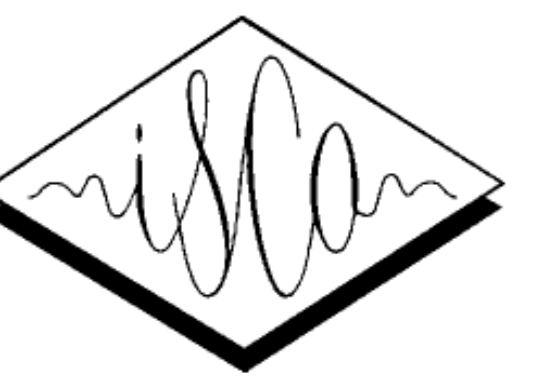


Crosslinguistic Corpus of Hesitation Phenomena:

A corpus for investigating first and second language speech performance

Ralph L. Rose <rose@waseda.jp> Waseda University Faculty of Science and Engineering, Tokyo Japan

INTERSPEECH
LYON 2013

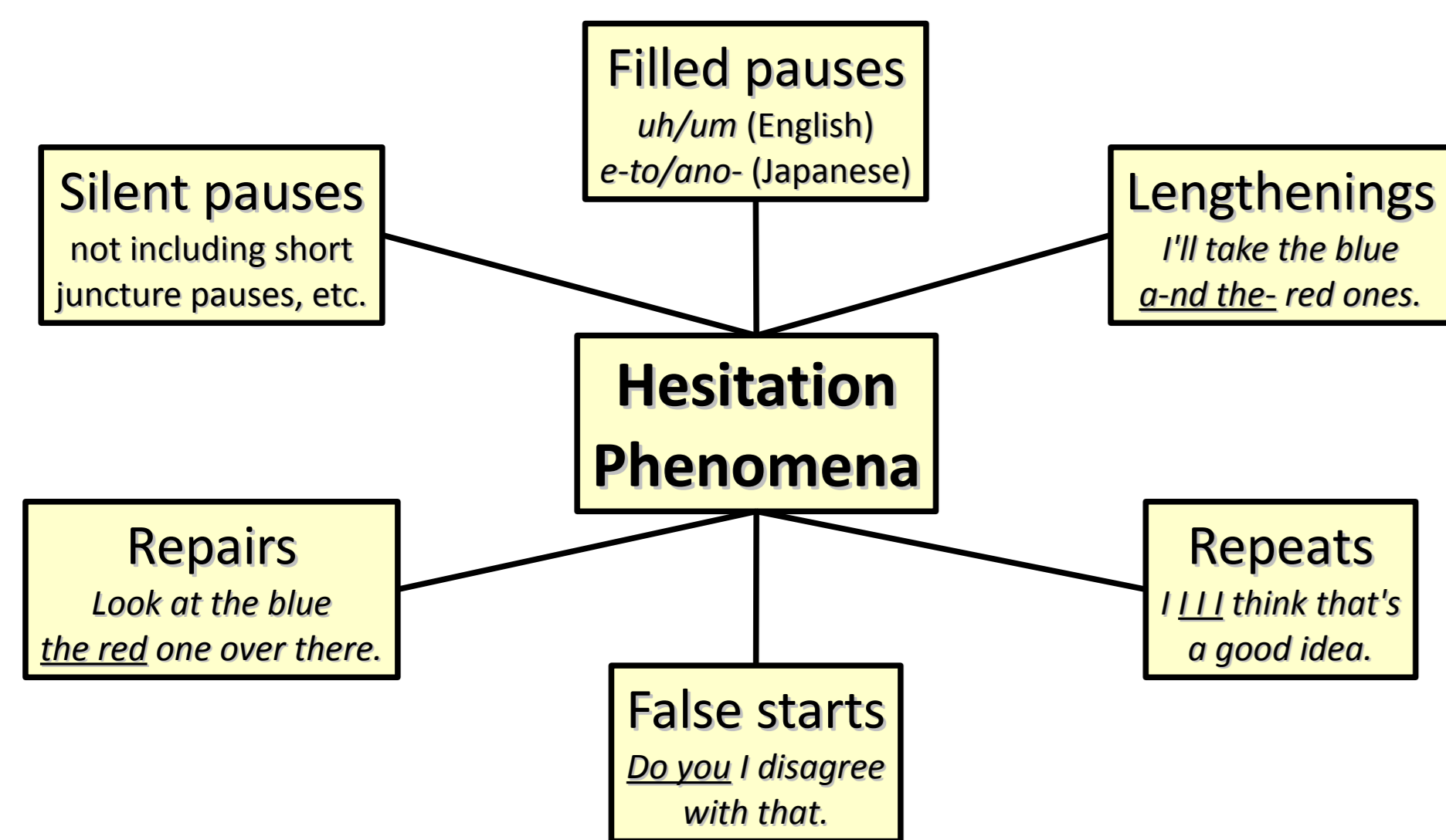


Abstract

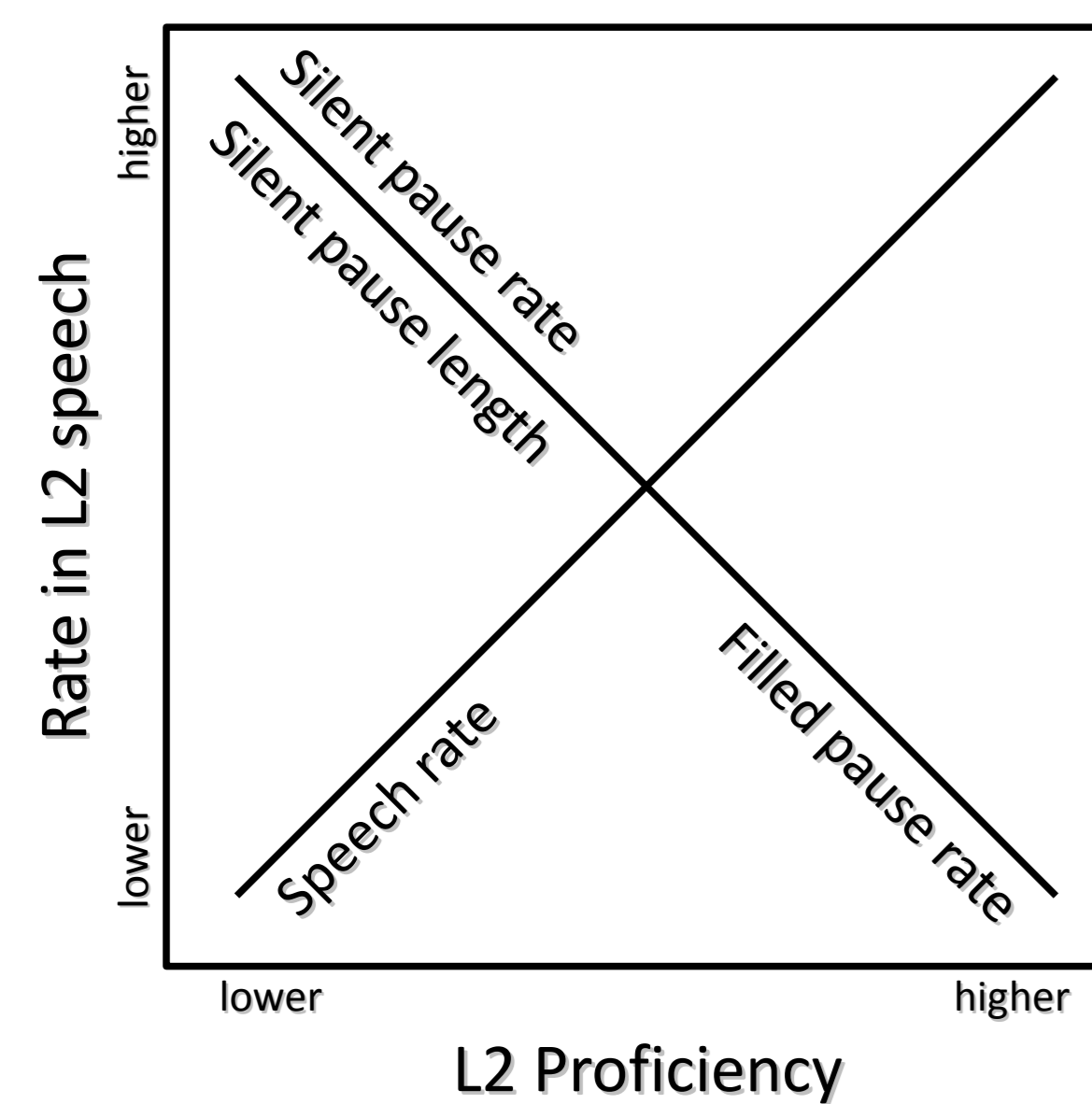
A consensus is forming that there is a need to evaluate second language speech performance with respect to first language speech behavior. To support this need, the Crosslinguistic Corpus of Hesitation Phenomena was developed. This freely available corpus is designed to investigate the crosslinguistic influence of speech patterns and consists of recordings of speakers producing first and second language speech samples in response to parallel elicitation tasks in each language. Preliminary results from the corpus are consistent with other findings that second language performance is sometimes correlated with first language speech behavior [3,7]. In particular, findings show that silent pause rate and duration correlate with first language performance while speech rate and filled pause rate do not. Interestingly, repeats also differ from first language production. Results show that the corpus may be a useful tool for researchers who wish to investigate the correspondence between first and second language speech, particularly with respect to the use of hesitation phenomena.

Background

Hesitation phenomena [9,11] may be categorized as follows.



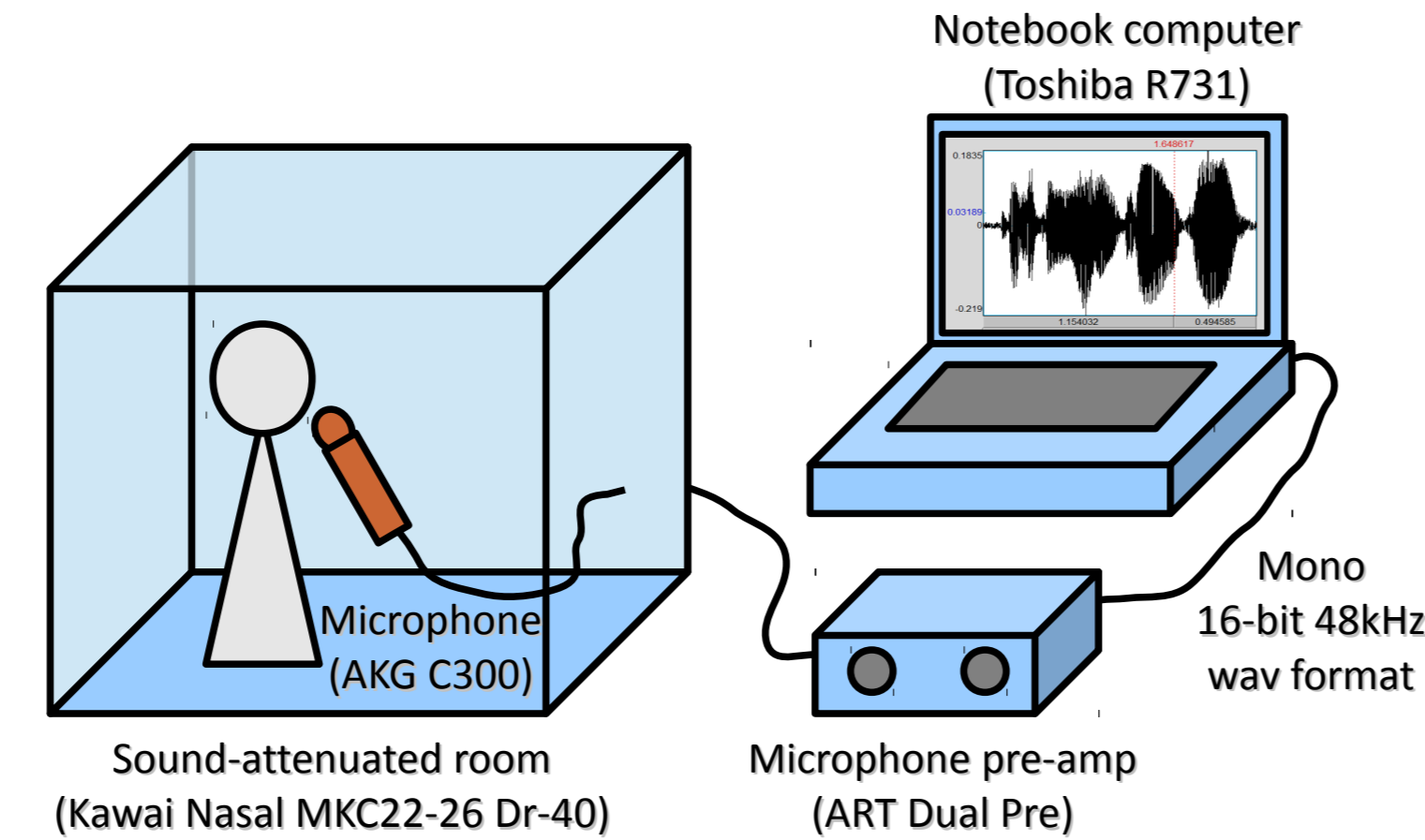
Studies of hesitation phenomena in second language (L2) speech production have observed several correlations with L2 proficiency [5, 13,14,16,17].



One problem with many of these studies is they do not account for first language (L1) speech behavior (cf., [6]): A learner who pauses frequently in L1 might do so in L2, regardless of proficiency. The goal of this study is to provide a resource for the study of L2 hesitation patterns with respect to L1 speech.

Corpus Design

Recording

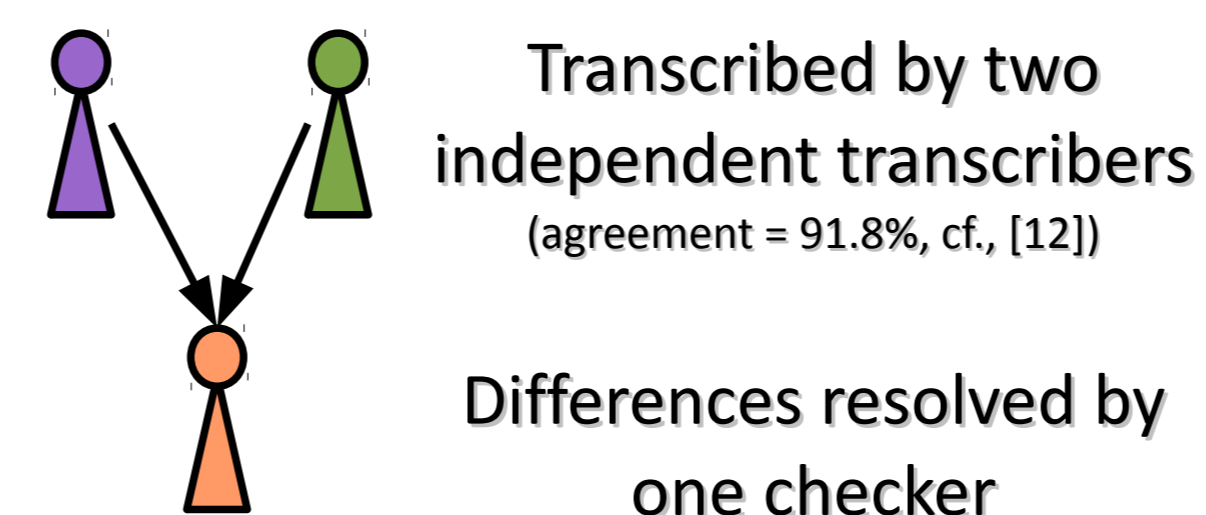


	Reading aloud	Picture description	Topic narrative
Japanese (L1)	"The farm script" (Japanese translation)	Various b/w pictures and cartoon strips	Explain what a television is to someone who has never seen one.
English (L2)	"The farm script" [4]	Various b/w pictures and cartoon strips (different from those in L1)	Explain basketball to someone who has never seen the sport

Participants (n=35) were recorded for 3-4 minutes in each task.

Transcription

- Whole and partial word tokens
- Filled pauses
- False starts
- Repair sequence structure: reparandum, editing terms, and repairs (cf., [10], [15])
- Other minor audible phenomena: e.g., coughs, throat-clearing



Pause and word interval durations were detected using Praat [2]. Transcripts are stored in XML format.

```

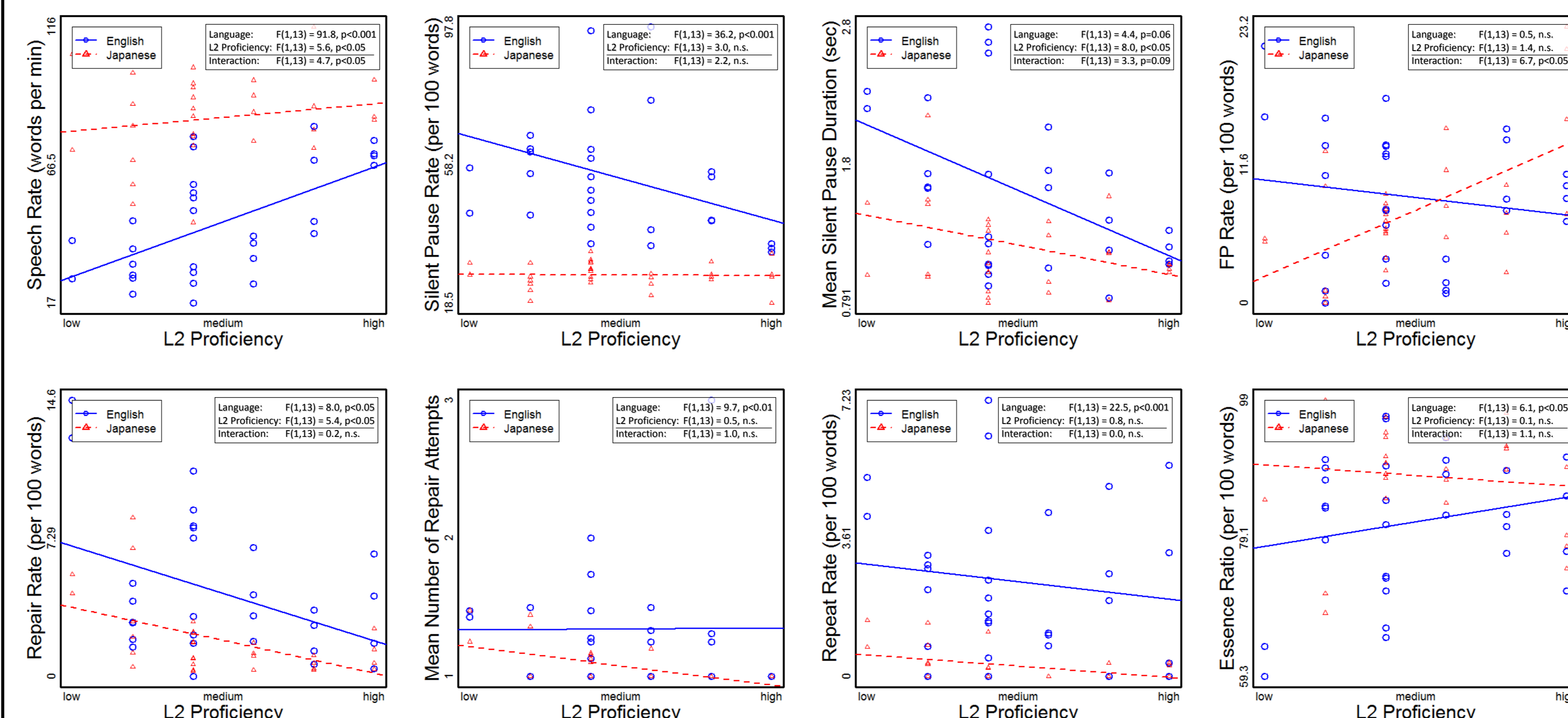
<UTTERANCE>
<T>in</T>
<T>America</T>
<T FILLED-PAUSE="yes">uh</T>
<T>there's</T>
<T>a</T>
<T FILLED-PAUSE="yes">uh</T>
<T>very</T>
<T>famous</T>
<T FILLED-PAUSE="yes">uh</T>
<T>and</T>
<T>loved</T>
<T FILLED-PAUSE="yes">uh</T>
<T>basketball</T>
<RP>
<O>
<T>c1#</T>
</O>
<T FILLED-PAUSE="yes">uh</T>
<E>
<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>
  
```

Results

The transcription process is ongoing. Results presented here represent n=18 native Japanese participants.

	Reading aloud	Picture description	Topic narrative	Total
Japanese	4,246 words 31.1 min.	4,375 words 56.6 min.	5,086 words 56.3 min.	12,707 words 144.0 min.
English	4,897 words 39.4 min.	2,960 words 61.9 min.	2,637 words 58.2 min.	10,494 words 160.0 min.

	Japanese	English
Silent pauses	3,106	3,841
Filled pauses, total	742	535
Open type (uh)	572	324
Closed type (um)	170	211
Repair sequences	231	348
Repeats	28	149



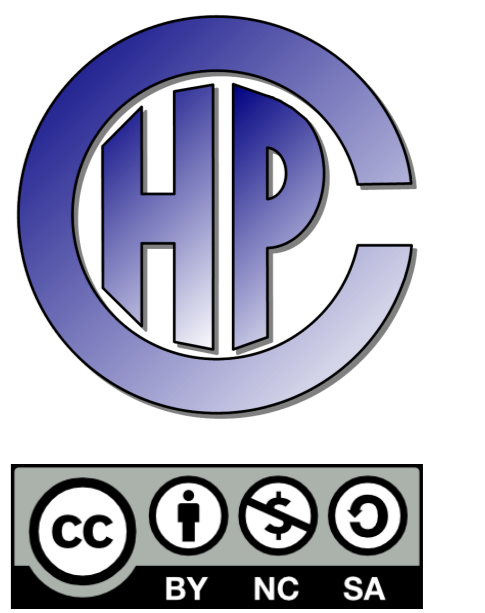
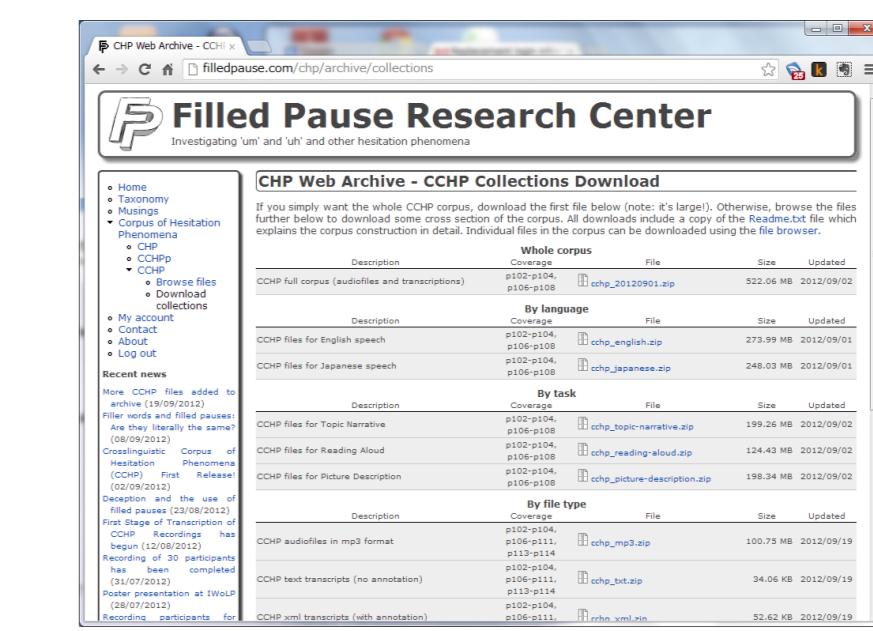
Results show that speech rate and filled pause rate in L2 speech are not correlated with L1 speech patterns (contra [3,7]) while silent pause rate and duration are correlated (consistent with [3,7]). Furthermore, results show clear linguistic differences in the use of repairs, repeats (rare in Japanese [8]), and overall use of hesitation phenomena, but not in ways that are indicative of L2 proficiency development.

Conclusions

The preliminary results from the Crosslinguistic Corpus of Hesitation Phenomena show that it can be a useful tool to investigate the relationship between first language speech behavior and second language speech performance, by evaluating the latter with respect to the former on an individual basis. Current results suggest that as second language learners develop higher proficiency in the second language, they speak faster and use fewer filled pauses in a manner that is not necessarily related to their first language rate of speech. On the other hand, their use of other hesitation phenomena—including silent pauses—as their proficiency develops may be more closely related to their first language speech patterns. Future work includes annotation of discourse and clause structure, part-of-speech mark-up, and syllable detection for the purpose of examining lengthenings.

Corpus Distribution

The recordings and transcripts are freely available via an online archive (<http://filledpause.com/chp/cchp>) under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. Teachers and researchers may make use of the corpus for research and educational purposes.



References

- Alexander, L. G., First Things First, Longman, 1967.
- Boersma, P., "Praat, a system for doing phonetics by computer", Glot International, 5(9/10):341-345, 2010.
- Cox, T. and Baker-Smemeo, W., "The relationship between L1 fluency and L2 fluency across different proficiency levels and L1s", presentation at Workshop Fluent Speech (Utrecht University, The Netherlands), 2012.
- Crystal, T. H., & House, A. S., "Segmental durations in connected speech signals: Preliminary results", Journal of Acoustical Society of America, 72(3):705-716, 1982.
- Cucchiarini, C., van Doremalen, J. and Strik, H., "Fluency in non-native read and spontaneous speech", Proceedings of Disfluency in Spontaneous Speech (DiSS) and Linguistic Patterns in Spontaneous Speech (LPS) Joint Workshop, pages 15-18, 2010.
- de Leeuw, E., "Hesitation markers in English, German, and Dutch", Journal of Germanic Linguistics, 19(2):85-114, 2007.
- Derwing, T. M., Munro, M. J., Thomson, R. I. and Rossiter, M. J., "The relationship between L1 fluency and L2 fluency development", Studies in Second Language Acquisition, 31(4):533-557, 2009.
- Fox, B., Hayashi, M., and Jasperson, R., "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, 1996.
- Goldman-Eisler, F., "A Comparative Study of Two Hesitation Phenomena", Language and Speech, 4(1):18-26, 1961.
- Levelt, W. J. M., "Monitoring and self-repair in speech", Cognition, 14(1):41-104, 1983.
- MacLay, H. and Osgood, C., "Hesitation Phenomena in Spontaneous English Speech", Word, 15:19-44, 1959.
- Patterson, B. R., Neupauer, N. C., Burant, P. A., Koehn, S. C., and Reed, A. T., "A preliminary examination of conversation analytic techniques: Rates of inter-transcriber reliability", Western Journal of Communication, 60(1):76-91, 1996.
- Riazantseva, A., "Second language proficiency and pausing: A study of Russian speakers of English", Studies in Second Language Acquisition, 23(4):497-526, 2001.
- Rieger, C. L., "Disfluencies and hesitation strategies in oral L2 tests", Proceedings of DiSS '03, Disfluency in Spontaneous Speech Workshop, 2003.
- Shriberg, E., Preliminaries to a theory of speech disfluencies, doctoral dissertation, University of California, Berkeley, 1994.
- Tavakoli, P., "Pausing patterns: differences between L2 learners and native speakers", ELT Journal, 65(1):71-79, 2011.
- Trofimovich, P. and Baker, W., "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, 2007.

Acknowledgments

This research has been funded 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). I am also grateful to Hiroaki Suzuki, Junichi Inagaki, Masayuki Motoori, Yukikatsu Fukuda, and Tatsuhiro Nomaguchi for their work as research assistants.