Differences in second language speech fluency ratings:
Native versus nonnative listeners

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

Fluency & Disfluency Across Languages and Language Varieties
Louvain-la-Neuve, Belgium; 15-17 February 2017

15-17 February 2017 Louvain-la-Neuve Belgium
International Conference
Fluency & Disfluency
Across Languages and Language Varieties

Crosslinguistic communication
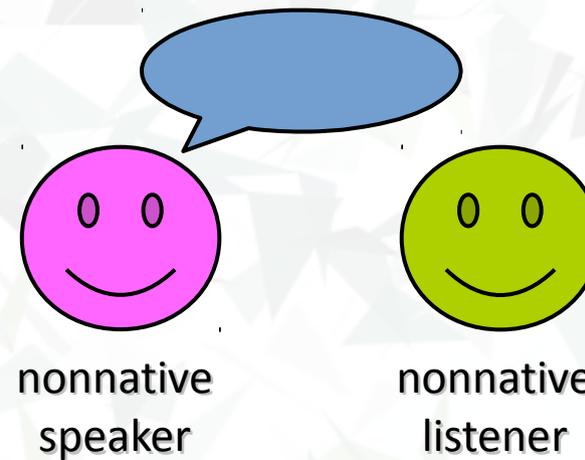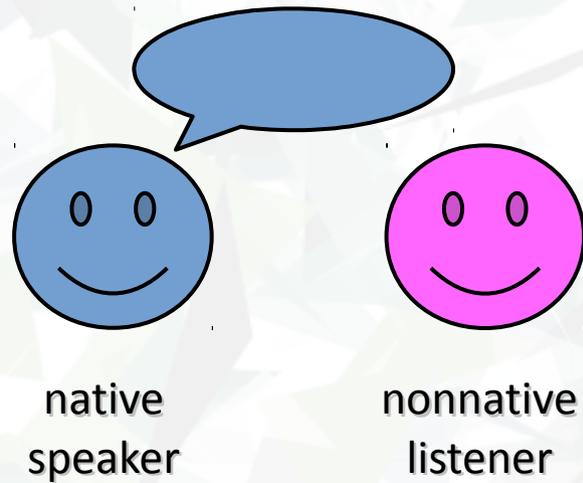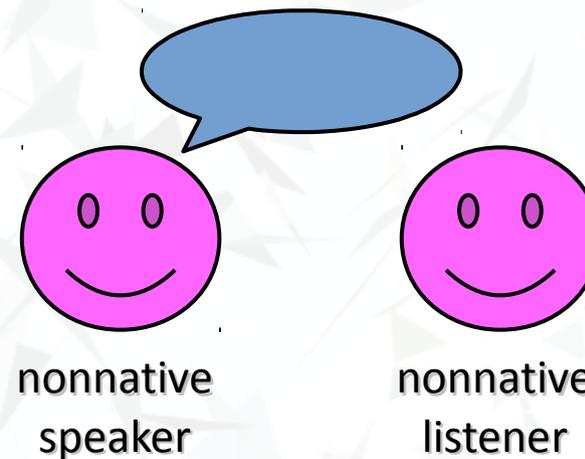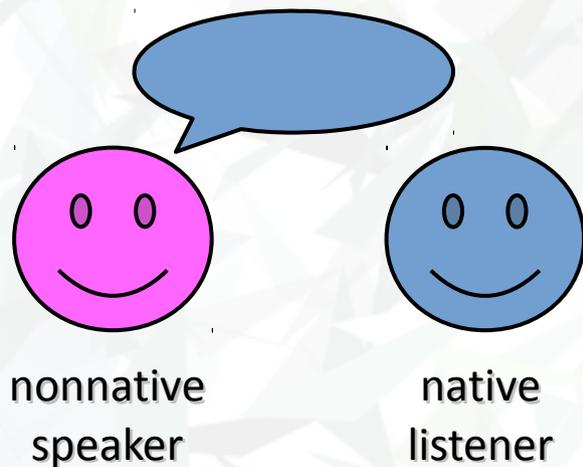
Overview

- Background
  - Crosslinguistic speech perception
- Method
  - Crosslinguistic Corpus of Hesitation Phenomena
  - Fluency ratings
    - by native listeners
    - by nonnative listeners
- Results
- Discussion
- Accessing the CCHP

Crosslinguistic speech perception

- Perception of nonnative speech modulated by listener's language background
  - Wester and Mayo 2014 – nonnative listeners judge accentedness more harshly than native listeners
  - Bent and Bradlow 2003 – nonnative listeners comprehend better than native listeners
- Crosslinguistic perceptions of fluency, too
  - Rossiter 2009
    - Fluency ratings: nonnative listeners < native listeners
    - Both native/nonnative listeners' ratings correlated with articulation rate and pause frequency
  - Foote and Trofimovich 2016 – native listeners attend to pause frequency; nonnative listeners to speech rate

nonnative speaker

nonnative listener

nonnative speaker

nonnative listener
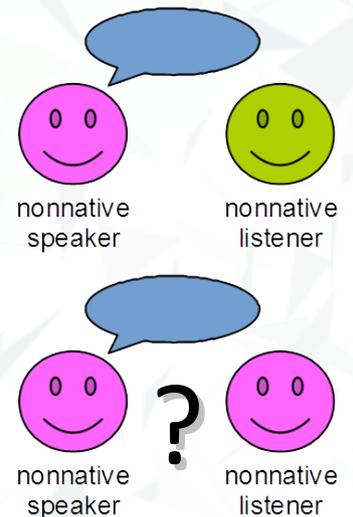
Research questions

- Do native listeners and nonnative listeners rate the fluency of nonnative speech differently?
  - Particularly when the nonnative listeners share the same native language as the speakers?
- What temporal features of nonnative speech do native and nonnative listeners each attend to in their judgments of fluency?

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 FPs, repairs, etc.
  - Two annotators, one checker
  - Temporal measurements (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>
```

CCHP: Basic Statistics

- Participants: 35 Japanese L1 / English L2 speakers

	Word count	Time
Read speech	21,406	2 hr, 41 min
Picture description	19,732	4 hr, 39 min
Topic narrative	21,138	4 hr, 35 min
Total	62,276	11 hr, 56 min

Transcriber agreement = 91.5%

Hesitation phenomena

- 15,480 silent pauses
- 3,741 filled pauses
- 1,635 repairs
- 566 repeats
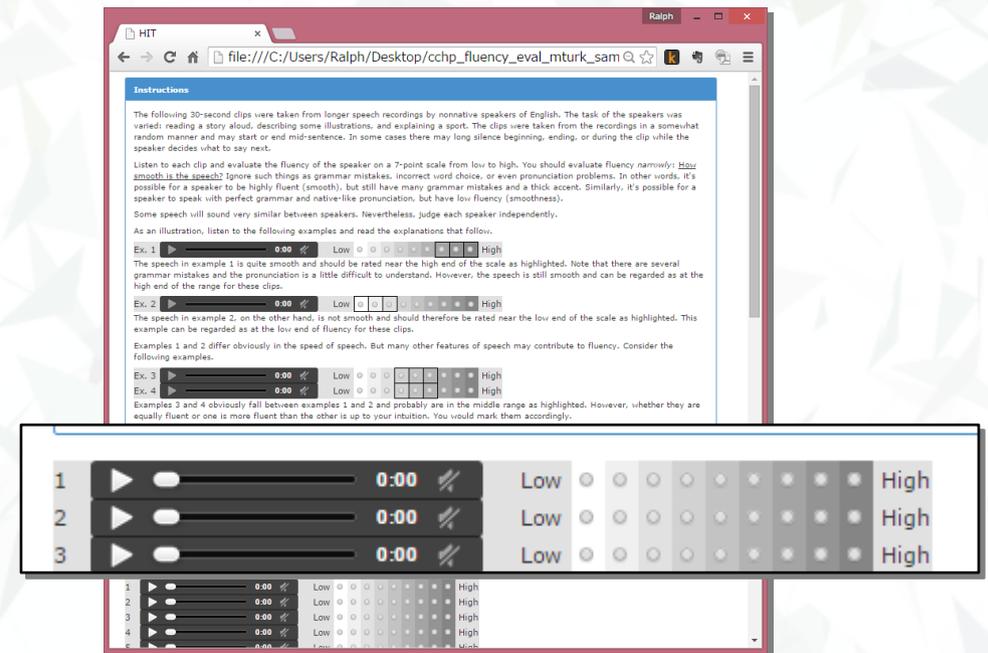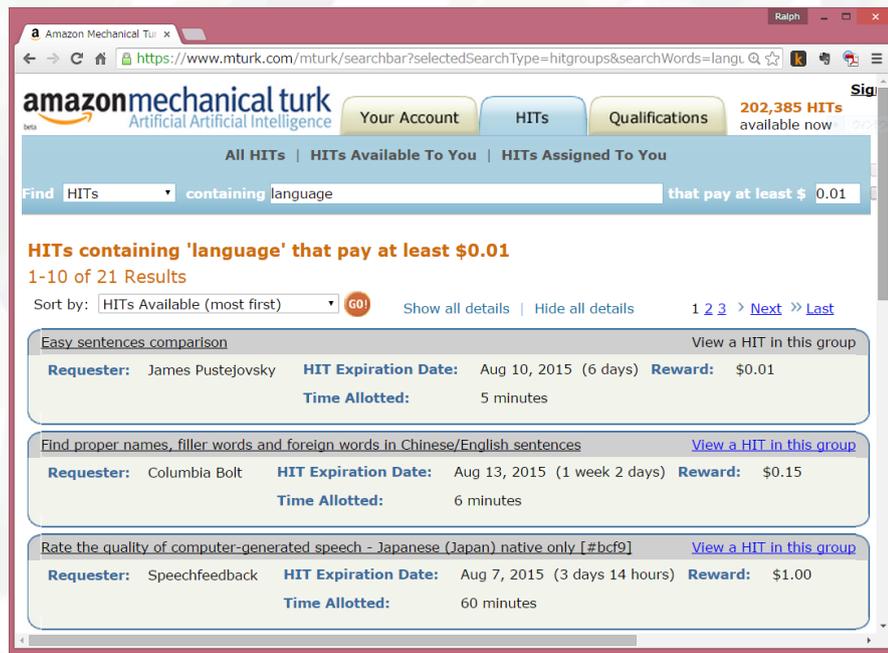
Fluency ratings

- Extracted 7 30-second clips of English speech per speaker
  - Reading aloud x 1
  - Picture description x 3
  - Topic narrative x 3
- Rater instructions
  - Rate fluency on 9-point scale (1 – low … 9 – high)
  - Rate "smoothness" of the speech
  - Ignore pronunciation, grammar, word-choice, etc.

Fluency ratings by native listeners

- Obtained via Amazon Mechanical Turk
- Workers limited to native English speakers (self-reported)
- Used attention checks and monitoring of audio player activity to check that instructions were followed.

(Previously reported in part in Rose 2015)

Fluency ratings by nonnative listeners

- Obtained via local web pages in computer lab
- Participants limited to native Japanese speakers (recruited)
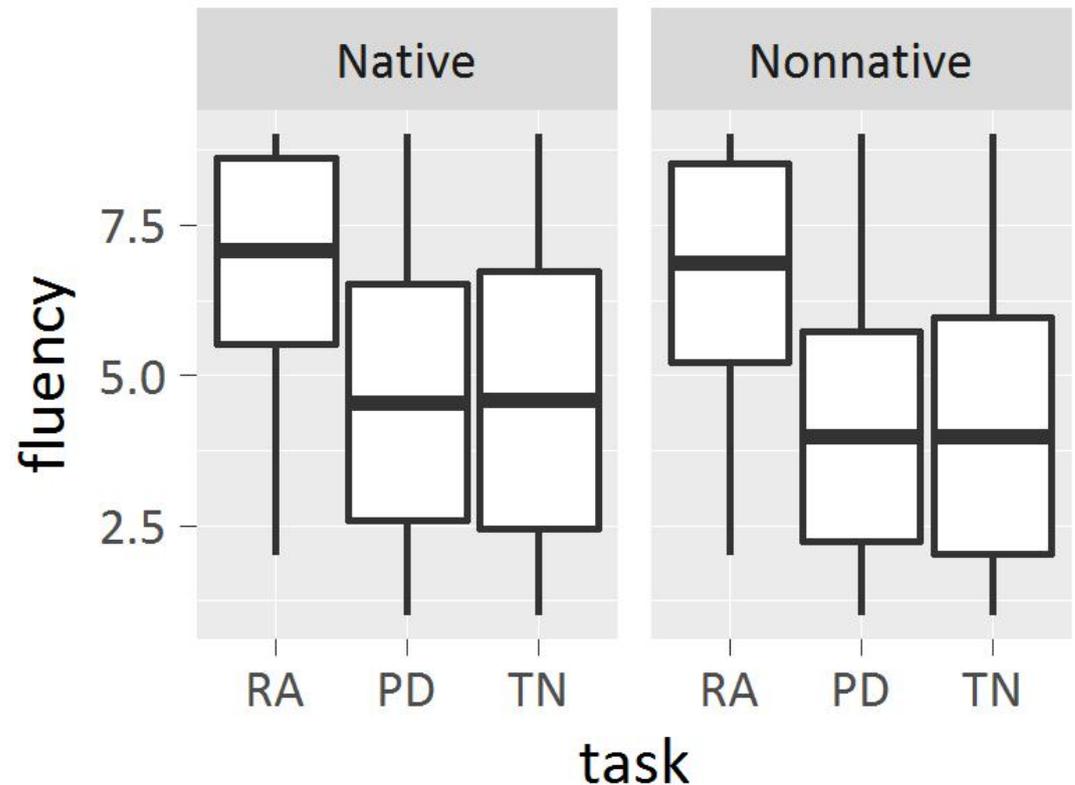- Used attention checks and monitoring of audio player activity to check that instructions were followed.

Results

- Participants: 34 native English raters; 20 native Japanese raters
- Nonnative (Japanese) raters judge fluency lower than native (English) raters (similar to previous work)
- Reading aloud judged more fluent than other tasks

Effect	DFn	DFd	F	p	p<.05
raterlang	1	52	7.581507	8.104075e-03	*
task	2	104	485.145647	1.850916e-53	*
raterlang:task	2	104	2.312539	1.040802e-01	

Results

- Linear regression modeling (using `lm` in `R`)
  - Dep. variable: fluency rating
  - Ind. variables: articulation rate, pause rate, pause length, filled pause rate, repair rate, rater's listener status, speech task
- With full model, all variables significant except task
- But by rater group, relevant temporal features are different

Higher fluency associated with	native raters	nonnative raters
• higher articulation rate	✓	✓
• lower pause rate		✓
• shorter pause length	✓	✓
• *higher* filled pause rate		✓
• lower repair rate	✓	

$[F(7,406)=64.9,\ p<0.001;\ \text{adjusted } R^2=0.52]$ $\qquad [R^2=0.45]\ [R^2=0.57]$

Discussion

- Native and nonnative raters both attend to articulation rate and silent pause length.
- Native raters do not attend to silent pause rate
  - Native listeners have comprehension disadvantage (Bent & Bradlow 2003)
  - Increased silent pauses may aid native listeners' comprehension, …
  - But they may not be perceptually salient
- Nonnative raters do not attend to repair rate
  - Recognition of repairs requires higher syntactic competence.
  - Some participants may lack this competence.
 - Hence, they do not notice repairs.

Discussion

- Native raters do not attend to filled pause rate

- Nonnative raters evaluate filled pauses more highly.

- Compare to other advantages for filled pauses in speech
  - Cues to greater complexity (Watanabe et al 2008)
  - Feeling of another's knowing (Brennan and Williams 1995)

- Filled pauses are more frequent in L1 Japanese than in L1 English (Watanabe and Toyama 2016)

- Perhaps raters regarded a higher filled pause rate as more natural, and hence "smoother".

Summary

- Nonnative listeners who share native language with speaker judge fluency more harshly than do native listeners.
- Nonnative and native listeners both judge fluency based on articulation rate and silent pause length.
  - Native listeners also attend to repair rate.
  - Nonnative listeners attend to silent pause rate, and
  - Judge a higher filled pause rate more positively.
- Differences between listener groups may stem from different processing difficulties and language backgrounds.
- Yet undetermined is whether the non-native-language-sharing listeners would perform differently. → Future work

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
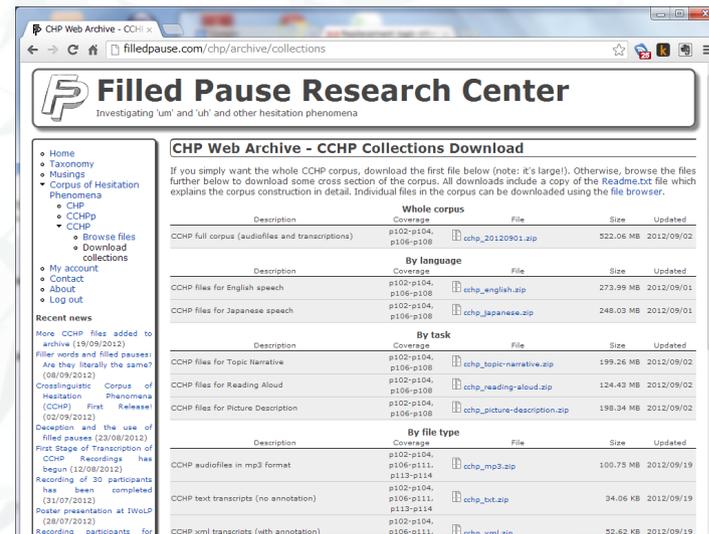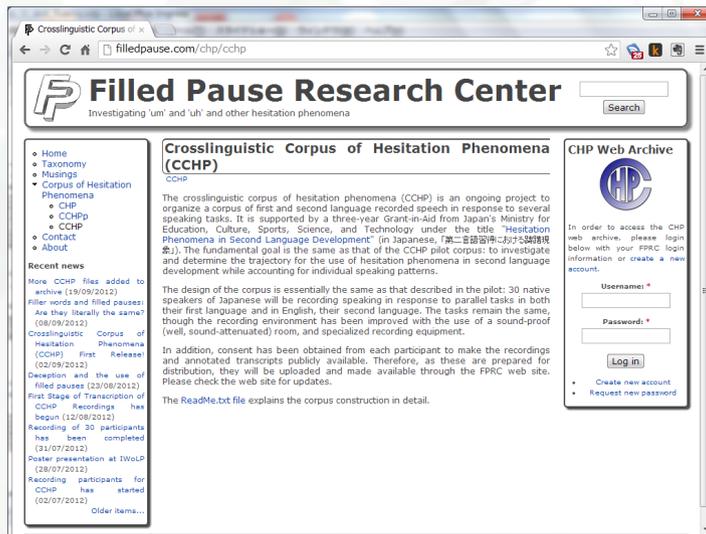
References

Amazon Mechanical Turk web site. www.mturk.com

Bent, T. and Bradlow, AR. (2003) The interlanguage speech intelligibility benefit. In Journal of the Acoustical Society of America, 114(3): 1600-1610.

Brennan, S., & Williams M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. Journal of Memory and Language. 34(3), p. 383-398.

Foote, J. A., & Trofimovich P. (2016). A Multidimensional Scaling Study of Native and Non-Native Listeners' Perception of Second Language Speech. Perceptual and Motor Skills. 122(2), p. 470-489.

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.

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.

Rose, R. (2015). Temporal Variables in First and Second Language Speech and Perception of Fluency. Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015), p. 0405.1-5. the University of Glasgow: Glasgow, UK.

Rossiter, MJ. (2009) Perceptions of L2 Fluency by Native and Non-native Speakers of English. In Canadian Modern Language Review, 65(3):395.

Watanabe, M., Hirose K., Den Y., & Minematsu N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. Speech Communication. 50(2), p. 81-94.

Watanabe, M. & Toyama, S. (2016) Comparison of filled pause durations and frequencies in parallel corpus of English and Japanese. 16th Australasian International Conference on Speech Science and Technology (SST2016), p.42.

Wester, M. and Mayo, C. (2014) Accent rating by native and non-native listeners. In Proceedings of International Conference on Acoustics, Speech, and Signal Processing, 7749-7753.