Interpreting the temporal characteristics of second language speech

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, & Tatsuhiro Nomaguchi

This research is partially supported 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)
Well, in my own life I'd break it up in stages, when I had a difficult youth. My father wasn't in the house. I've written about this. You know there were times where I've experimented with drugs, and I drank in my teenage years. And what I trace this to is a certain selfishness on my part. I was so obsessed with me, and the reasons that I might be dissatisfied, that I couldn't focus on other people. And I think the process for me of growing up was to recognize that it's not about me, it's about ...

Absolutely. But look, you know, when I find myself 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 so that I think is my own failure
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, and uh what I trace this to is uh a certain selfishness on my part, I- 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 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

Barack Obama (August, 2008); Saddleback Presidential Forum
Overview

- Temporal variables of speech
  - Overview
  - In L2 speech
- Crosslinguistic Corpus of Hesitation Phenomena
  - Description
  - Results
- Developmental Trajectory
- Accessing the CCHP
Temporal Variables

Silent pauses
- longer than 0.3-1.0 sec

Self-corrections (repairs)
- Sequence that repairs a preceding sequence
- *Look at the blue the red one over there.*

Filled pauses
- *uh/um* (English)
- *e-to/ano-* (Japanese)

False starts
- Beginning of utterance that is abandoned
  - *Do you I disagree with that.*

Lengthenings
- Prolongation of one or more syllables
  - *I'll take the blue a-nd the- red ones.*

Repeats/Restarts
- Repetition of a sequence of words
  - *I I I think that's a good idea.*

False starts
- Beginning of utterance that is abandoned
  - *Do you I disagree with that.*

Speech rate
- by word, by syllable, with/without pauses
  - (Cucchiarini et al 2010)

Temporal Variables in L2 Production

Temporal Variables 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 relationship between L1 temporal patterns and L2 temporal patterns.
Research Questions

- What is the relationship between temporal patterns in L1 and L2 speech?
- What is the developmental trajectory of temporal variables 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
Crosslinguistic Corpus of Hesitation Phenomena (CCHP)

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

<UTTERANCE>
<T>in</T>
<T>in</T>
<T>America</T>
<T>there's</T>
<T>a</T>
<T>very</T>
<T>famous</T>
<T>and</T>
<T>loved</T>
<T>basketball</T>
<T>cl#</T>
<T>association</T>
<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 Results: Basic Statistics

- **Participants:** 15 Japanese L1, English L2 speakers
- **Full corpus**
  - 24,201 words
  - 5 hrs, 4 min
- **Spontaneous speech**
  - 15,058 words
  - 3 hrs, 56 min
- **Read speech**
  - 9,143 words
  - 1 hr, 10 min

- **Sample recordings**
  - **Speaker 1** (low L2 proficiency)
    - Topic Narrative Japanese “Explain TV”
  - **Speaker 2** (high L2 proficiency)
    - Topic Narrative Japanese “Explain TV”
    - Topic Narrative English “Explain basketball”

- **6,947 silent pauses**
- **1,277 filled pauses**
- **579 self-corrections**
- **177 repeats**
CCHP Results: Speech Rate

Consistent with Derwing et al (2009) and Cox and Baker-Smemoe (2012)
CCHP Results: Silent Pauses

Silent Pause Rate

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

Silent Pause Duration

CCHP Results: Filled Pauses

Filled Pause Rate

Interaction

FP Rate (per 100 words)

L2 Proficiency

English
Japanese
Developmental Trajectory

To-do:
- Filled pause duration
- Lengthenings
- Repair type distribution
- Structural distribution
- Syllable counts
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 temporal patterns.
- Results show that silent and filled pause features indicate learners' L2 proficiency.
- Results suggest that other L2 temporal variables correlate with those of L1.
Assembling a larger (n=35), 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


