

Temporal Variables in First and Second Language Speech and Perception of Fluency

Ralph L. Rose

<rose@waseda.jp> <@fildpauz>

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



Acknowledgments

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

International Congress of Phonetic Sciences
Glasgow, Scotland, UK
10-14 August 2015

This research is partially supported by Japan
Society for the Promotion of Sciences (JSPS)
Grants-in-Aid (#24520661, #15K02765)

Fluency

- Segalowitz (2010): levels of fluency

- Cognitive fluency: ease of mental preparation

De Jong et al (2012)

- Utterance fluency: smoothness of articulation

- Perceptual fluency: hearer's view of smoothness

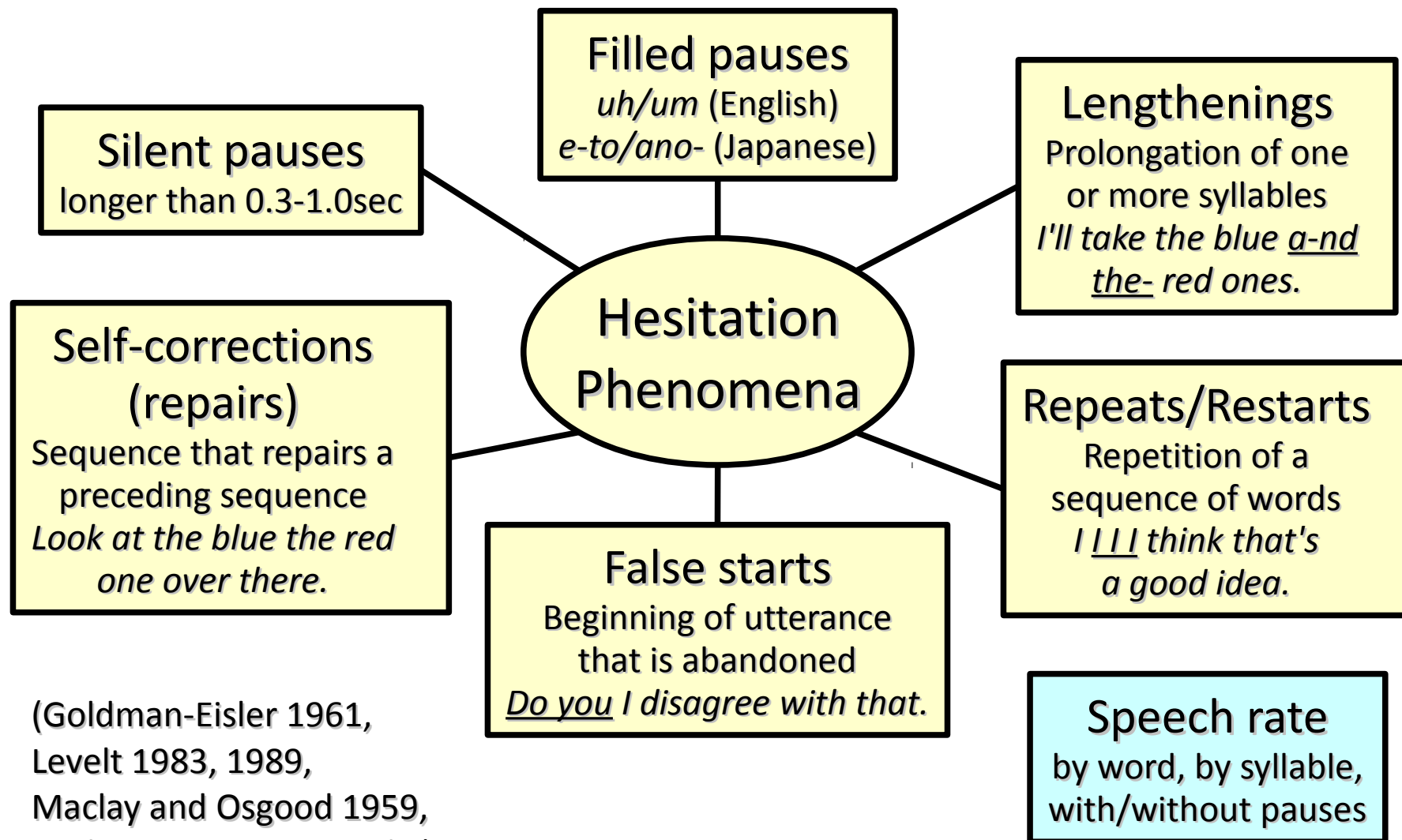
This study

- De Jong et al (2012) investigated relationship between cognitive fluency and utterance fluency.

- L2 speech rate related to cognitive fluency

- L2 Silent pause duration weakly related

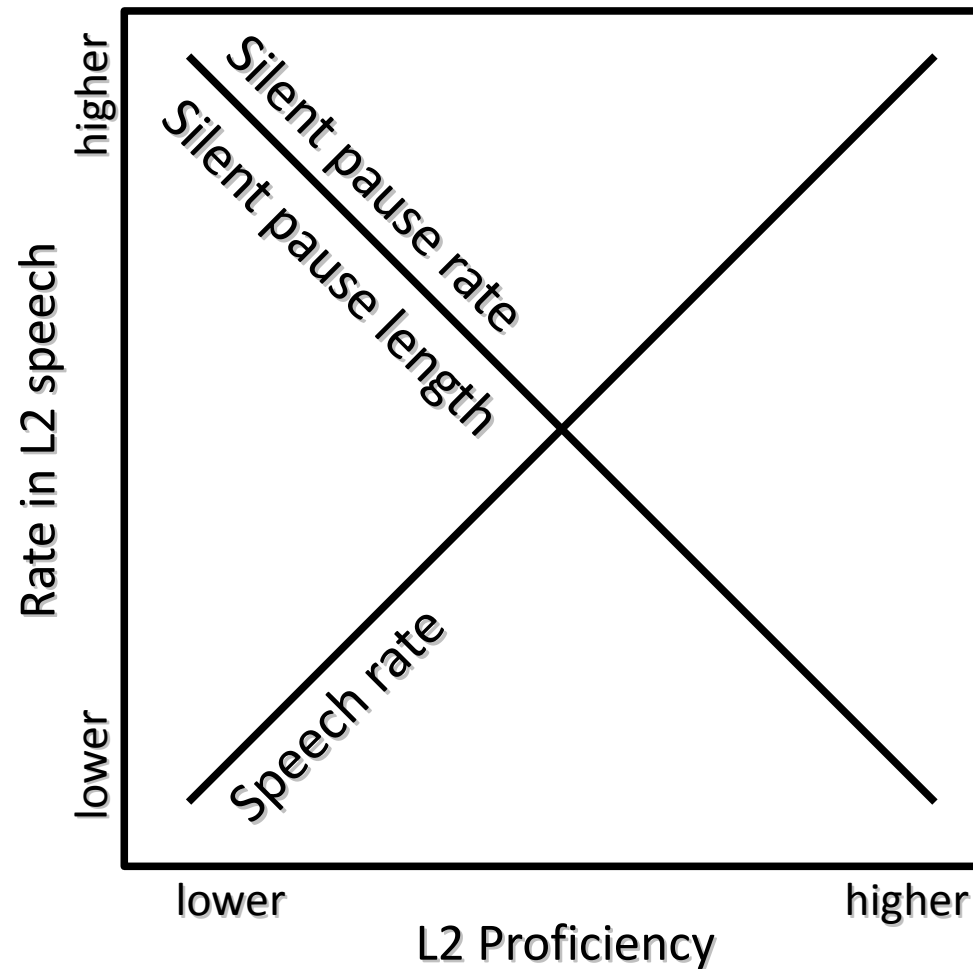
Observations of utterance fluency: Temporal variables



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

(Cucchiarini et al 2010)

Temporal variables in L2 production



(Cucchiari et al 2010, Kormos and Dénes 2004, Riazantseva 2001, Tavakoli 2011, Trofimovich and Baker 2006, 2007, Wu 2008)

Temporal variables in L2 production

- As a whole, work has been quite comprehensive.
- Lack of L1-L2 data from same speaker (cf., Cutler plenary)
- 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.
- Aim of the present research: Examine which utterance fluency characteristics correlate with perceptions of fluency by hearers.

Crosslinguistic Corpus of Hesitation Phenomena (CCHP)

- Participants: L2 learners of varying proficiency levels
- Elicitation tasks
 - Spontaneous speech: picture description, topic narrative
 - Reading aloud
 - Performed in both L1 and L2
- Annotation
 - Transcripts, HP, word and 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>
```

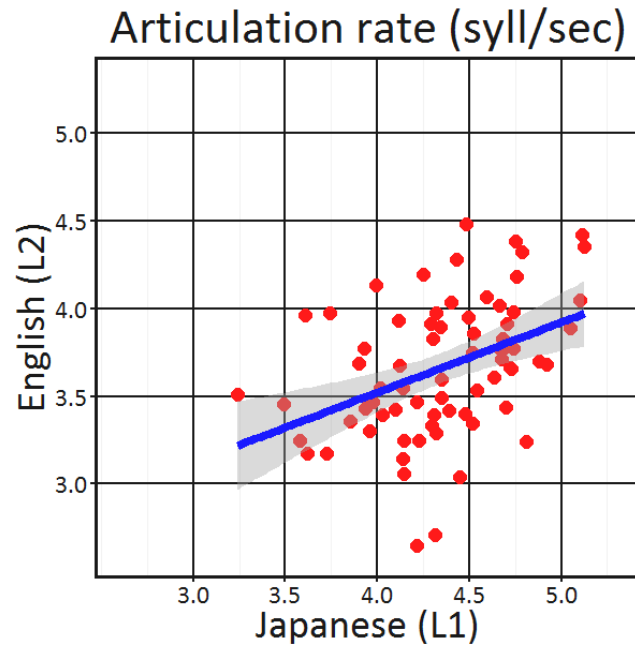
CCHP: Basic Statistics

- Participants: 36 Japanese L1 / English L2 speakers
- L1-L2 utterance fluency factors measured with Praat script (Quené et al 2011)

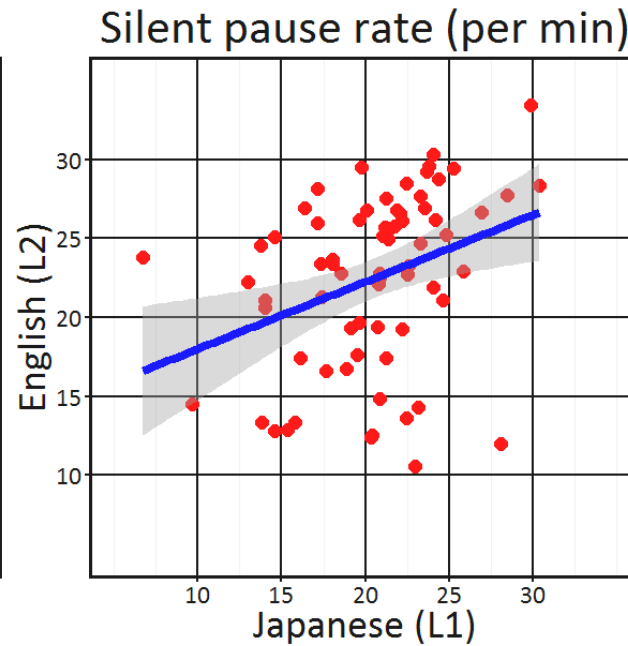
	Word count	Time
Read speech	22,336	2 hr, 48 min
Spontaneous speech	40,296	8 hr, 43 min
Total	62,632	11 hr, 31 min

Transcriber agreement = 91.5%

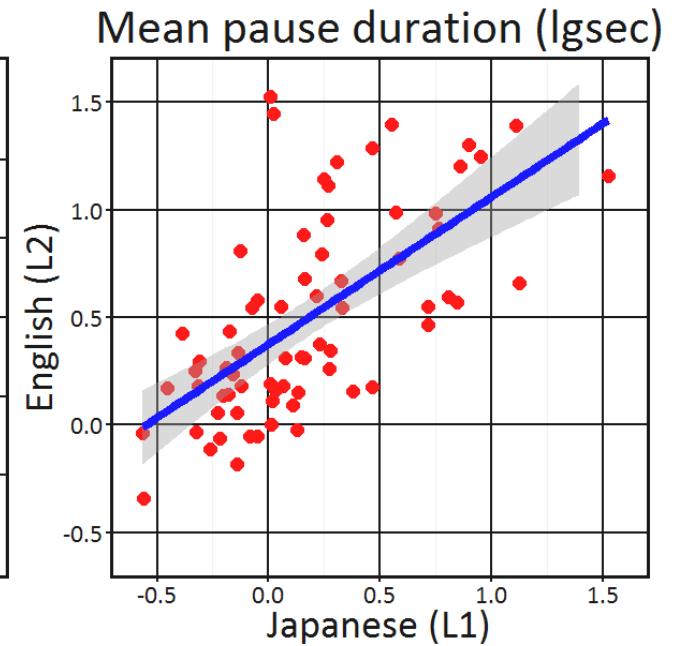
L1-L2 Utterance Fluency



$r=0.413$
 $p<0.001$
 $r^2=0.17$



$r=0.341$
 $p<0.005$
 $r^2=0.12$



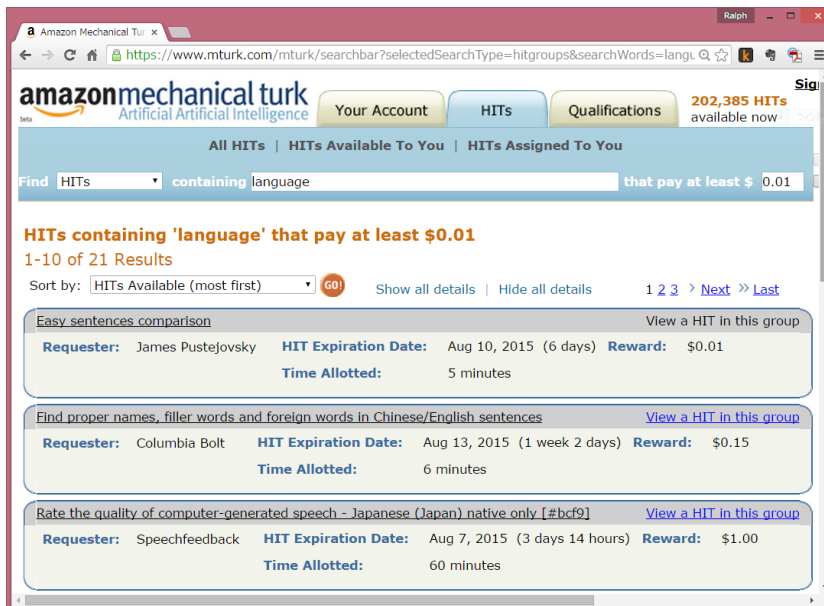
$r=0.636$
 $p<0.001$
 $r^2=0.40$

Pause duration > Articulation rate > Pause rate

Predictive of L2 proficiency

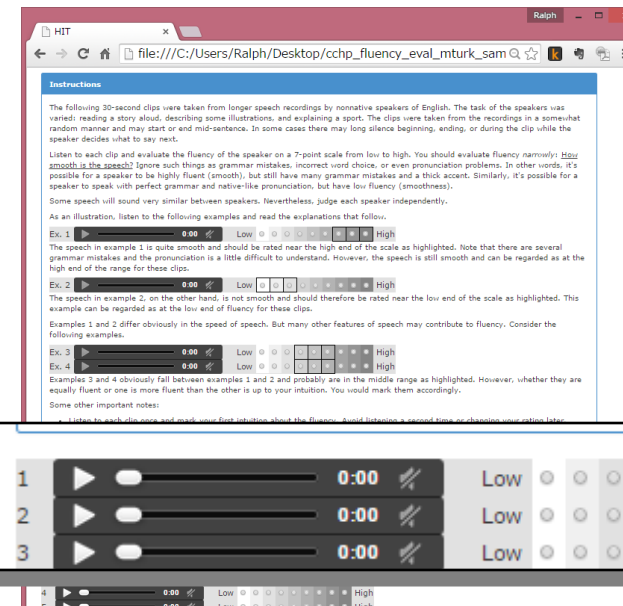
L2 Perceptual Fluency

- 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.



Amazon Mechanical Turk search results for HITs containing 'language' that pay at least \$0.01. The page shows 21 results, with the first three visible:

HIT Title	Requester	HIT Expiration Date	Reward	Time Allotted
Easy sentences comparison	James Pustejovsky	Aug 10, 2015 (6 days)	\$0.01	5 minutes
Find proper names, filler words and foreign words in Chinese/English sentences	Columbia Bolt	Aug 13, 2015 (1 week 2 days)	\$0.15	6 minutes
Rate the quality of computer-generated speech - Japanese (Japan) native only [#bcf9]	Speechfeedback	Aug 7, 2015 (3 days 14 hours)	\$1.00	60 minutes



HIT instructions page for the 'Rate the quality of computer-generated speech' HIT. The instructions describe the task: listening to 30-second clips and evaluating the fluency of the speaker on a 7-point scale from low to high. The instructions include examples and a rating scale.

Instructions:

The following 30-second clips were taken from longer speech recordings by nonnative speakers of English. The task of the speakers was varied: reading a story aloud, describing some illustrations, and explaining a sport. The clips were taken from the recordings in a somewhat random manner and may start or end mid-sentence. In some cases there may be long silence beginning, ending, or during the clip while the speaker decides what to say next.

Listen to each clip and evaluate the fluency of the speaker on a 7-point scale from low to high. You should evaluate fluency narrowly: *How smooth is the speech?* Ignore such things as grammar mistakes, incorrect word choice, or even pronunciation problems. In other words, it's possible for a speaker to be highly fluent (smooth), but still have many grammar mistakes and a thick accent. Similarly, it's possible for a speaker to speak with perfect grammar and native-like pronunciation, but have low fluency (smoothness).

Some speech will sound very similar between speakers. Nevertheless, judge each speaker independently.

As an illustration, listen to the following examples and read the explanations that follow.

Ex. 1: [Audio player] Low High
The speech in example 1 is quite smooth and should be rated near the high end of the scale as highlighted. Note that there are several grammar mistakes and the pronunciation is a little difficult to understand. However, the speech is still smooth and can be regarded as at the high end of the range for these clips.


Ex. 2: [Audio player] Low High
The speech in example 2, on the other hand, is not smooth and should therefore be rated near the low end of the scale as highlighted. This example can be regarded as at the low end of fluency for these clips.

Examples 1 and 2 differ obviously in the speed of speech. But many other features of speech may contribute to fluency. Consider the following examples.

Ex. 3: [Audio player] Low High
Ex. 4: [Audio player] Low High
Examples 3 and 4 obviously fall between examples 1 and 2 and probably are in the middle range as highlighted. However, whether they are equally fluent or one is more fluent than the other is up to your intuition. You would mark them accordingly.

Some other important notes:

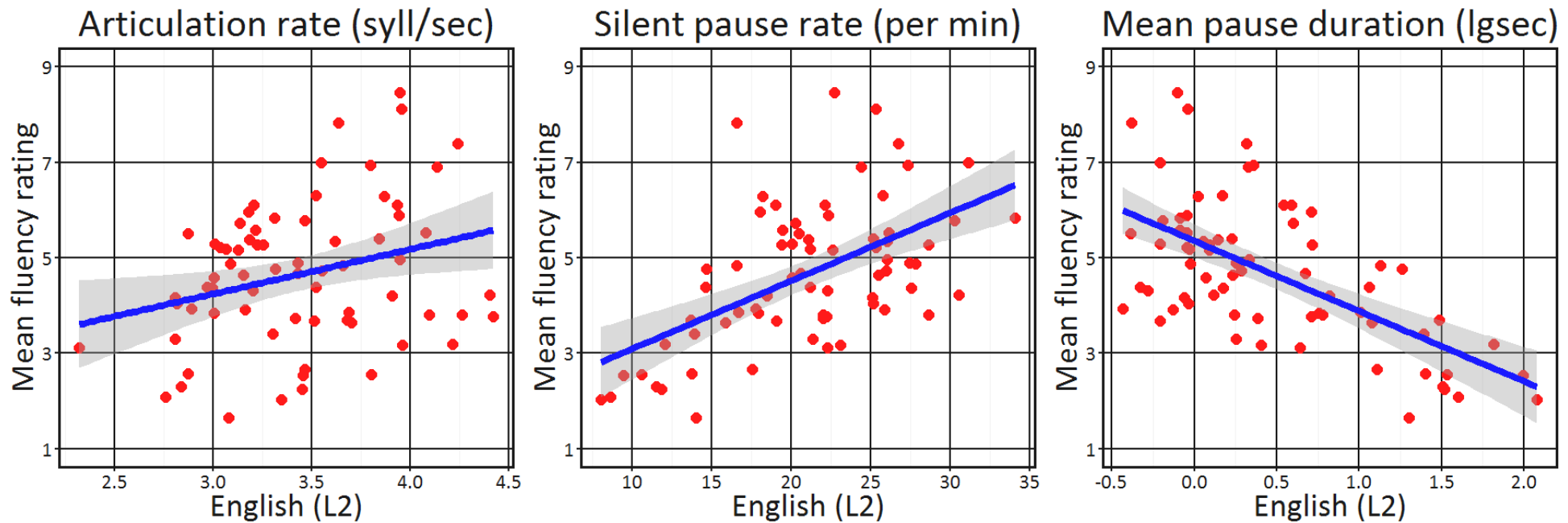
- Listen to each clip once and make your first intuition about the fluency. Avoid listening a second time to reassess your rating.



Rating scale for the HIT. The scale is a 7-point scale from Low to High. The first three examples are highlighted in the middle range.

Example	Rating
1	High
2	Low
3	Low

Utterance Fluency vs. Perceptual Fluency



	Est.	Std. Error	t	p
(Intercept)	2.1831	1.0524	2.074	<0.05
Articulation rate	1.0268	0.2997	3.426	=0.001
Mean pause duration	-0.6138	0.0861	-7.130	<0.001
Adjusted R ² = 0.4638; F(2, 67) = 30.84, p<0.001				

Summary and Implications

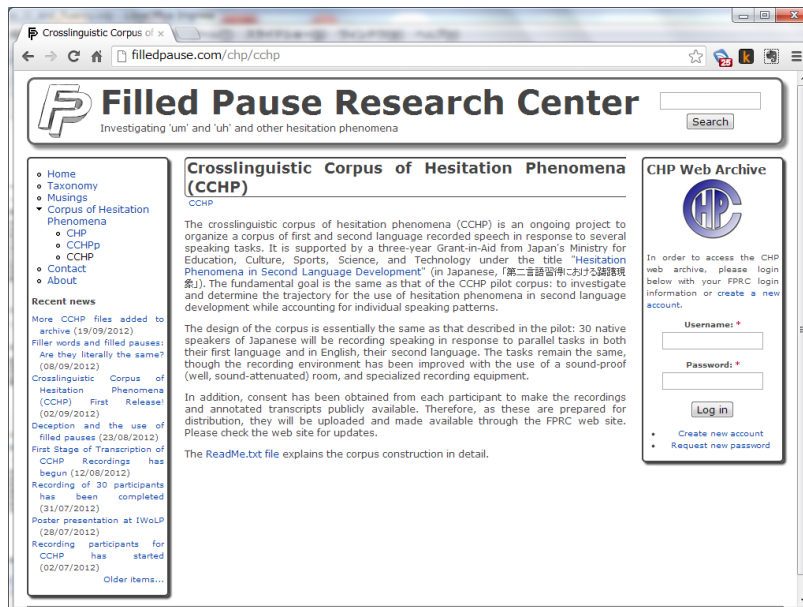
- Results show that for utterance fluency, silent pause rate is most indicative of learners' L2 proficiency.
 - Other L2 temporal variables correlate with those of L1.
- Fluency raters, however, seem to rely on articulation rate and silent pause duration instead.
- Implications for pedagogy
 - Raise awareness among L2 fluency raters of which temporal variables are truly indicative of L2 proficiency development.
 - Train speakers to speak faster and use shorter pauses.

Silent pause duration	Articulation rate
Strongest influence on perception of fluency	

Silent pause rate
Best predictor of L2 Proficiency

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>



The screenshot shows the homepage of the Filled Pause Research Center. The header includes the logo and the text "Investigating 'um' and 'uh' and other hesitation phenomena". A search bar is located in the top right. The main content area is divided into several sections: a navigation menu on the left, a central section titled "Crosslinguistic Corpus of Hesitation Phenomena (CCHP)" with a description of the project, and a "CCHP Web Archive" section with a login form. The login form includes fields for "Username" and "Password" and a "Log in" button. Below the login form are links for "Create new account" and "Request new password".



The screenshot shows the "CCHP Web Archive - CCHP Collections Download" page. It features a navigation menu on the left and a main section with a table of collections. The table is organized into three categories: "Whole corpus", "By language", and "By task". Each category has a sub-table with columns for "Description", "Coverage", "File", "Size", and "Updated".

Whole corpus				
Description	Coverage	File	Size	Updated
CCHP Full corpus (audiofiles and transcriptions)	p102-p104, p106-p108	cchp_20120901.zip	522.06 MB	2012/09/02
By language				
Description	Coverage	File	Size	Updated
CCHP files for English speech	p102-p104, p106-p108	cchp_english.zip	273.99 MB	2012/09/01
CCHP files for Japanese speech	p102-p104, p106-p108	cchp_japanese.zip	248.03 MB	2012/09/01
By task				
Description	Coverage	File	Size	Updated
CCHP files for Topic Narrative	p102-p104, p106-p108	cchp_topic-narrative.zip	199.26 MB	2012/09/02
CCHP files for Reading Aloud	p102-p104, p106-p108	cchp_reading-aloud.zip	124.43 MB	2012/09/02
CCHP files for Picture Description	p102-p104, p106-p108	cchp_picture-description.zip	198.34 MB	2012/09/02
By file type				
Description	Coverage	File	Size	Updated
CCHP audiofiles in mp3 format	p102-p104, p106-p111, p113-p114	cchp_mp3.zip	100.75 MB	2012/09/19
CCHP text transcripts (no annotation)	p102-p104, p106-p111, p113-p114	cchp_txt.zip	34.06 KB	2012/09/19
CCHP xml transcripts (with annotation)	p102-p104, p106-p111	rnhn_xml.zip	52.62 KB	2012/09/19

References

- 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).
- Cucchiari, 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.
- 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.
- 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.
- Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech*, 4(1), 18-26.
- 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.
- Riazantseva, A. (2001). Second language proficiency and pausing a study of Russian speakers of English. *Studies in Second Language Acquisition*, 23(4), 497-526.
- 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>.
- Rochester, S. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2(1), 51-81.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.
- Tavakoli, P. (2011). Pausing patterns: differences between L2 learners and native speakers. *ELT Journal*, 65(1), 71-79.
- Trofimovich, P., & Baker, W. (2006). Learning 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.
- 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.