

The Developmental Trajectory of Second Language Learners' Use of Hesitation Phenomena

Ralph L. Rose
<rose@waseda.jp>

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



Acknowledgments

Hiroaki Suzuki, Junichi Inagaki, Masayuki Motoori, & Yukikatsu Fukuda

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

Overview

- Hesitation phenomena
 - Overview
 - HP in L2 speech
- Crosslinguistic Corpus of Hesitation Phenomena
 - Description
 - Results
- HP Developmental Trajectory
- Accessing the CCHP

Overview of types of HP

- Long investigative history
 - Goldman-Eisler 1961, Levelt 1983, 1989, Maclay and Osgood 1959, Rochester 1973, inter alia
- Types
 - Silent pauses (SP): longer than 0.3-1.0 sec
 - Filled pauses (FP): *uh/um* in English, *e-to/ano-* in Japanese
 - Lengthenings: prolongation of one or more syllables
 - Repeats/restarts: repetition of a sequence of words
 - False starts: beginning of an utterance that is abandoned
 - Self-corrections: a sequence of words that repairs an immediately preceding sequence

HP in L2 production

- Findings (Cucchiariini et al 2010, Kormos and Dénes 2004, Riazantseva 2001, Rieger 2003, Tavakoli 2011, Trofimovich and Baker 2006, 2007, Wu 2008)
 - SP duration and rate: higher proficiency → shorter and fewer silent pauses
 - FP rate: higher proficiency → fewer filled pauses
 - Differences between read and spontaneous speech
- Related
 - Speech rate: higher proficiency → faster rate

HP in L2 production

- As a whole, work has been quite comprehensive.
- However, individual works are limited in that many do not take individual variation into account (cf., de Leeuw 2007).
- 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.
- The current research is designed to contribute to greater understanding of the influence of L1 hesitation on L2 hesitation.

Research Questions

- What is the relationship between hesitation patterns in L1 and L2 speech?
- What is the developmental trajectory of HP use in L2?

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
- Demographic information: age, gender, L2 proficiency (standardized test scores, experience abroad, self-assessment)
- Annotation
 - Transcripts, HP, word & pause intervals
 - Two annotators, one checker

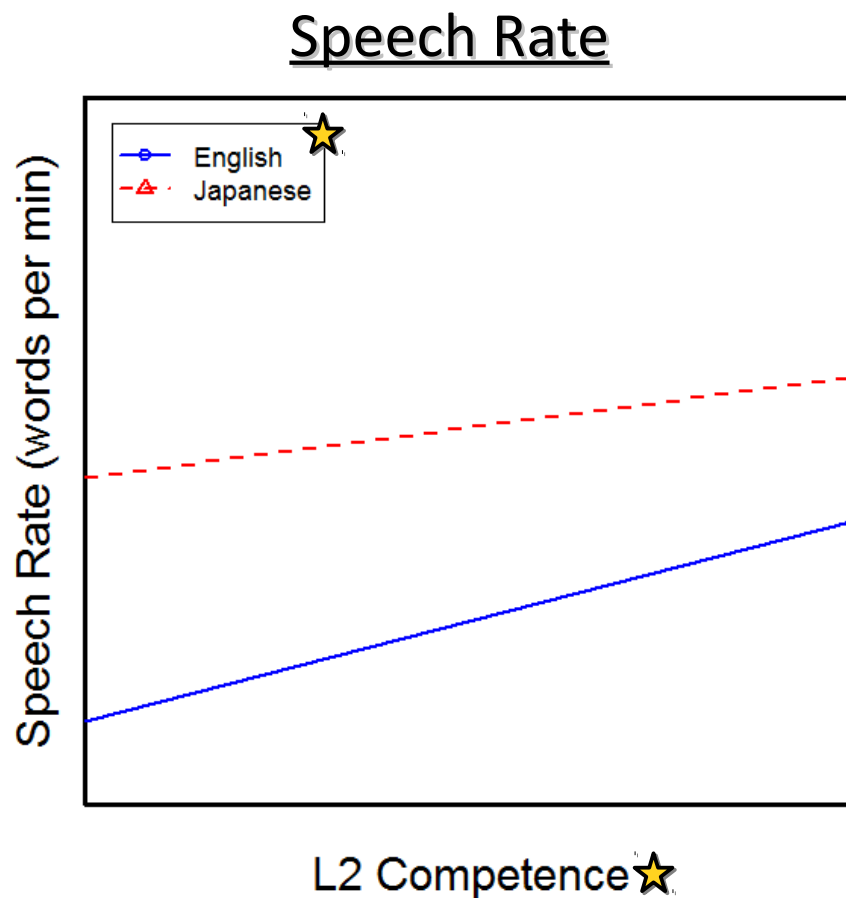
CCHP Results: Basic Statistics

- Participants: 25 Japanese L1, English L2 speakers
- Full corpus
 - 42,972 words
 - 8 hrs, 9 min
- Spontaneous speech
 - 27,416 words
 - 6 hrs, 12 min
- Read speech
 - 15,556 words
 - 1 hr, 57 min
- 11,091 silent pauses
- 2,404 filled pauses
- 1,080 self-corrections
- 309 repeats

CCHP Results: Analysis

- Used spontaneous speech data only.
- Extracted counts for speech rate, silent pauses, filled pauses, repeats, and self-corrections.
- Performed repeated measures ANOVA
 - (between) L2 Proficiency as numerical variable, estimated from test scores, experience abroad, self-assessment
 - (within) Language as categorical variable: Japanese, English
- Used $\alpha = 0.05$ for significance testing (marked with ★).

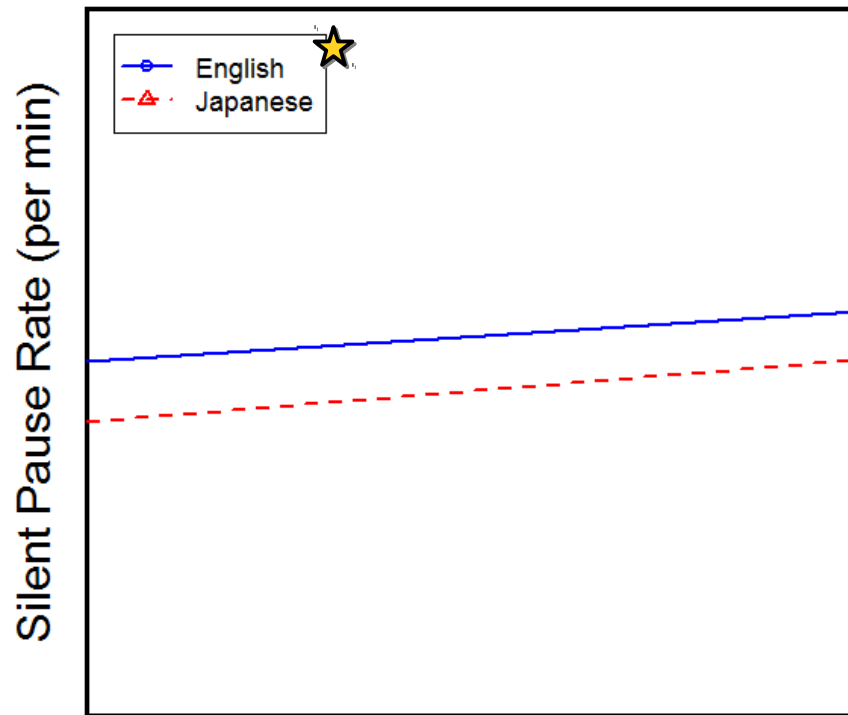
CCHP Results: Speech Rate



Consistent with Derwing et al (2009) and Cox and Baker-Smemoe (2012)

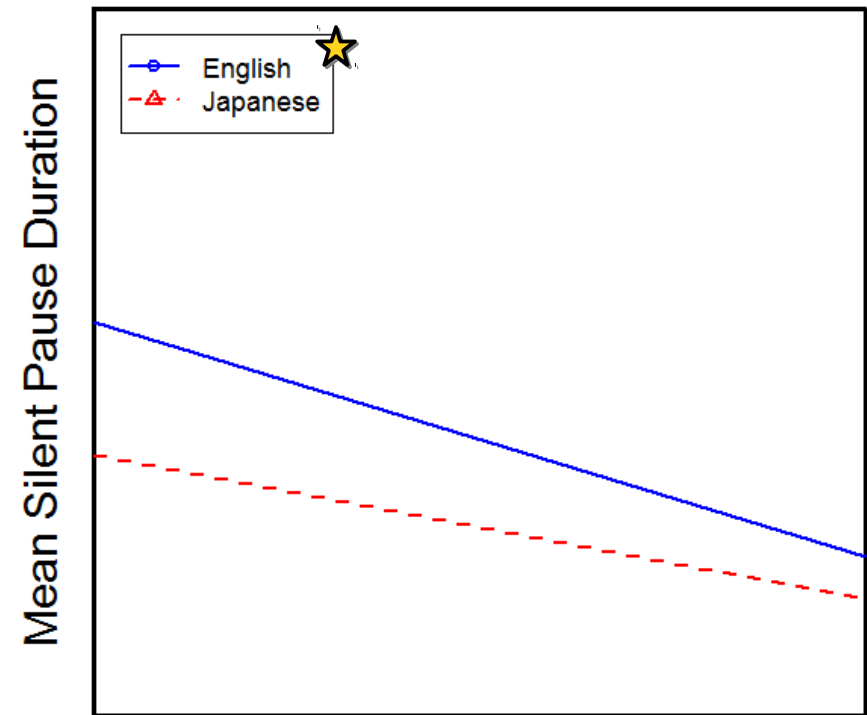
CCHP Results: Silent Pauses

Silent Pause Rate (per min)



L2 Competence

Silent Pause Duration

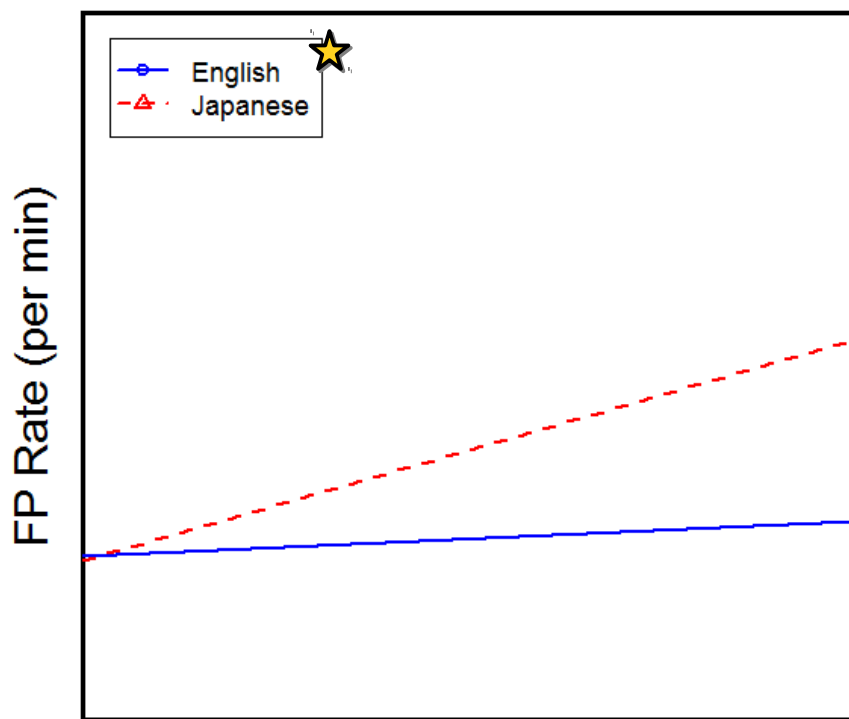


L2 Competence★

Consistent with Derwing et al (2009) and Cox and Baker-Smemoe (2012)

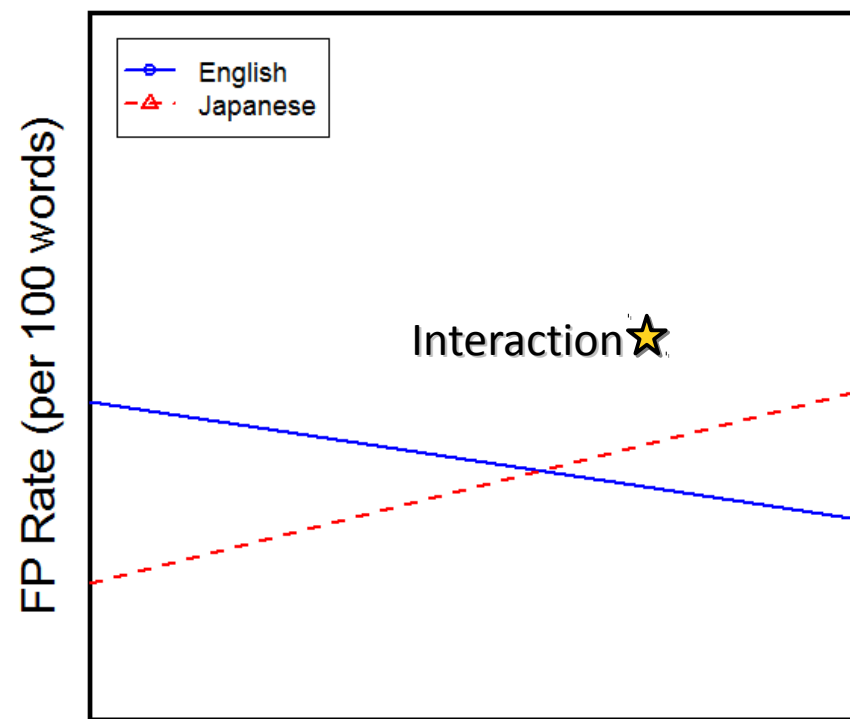
CCHP Results: Filled Pauses

Filled Pause Rate (per min)



L2 Competence

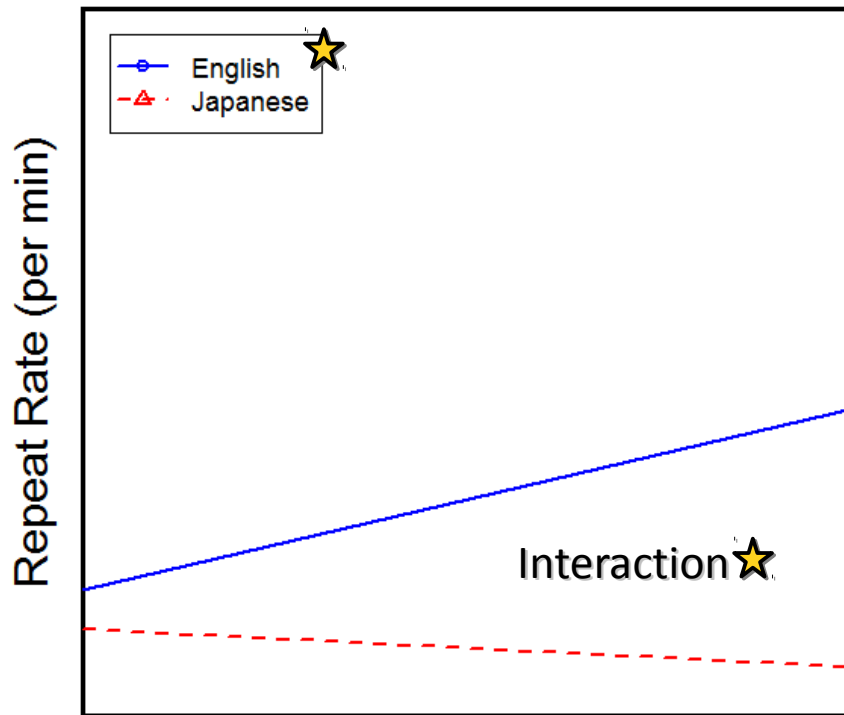
Filled Pause Rate (per word)



L2 Competence

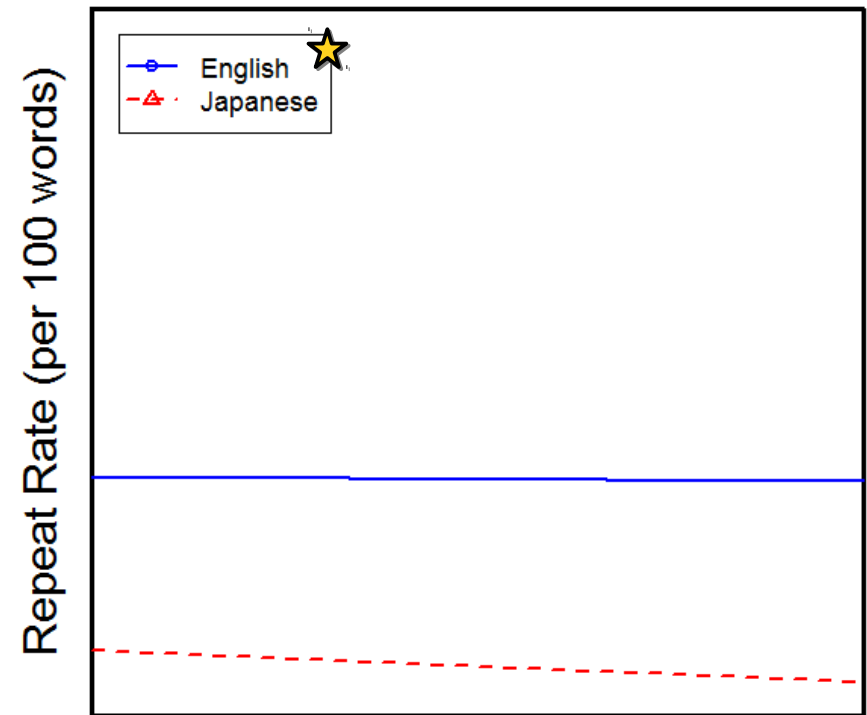
CCHP Results: Repeats

Repeat Rate (per min)



L2 Competence

Repeat Rate (per word)

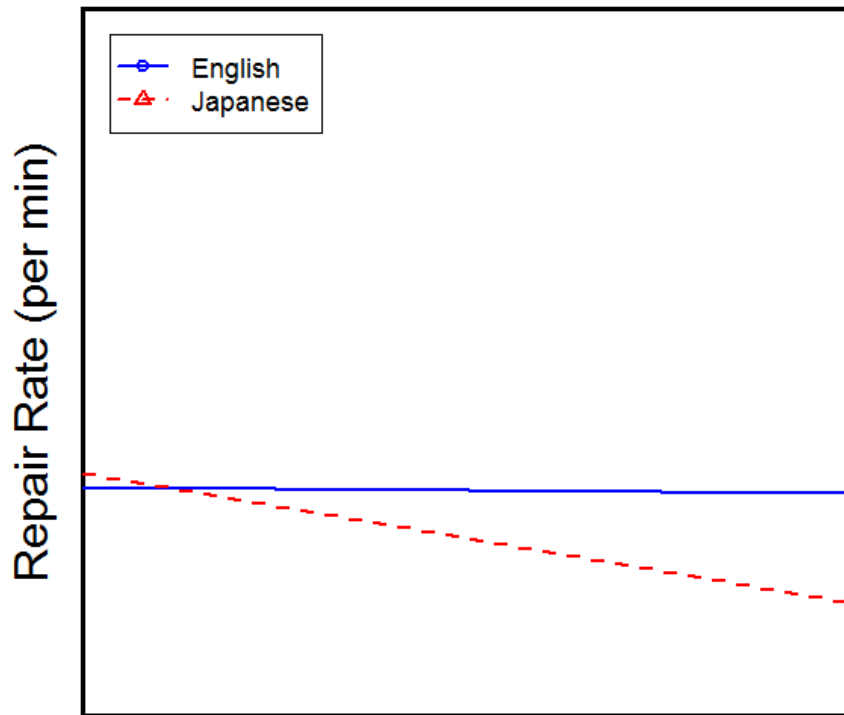


L2 Competence

Repeats are uncommon in Japanese (Fox et al 1996)

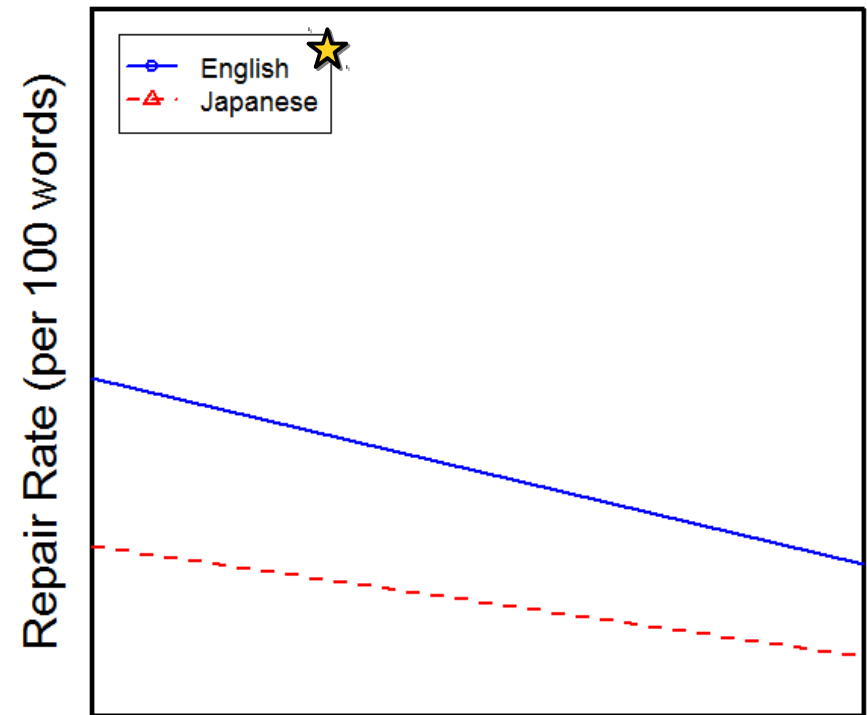
CCHP Results: Self-corrections

Self-correction Rate (per min)



L2 Competence

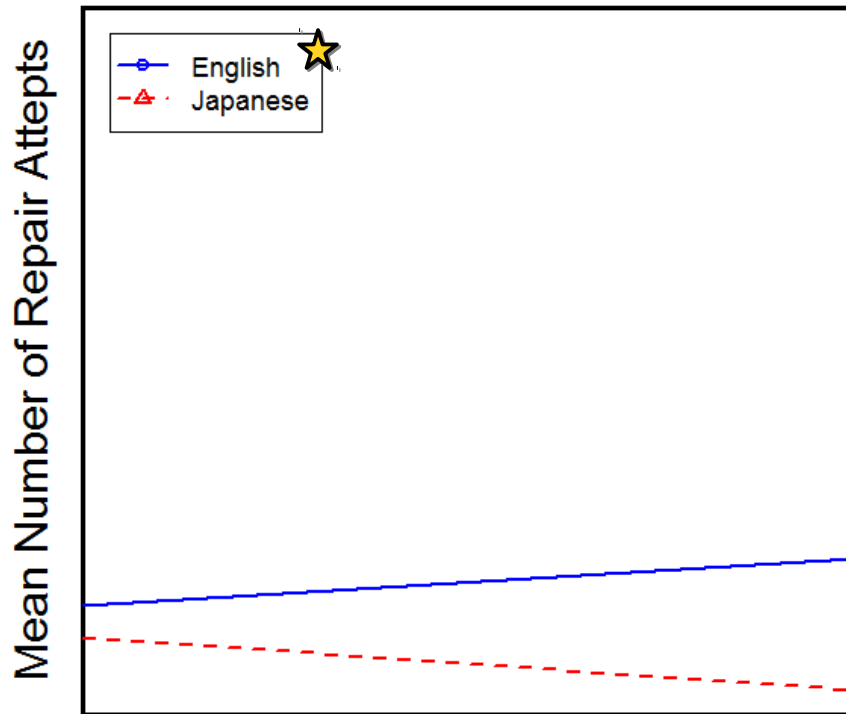
Self-correction Rate (per word)



L2 Competence★

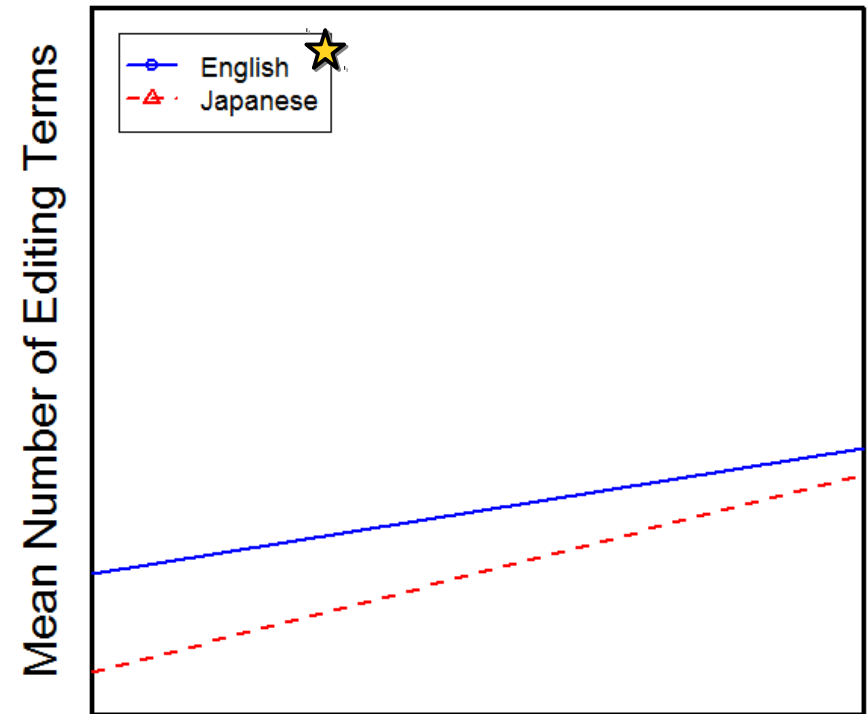
CCHP Results: Other Repair Measures

Mean Num Repair Attempts



L2 Competence

Mean Num Editing Terms



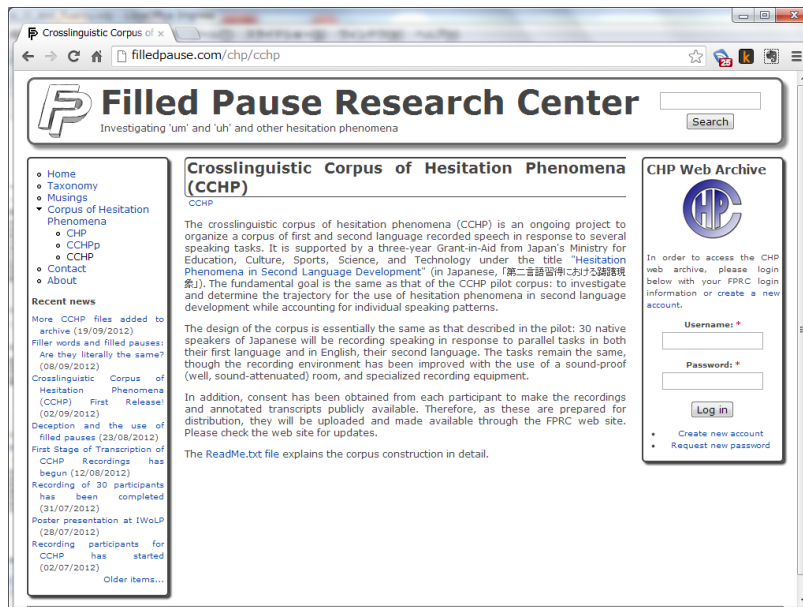
L2 Competence

Summary

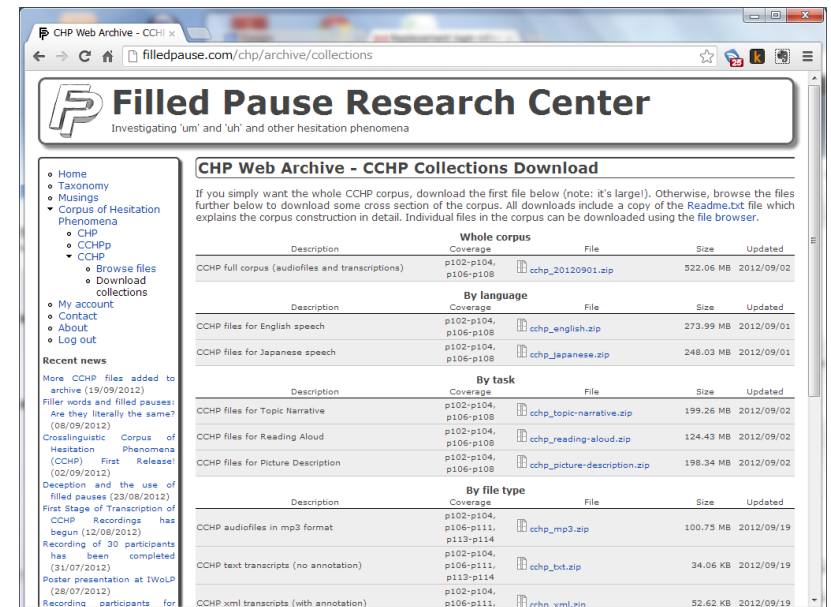
- Recent studies of L2 speech performance are taking L1 speech performance more and more into account.
- The Crosslinguistic Corpus of Hesitation Phenomena allows us to account for L1 factors in the study of L2 hesitation patterns.
- Results show that learners' use of filled pauses change with increased proficiency, independent of L1 speech factors.
- Results show that speakers at all proficiency levels use more repeats.
- Results suggest that other aspects of L2 hesitation use correlate with that of L1.

CCHP Public Corpus

- Assembling a larger (N=30), 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 titled "Crosslinguistic Corpus of Hesitation Phenomena (CCHP)" and contains a paragraph describing the project. Below this, there is a "CCHP Web Archive" section with a login form and a "Log in" button. The left sidebar contains a navigation menu with links to Home, Taxonomy, Musings, Corpus of Hesitation Phenomena, Recent news, and Older items... The right sidebar contains a "CCHP Web Archive" section with a login form and a "Log in" button.



The screenshot shows the "CCHP Web Archive - CCHP Collections Download" page. The header includes the logo and the text "Investigating 'um' and 'uh' and other hesitation phenomena". The main content area is titled "CCHP Web Archive - CCHP Collections Download" and contains a paragraph explaining the download options. Below this, there are three tables: "Whole corpus", "By language", and "By task". Each table lists the description, coverage, file name, size, and updated date. The "Whole corpus" table has one row: "CCHP Full corpus (audiofiles and transcriptions)" with a file size of 522.06 MB and an updated date of 2012/09/02. The "By language" table has two rows: "CCHP files for English speech" (273.99 MB) and "CCHP files for Japanese speech" (248.03 MB). The "By task" table has three rows: "CCHP files for Topic Narrative" (199.26 MB), "CCHP files for Reading Aloud" (124.43 MB), and "CCHP files for Picture Description" (198.34 MB). The "By file type" table has two rows: "CCHP audiofiles in mp3 format" (100.75 MB) and "CCHP text transcripts (no annotation)" (34.06 KB). The "CCHP xml transcripts (with annotation)" row has a file size of 52.62 KB and an updated date of 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.
- de Leeuw, E. (2007). Hesitation markers in English, German, and Dutch. *Journal of Germanic Linguistics*, 19(2), 85-114.
- 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.
- Fox, B., Hayashi, M., and Jasperson, R. (1996). Resources and repair: A cross-linguistic study of syntax and repair. In Ochs, E., Schegloff, E., and Thompson, S. (eds). *Interaction and Grammar*. Cambridge University Press, Studies in Interactional Sociolinguistics, volume 13: pages 185-237.
- 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.
- 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.
- 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.