Sensitivity of pausing strategies to utterance length: Evidence from a crosslinguistic speech corpus

Ralph L. ROSE

e-mail: rose@waseda.jp Twitter: @fildpauz

Center for English Language Education (CELESE)

Faculty of Science and Engineering

Waseda University; Tokyo, Japan

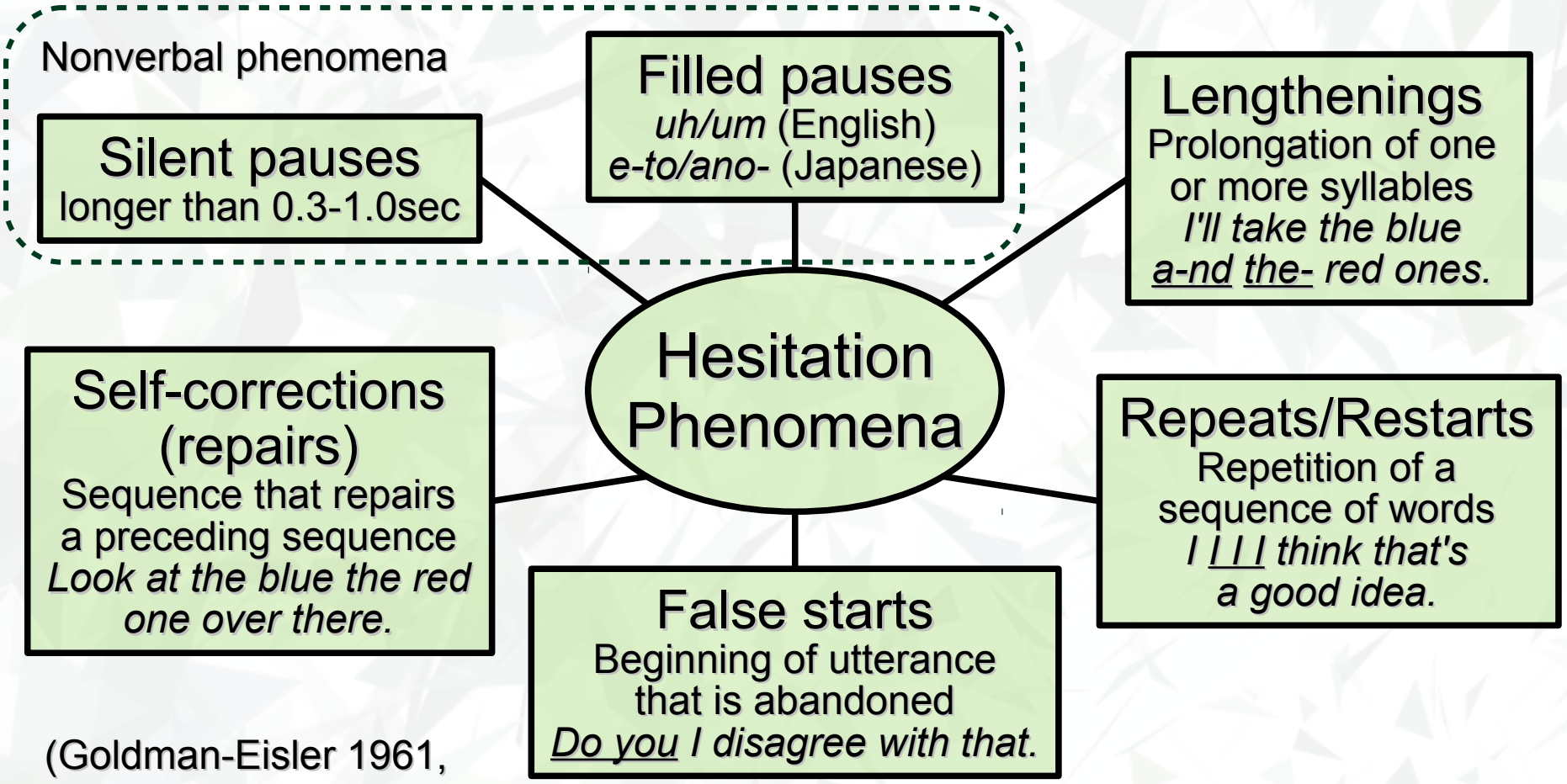
Acknowledgments: Michiko Watanabe, Hinako Masuda, Hiroaki Suzuki, Junichi Inagaki, Masayuki Motoori, Yukikatsu Fukuda, Tatsuhiro Nomaguchi, Aiko Ooe, & Maiko Serizawa

Supported by Japan Society for the Promotion of Sciences (JSPS) Grants-in-Aid #15K02765 (PI, Ralph Rose), #26284062 (PI, Kikuo Maekawa)

8th International Conference in Corpus Linguistics (CILC)

Malaga, Spain; 2-4 March 2016

When something goes wrong during spontaneous speech production...



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

Recent work on the distribution of silent and filled pauses

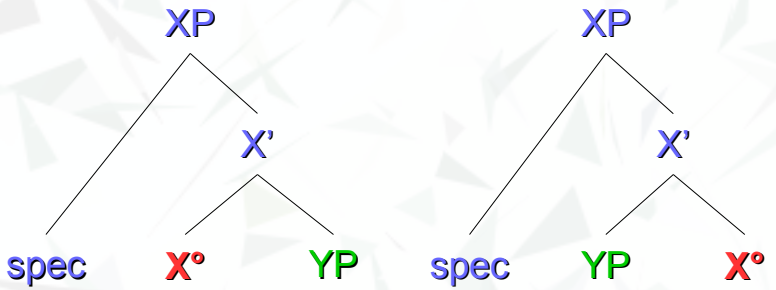
- Campione and Véronis 2005: Examined SPs and FPs in the Spoken French Reference Corpus
  - SPs and FPs overlap to mark syntactic structure
  - FPs (with following SPs) mark non-syntactic interruptions
- Watanabe, et al 2013, 2014, 2015: Examined clause boundary depth and clause length in Japanese
  - 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.

Open questions

- What are the triggering contexts for SPs and FPs?
- (i.e.,) What constitutes linguistic complexity?
  - Utterance length
  - Clause length
  - Clause depth
  - Other?
- How does this affect L2 versus L1 speech production?

Methodological questions

- Isn't this psycholinguistics? Why corpus linguistics?
  - More and more psycholinguists are finding it useful to evaluate hypotheses through corpus analysis (cf., Yamashita et al 2011)
- Why English and Japanese?
  - Short phrases vs. long phrases (Tanaka et al 2011)
    - English: Short before long
    - Japanese: Long before short
  - Head direction
    - English: head-initial
    - Japanese: head-final

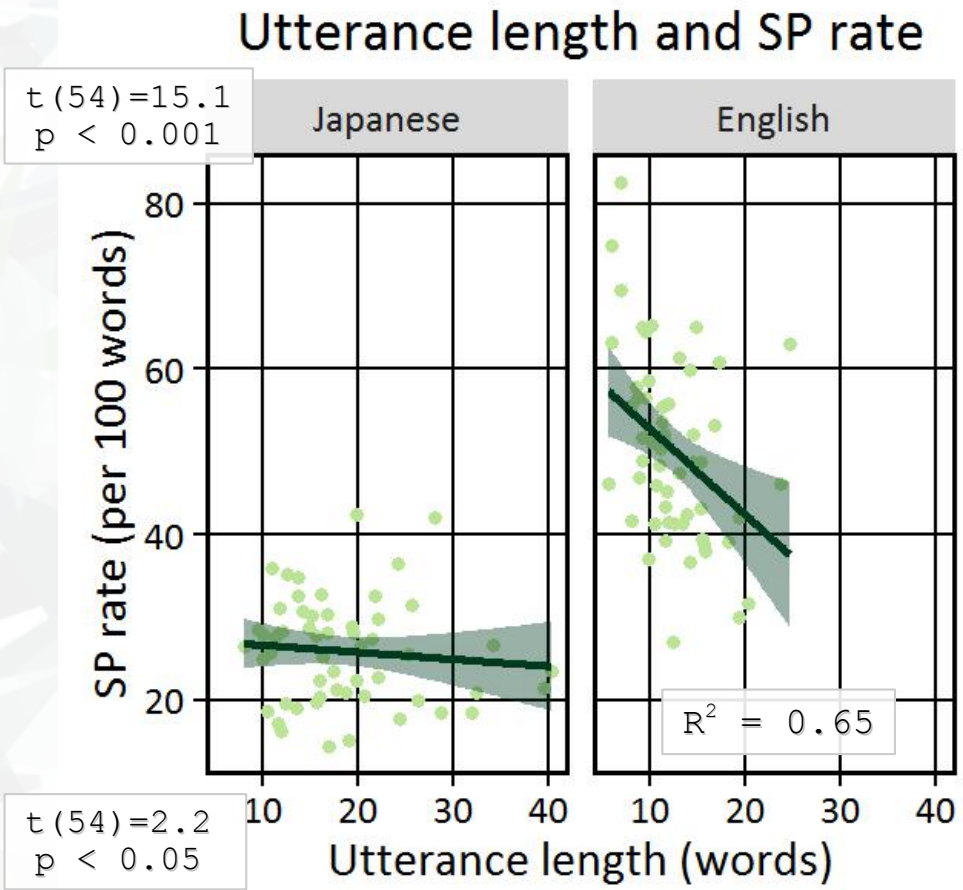


Crosslinguistic Corpus of Hesitation Phenomena

- CCHP (Rose 2013)
- Participants: L2 learners of varying proficiency levels
- Elicitation tasks (both L1 and L2)
  - Spontaneous speech: picture description, topic narrative
  - Reading aloud
- Annotation
  - Transcript with filled pauses
  - Two annotators, one checker
  - Pause info (Quené et al 2011)

```
<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>
```

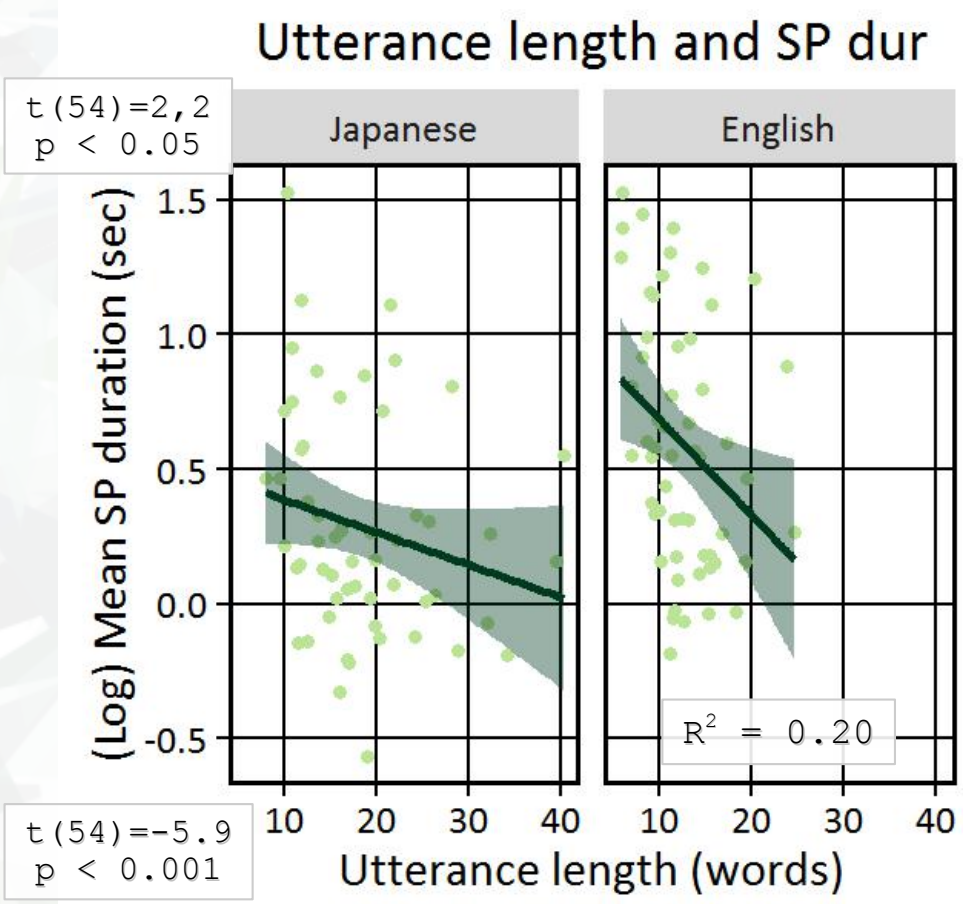
Influence of utterance length in L1 and L2



- Clear difference between Japanese and English speech.
- But,
  - Is utterance length just proxy for L2 proficiency?
  - Perhaps not: alternate model not significant

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

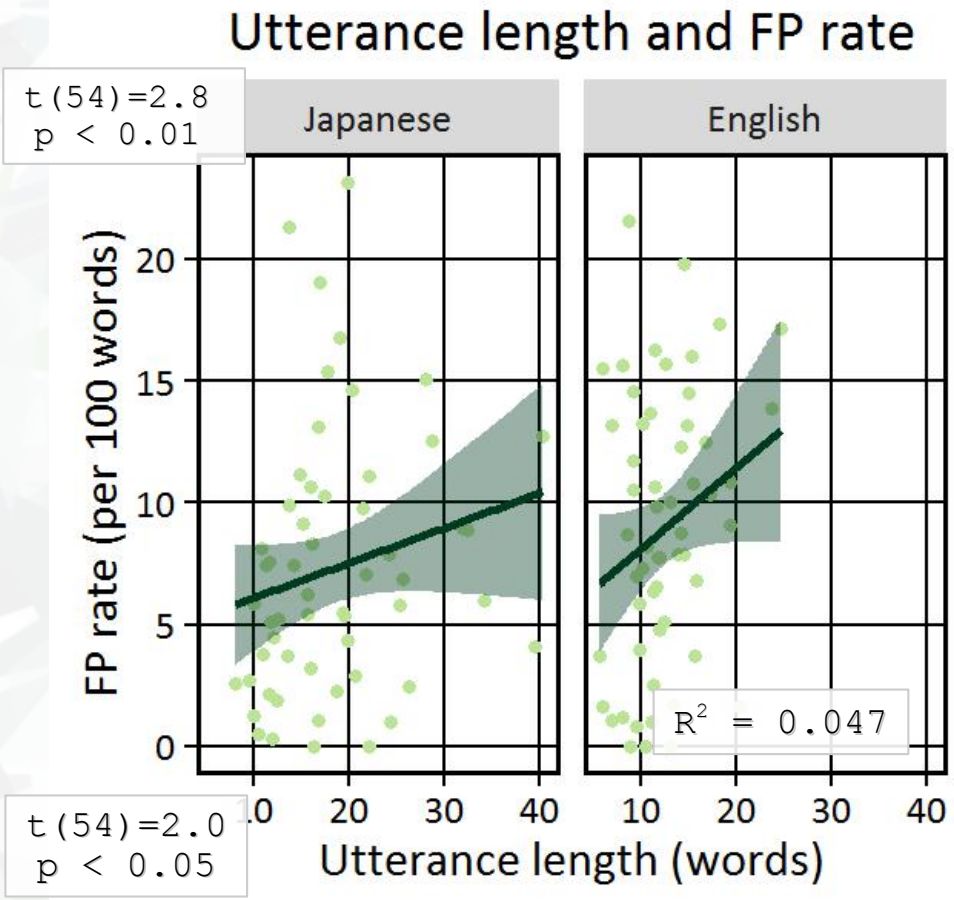
Influence of utterance length in L1 and L2



- In both L1 Japanese and L2 English, the longer the utterance, the shorter the mean silent pause duration.
  - Japanese data contrasts Watanabe et al 2015
  - Is utterance length just proxy for L2 proficiency?
  - Perhaps not: alternate model not significant

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

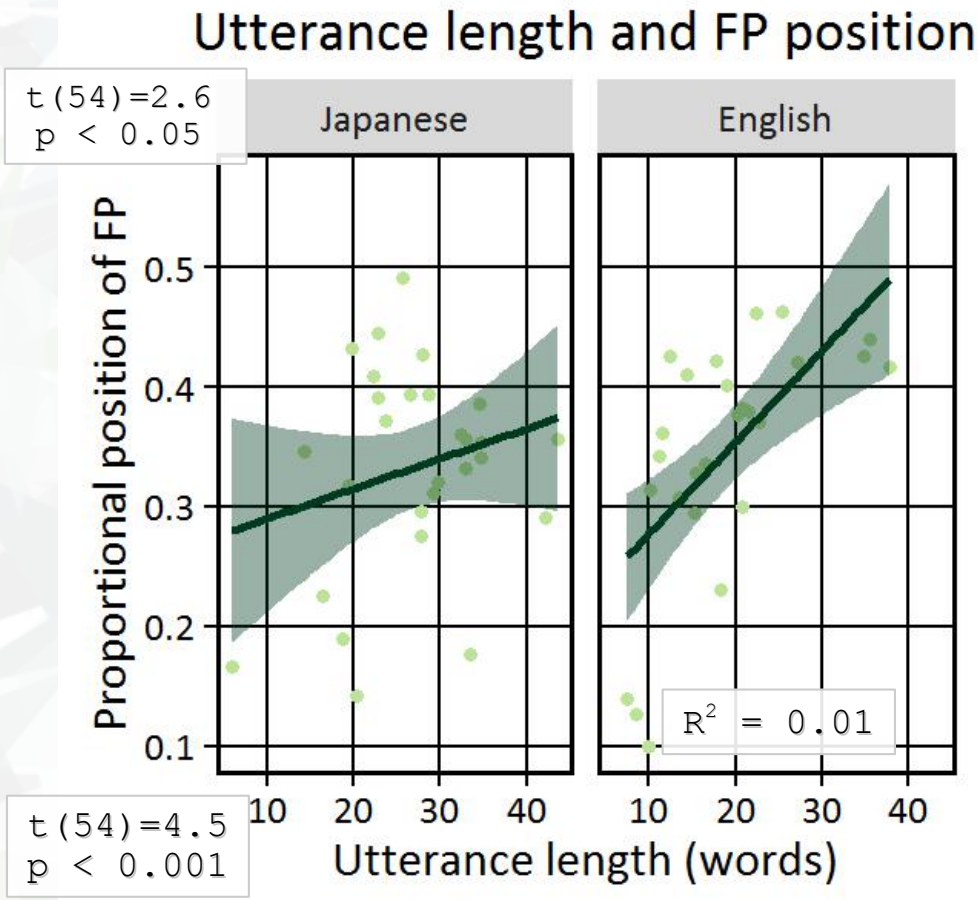
Influence of utterance length in L1 and L2



- In L1 Japanese, speakers' FP rate increases with longer utterances
- This trend is not significant in English.

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

Influence of utterance length in L1 and L2



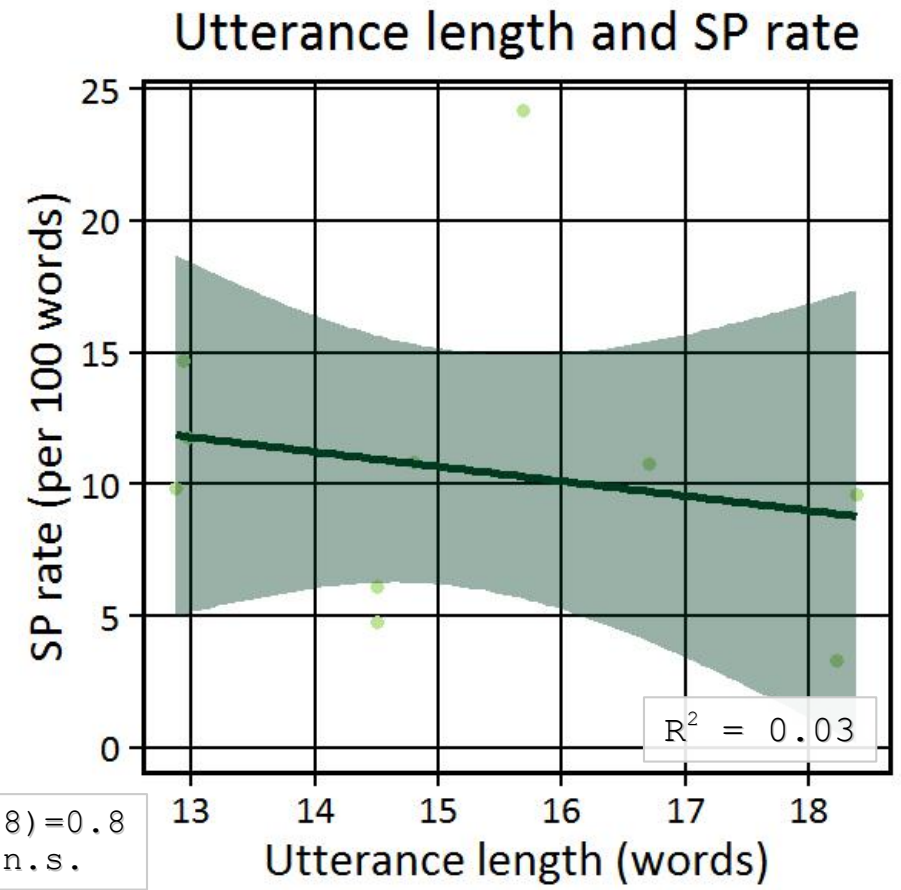
- In both L1 Japanese and L2 English, the longer a speakers' utterance (in words), the later (proportionally) that a FP is used in the utterance.
- This trend is greater in L2 English than in L1 Japanese.

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

Corpus of Oral Presentations in English

- COPE (Watanabe)
  - Complement to the Corpus of Spontaneous Japanese (CSJ: Maekawa 2003)
- Participants: 20 native speakers of English
- Elicitation task: Unscripted monologue
  - Approx. 10 min. oral presentation with 10 min. preparation
- Annotation
  - Transcription with filled pauses
  - Shallow parse (utterance and clause boundaries)

Influence of utterance length in English

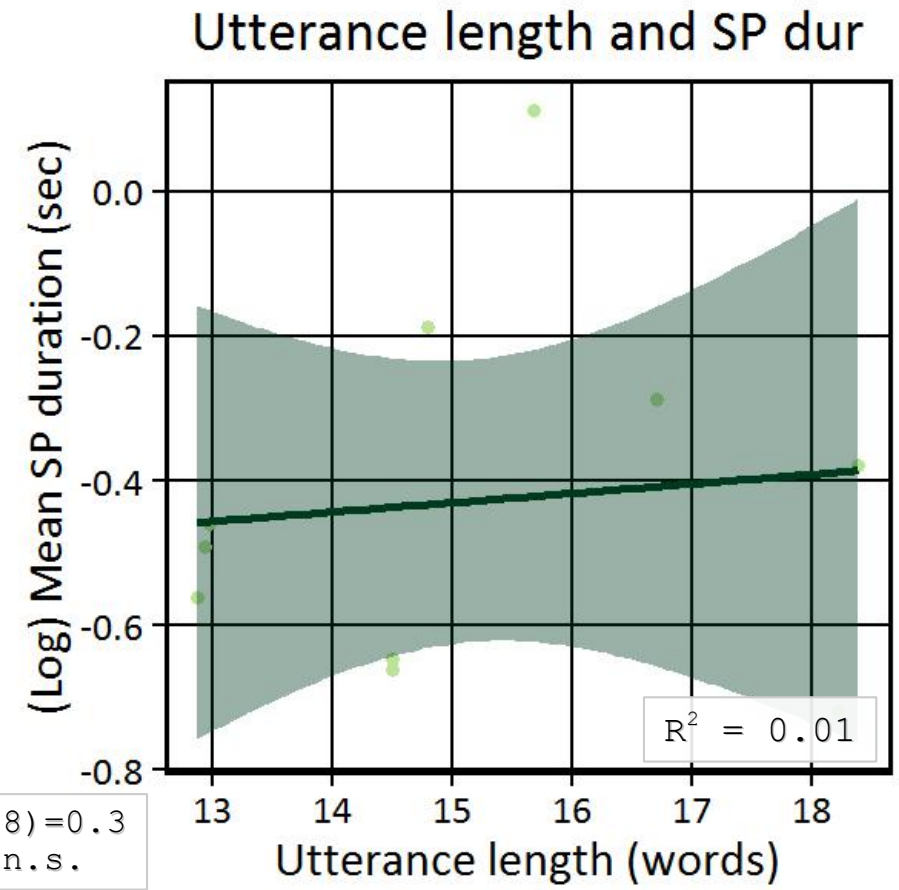


t(8)=0.8
n.s.

- No apparent relationship between utterance length and occurrence of SP.
- (But comparing broad means only)

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

Influence of utterance length in English

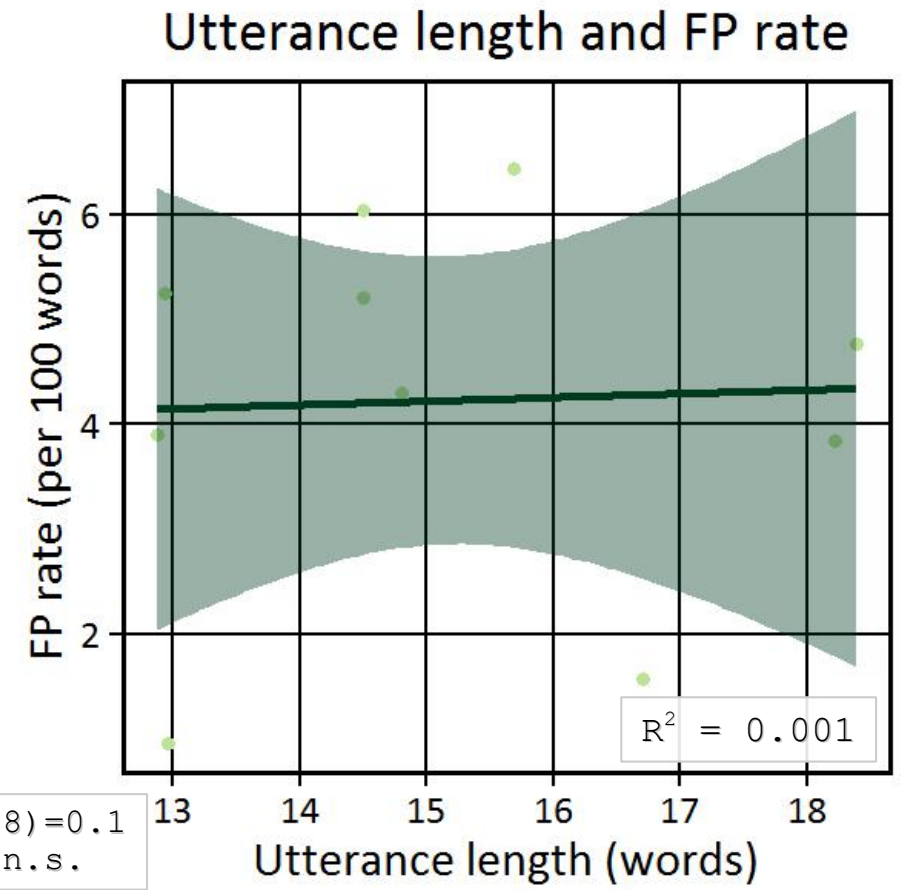


t(8)=0.3
n.s.

- No apparent relationship between utterance length and duration of SP.
- (But comparing broad means only)

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

Influence of utterance length in English

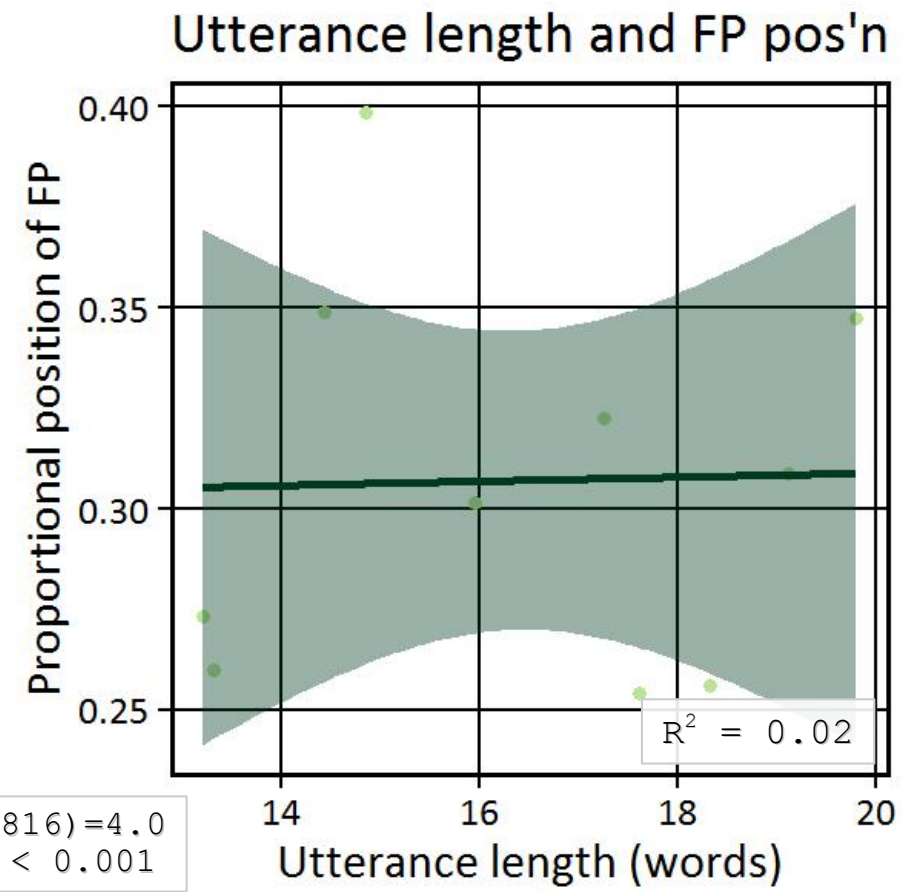


t(8)=0.1
n.s.

- No apparent relationship between utterance length and occurrence of FP.
- (But comparing broad means only)

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

Influence of utterance length in English



t(816)=4.0
p < 0.001

- The longer the utterance, the later the FP will be used.

All stats using lme in R; fixed effects as shown and participants as random effects. T-values for relevant fixed effect(s) and marginal R^2 are shown.

Summary

	SP rate wrt Utterance length	SP dur wrt Utterance length	FP rate wrt Utterance length	FP pos'n wrt Utterance length
L1 English	n.s.	n.s.	n.s.	+
L1 Japanese	n.s.	-	+	+
L2 English	-	-	n.s.	+

As L2 learners become more proficient in the target language, their pausing rates and durations as well as utterance length more closely resemble those of natives. However, the interaction of these features – i.e., their pausing strategies – may still show a very different production profile.

Future work

- Examine FP durations.
- Examine fine-grained placement of SPs.
- Examine influence of clause-related complexity.

References

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

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

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.

Maekawa, K. (2003) "Corpus of Spontaneous Japanese: Its Design and Evaluation", Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003), 4, 7-12.

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.

Rose, R. (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), Lyon, France, pp. 992-996.

Tanaka, M.N., Branigan, H.P., and Pickering, M.J. (2011). The production of Head-Initial and Head-Final Lanaguages. In Yamashita, Hirose, and Packard [Eds], pp. 113-129.

Watanabe, M. Corpus of Oral Presentations in English (COPE). Unpublished corpus data.

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).

Yamashita, H., Hirose, Y., Packard, J. (Eds.) (2011). *Processing and Producing Head-final Structures*. Dordrecht: Springer.