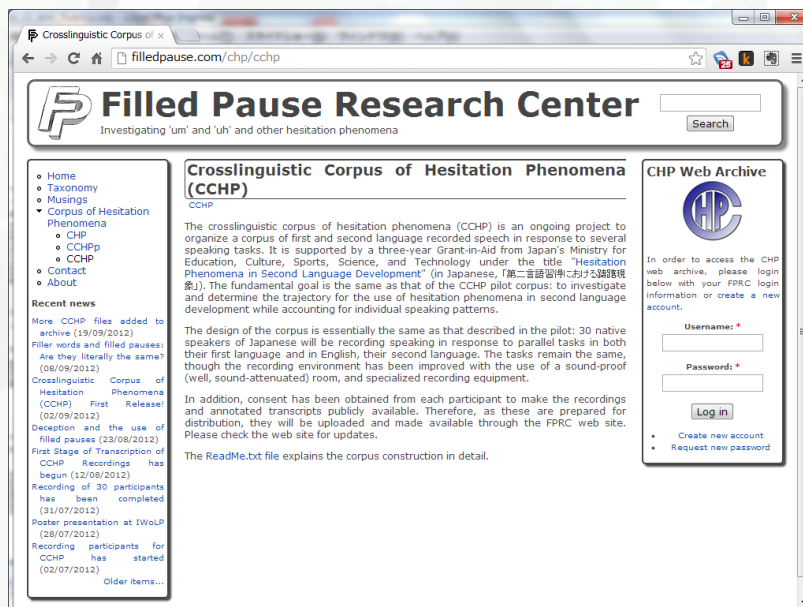
Automated assessment of L2 speech

- 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

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

detection settings

audio input settings

audio level meter

fluency measure indicators

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

Fluidity window contents:

File Help

Start

Stop

Speaking ideas: A description of a familiar place; How you spent your last vacation; Your most memorable experience; How to play a certain game or sport; Your opinion on an issue with reasons; and anything else...

Transition threshold 0.1
Pause threshold 0.3
Silence threshold -70
Filled pause sensitivity 0.1
Smoothing factor 10

Elapsed time: 00:00.0
Speech time: 00:00.0
Pause time: 00:00.0
Syllable count: 0
Silent pause count: 0
Filled pause count: 0

○ Port ツスツピツーツカツー (Realtek High D, version

○ Port Realtek Digital Output (Realtek, version 6.3

○ 8000 hz
○ 11050 hz
○ 16000 hz
○ 22050 hz
◉ 44100 hz
○ 48000 hz

○ 8-bit
◉ 16-bit

Fluidity: playback window

playback controls

waveform representation

fluency visualization

filled pauses speech silent pauses
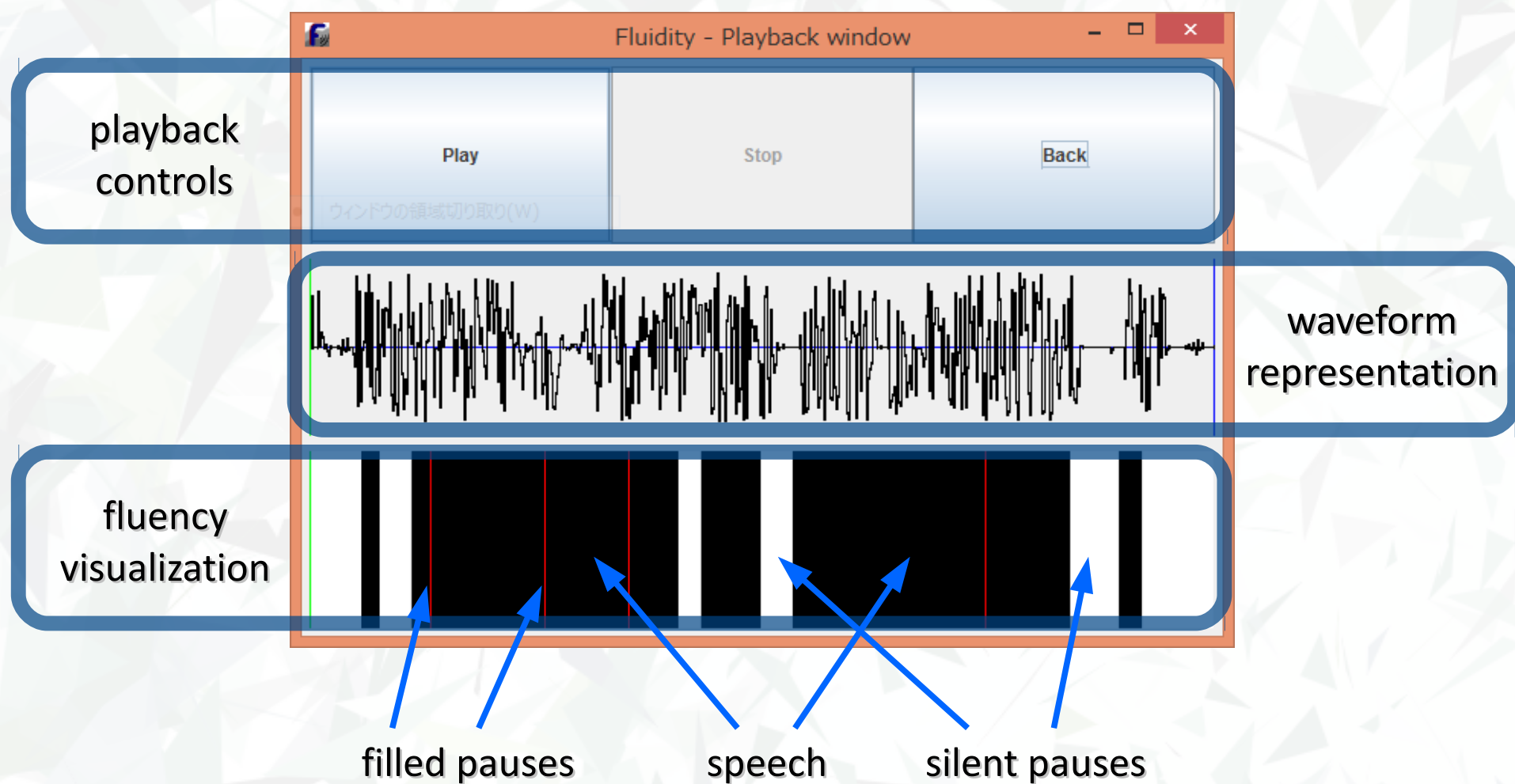
Fluidity: usability testing

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

Fluidity: user response
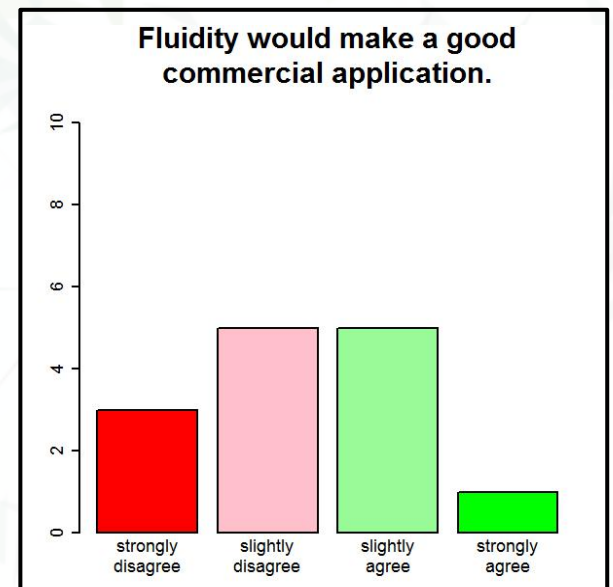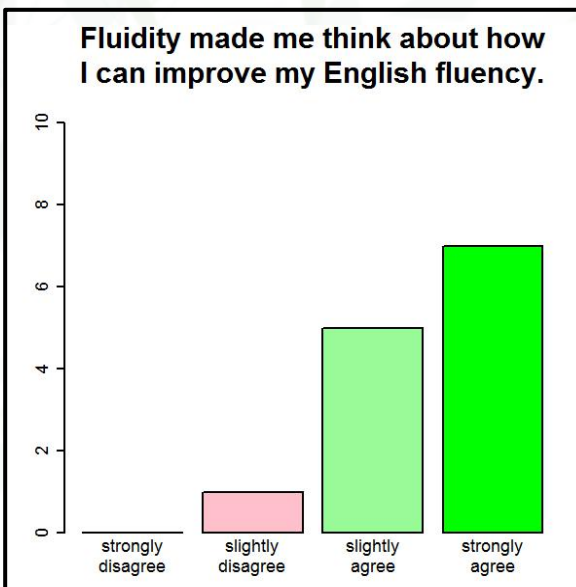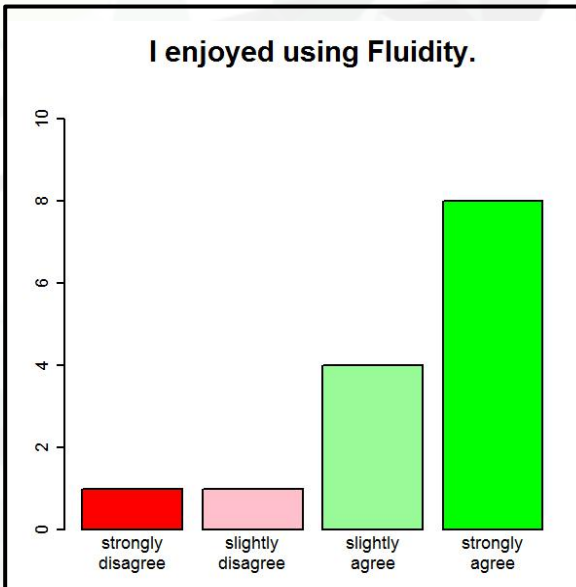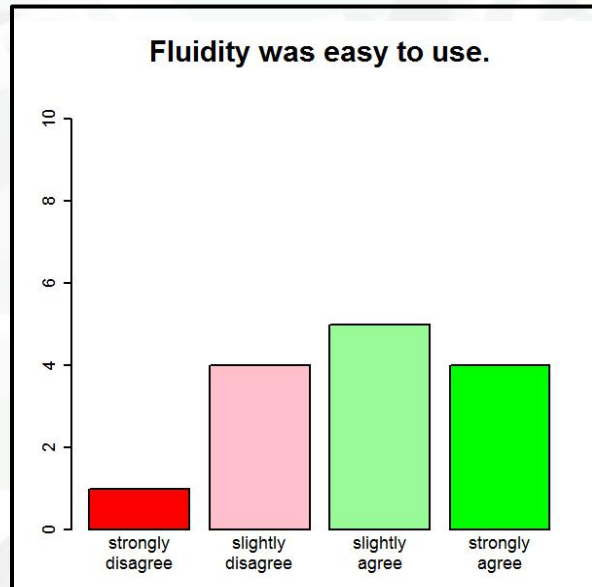
Fluidity: desirable features
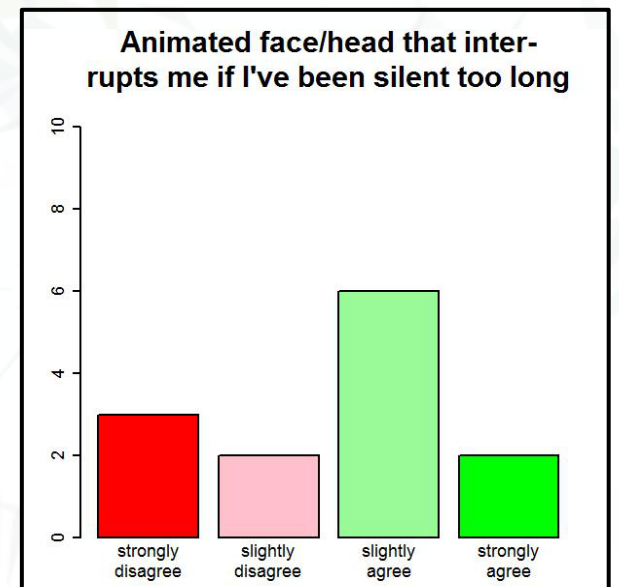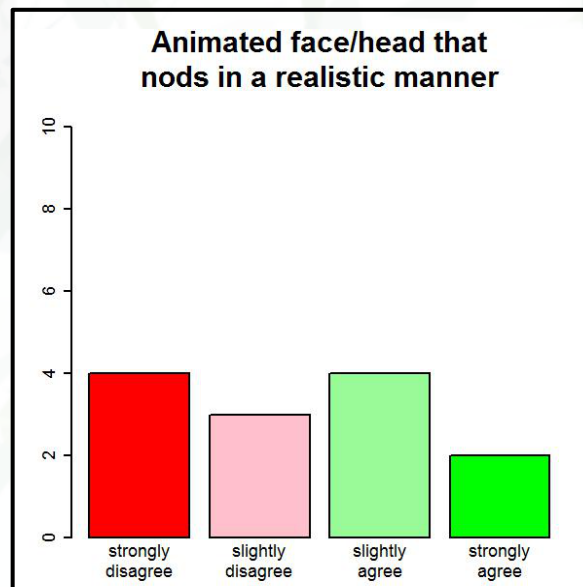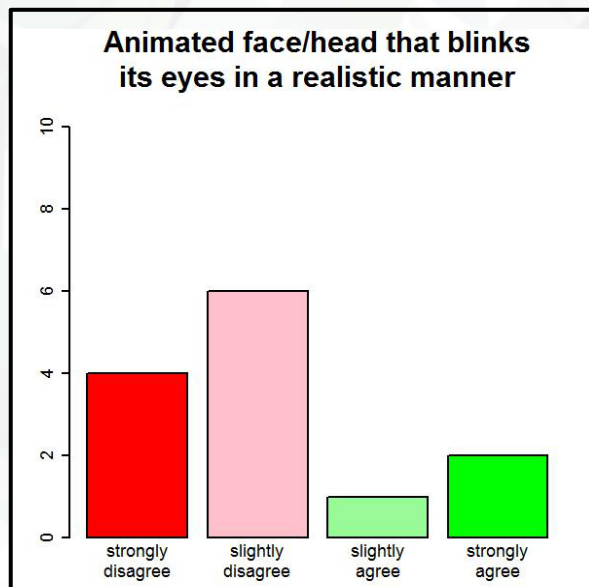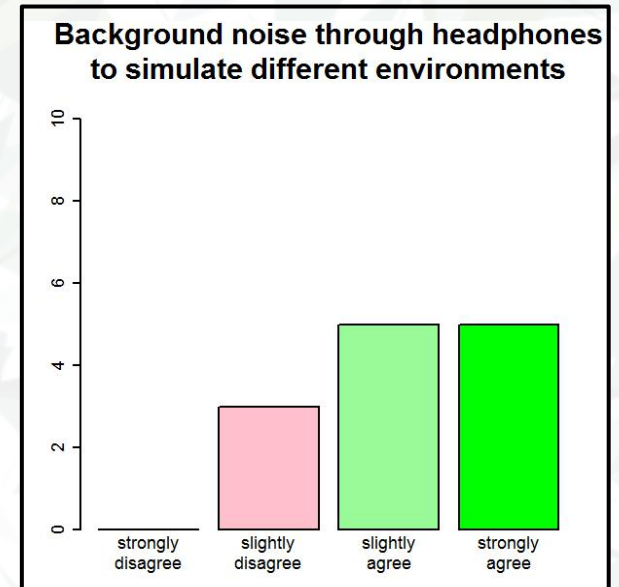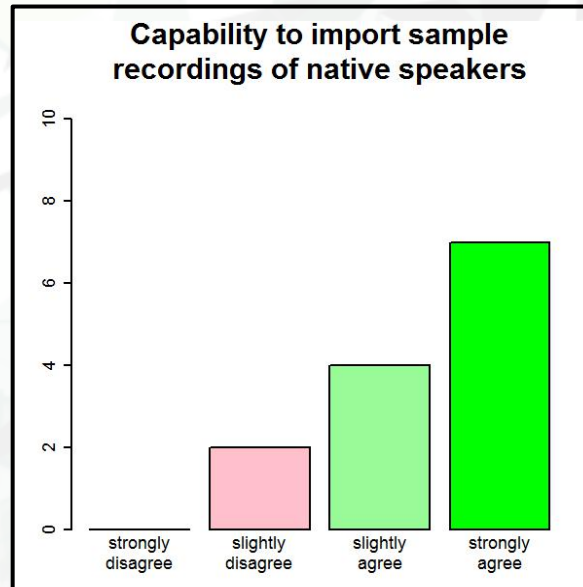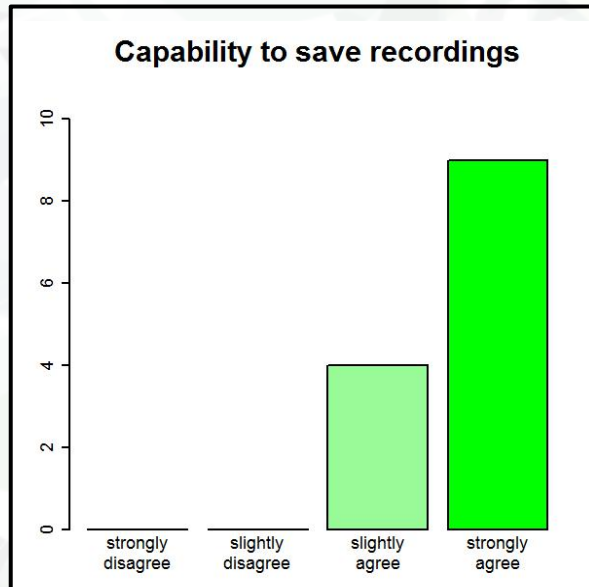
Fluidity: technical considerations
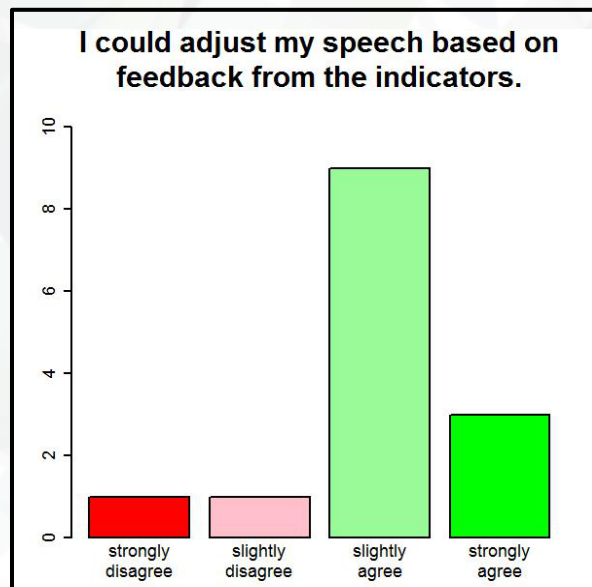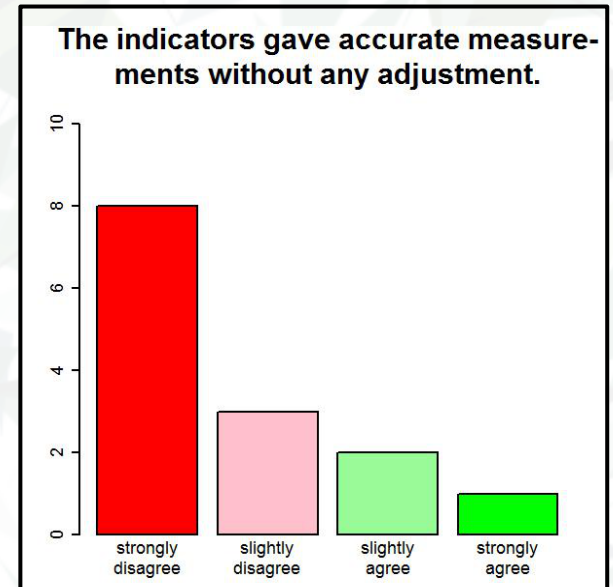
- Silence threshold often needs to be adjusted.
  - Solution: Do a better job of auto-detecting the sound configuration and microphone settings.
- Filled pause detection is difficult. Even after adjusting sensitivity, many participants did not see their filled pauses detected accurately (or at all).
  - 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.

Fluidity: 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.

Distribution of silent and filled pauses

- Campione and Véronis 2005: Examined SPs and FPs in the Spoken French Reference Corpus
  - SPs and FPs overlap to mark syntactic structure
  - FPs (with following SPs) mark non-syntactic interruptions
- Watanabe, et al 2013, 2014, 2015: Examined clause boundary depth and clause length in Japanese
  - SP duration is longer at stronger boundaries and before longer clauses
  - FP duration is not affected
- In short, SP and FP are not simply free alternatives when hesitation is needed.

Utterance length in L1 and L2
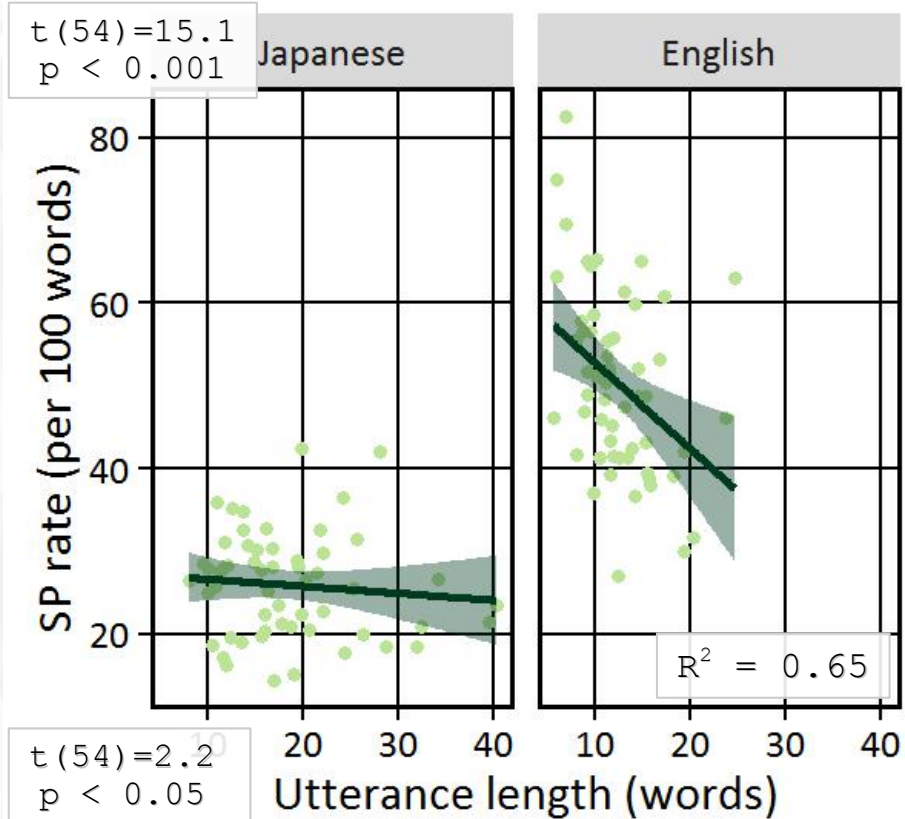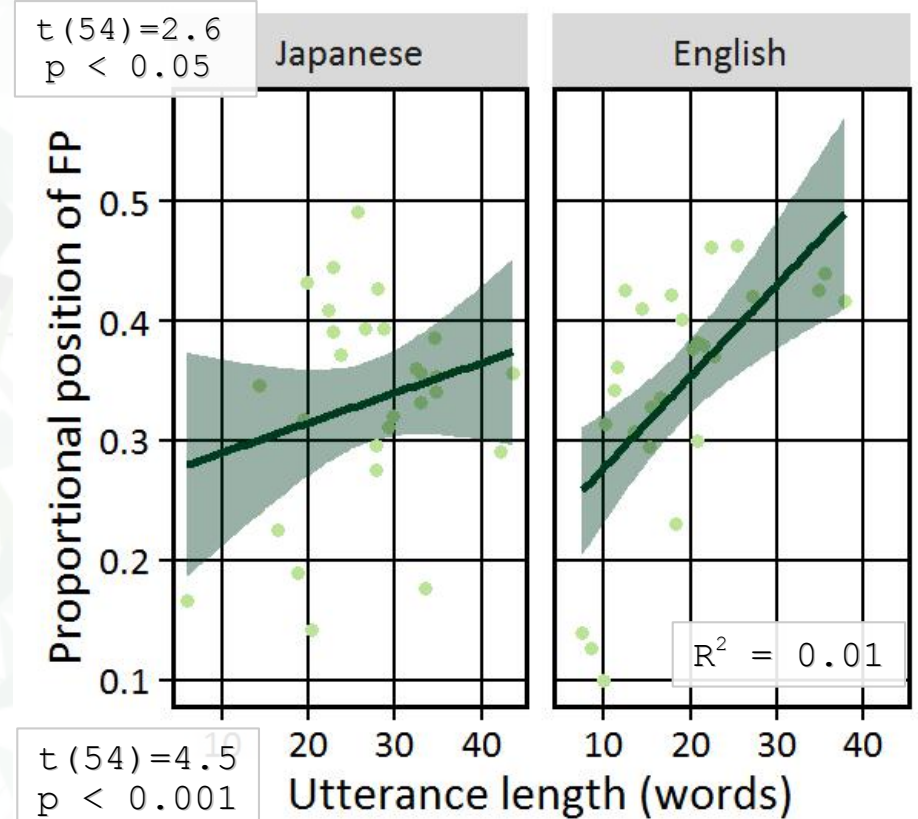
Fluency ... and complexity

- In pedagogical models of L2 development, complexity, accuracy, and fluency (CAF) are often treated as independent variables
- Some aspects of utterance fluency may give us insight on the complexity of L2 speakers' utterances.

Wrapping up

- **Determining L2 fluency**
  - While many temporal features of L2 utterances are observable, not all of them correlate to a perception of fluency on the part of listeners: Silent pause rate and duration > articulation rate and filled pause rate
- **Detecting L2 fluency**
  - Fluidity shows that certain aspects of utterance fluency can be detected in (nearly) real-time.
- **Developing L2 fluency**
  - Fluidity users feel that the real-time feedback on their utterance-level fluency features raises their awareness of their own fluency and provides a basis to improve it.

References

Amazon Mechanical Turk web site. www.mturk.com

Anderson-Hsieh, J. Using electronic visual feedback to teach supra-segmentals. *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.

Campione, E., & Véronis J. (2005). Pauses and hesitations in French spontaneous speech. The 4th Workshop on Disfluency in Spontaneous Speech, p. 43-46.

Cox, T. and Baker-Smemoe, W. (2012). The relationship between L1 fluency and L2 fluency across different proficiency levels and L1s. Presentation at Workshop Fluent Speech (Utrecht University, The Netherlands).

Cucchiarini, C., van Doremalen, J., & Strik, H. (2010). *Fluency in non-native read and spontaneous speech*. Paper presented at Proceedings of Disfluency in Spontaneous Speech (DiSS) and Linguistic Patterns in Spontaneous Speech (LPSS) Joint Workshop.

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.

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.

De Jong, N. Steinel, M.P., Florijn, A., Schoonen, R., and Hulstijn, J.H. (2012). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34(5): 893-916.

De Jong, N., Groenhout, R., Schoonen, R., Hulstijn, J.H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behaviour. *Applied Psycholinguistics*, 36(2): 223-243.

Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition, 31*(4), 533-557.

Don, S., & Lickley R. (2015). Uh I forgot what I was going to say: How memory affects fluency. The 7th Workshop on Disfluency in Spontaneous Speech.

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

Eskenazi M. (2009) An overview of spoken language technology for education, *Speech Communication*, 51(10): 832-844

Gamper, J. & Knapp, J. (2002) A Review of Intelligent CALL Systems, *Computer Assisted Language Learning*, 15(4): 329-342

Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech, 4*(1): 18-26.

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.

Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145-164.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41-104.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*, MIT Press, ACL-MIT Press series in natural-language processing.

Maclay, H., & Osgood, C. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15*, 19-44.

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

Quené, H., Persoon, I., de Jong, N. Praat Script Syllable Nuclei v2 [Praat Script]. Version 28 Feb 2011, retrieved 26 Dec 2014 from https://sites.google.com/site/speechrate/Home/praat-script-syllable-nuclei-v2.

Riazantseva, A. (2001). Second language proficiency and pausing a study of Russian speakers of English. *Studies in Second Language Acquisition, 23*(4), 497-526.

Rieger, C. L. (2003). *Disfluencies and hesitation strategies in oral L2 tests*. Paper presented at Proceedings of DiSS '03, Disfluency in Spontaneous Speech Workshop.

Rochester, S. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research, 2*(1), 51-81.

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.

Tavakoli, P. (2011). Pausing patterns: differences between L2 learners and native speakers. *ELT Journal, 65*(1), 71-79.

Trofimovich, P., & Baker, W. (2006). Lea,rning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*, 1-30.

Trofimovich, P., & Baker, W. (2007). Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics, 28*(2), 251-276.

Watanabe, M. (2013) Phrasal complexity and the occurrence of filled pauses in presentation speeches in Japanese, In Proceedings of Disfluency in Spontaneous Speech, Stockholm, pp. 71-72.

Watanabe, M. (2013) Complexity of upcoming phrases as a factor affecting the use of filled pauses in Japanese, The 327th Regular Meeting of The Phonetic Society of Japan, National Institute for Japanese Language and Linguistics, Tokyo, the abstract in Journal of Phonetic Society of Japan, vol. 17 No.2, p.90.

Watanabe, M., and Kashiwagi, Y. (2014) The effects of boundary type and length of the following clause on the duration of silent and filled pauses at clause boundaries in Japanese presentation speeches. Proceedings of the Autumn Meeting, Acoustical Society of Japan, Sapporo, pp. 281-282.

Watanabe, M., & Kashiwagi, Y. (2014) Syntactic complexity between dependent phrases and their heads affects filler rate immediately after dependent phrases in Japanese. Journal of Phonetic Society of Japan, vol. 18 No.1, pp. 45-56.

Watanabe, M., Kashiwagi Y., & Maekawa K. (2015). The relationship between preceding clause type, subsequent clause length and duration of silent and filled pauses at clause boundaries in Japanese monologues. The 7th Workshop on Disfluency in Spontaneous Speech (DiSS 2015).

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.

Wu, C.-H. (2008). *Filled pauses in L2 Chinese: A comparison of native and non-native speakers*. Paper presented at Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20), Columbus, Ohio.

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.