Spoken Word Recognition in First and Second Languages: The case of Japanese Listeners

Kiyoko Yoneyama
Daito Bunka University, Tokyo, Japan

Benjamin Munson
University of Minnesota, Minneapolis, US

Organization of the Talk

1. L2 perception research
2. L2 word recognition research
3. Research Questions
4. Experiment
5. Results
6. General Discussion

Levels of Representation, Levels of Processing

- Contemporary models of the phonological lexicon posit at least two distinct levels of representation, a lexical level and a sublexical level.
- The structure of the lexical level of representation is often examined by looking at how lexical similarity—often operationally defined in terms of phonological neighborhood density—affects lexical processing (e.g., Luce and Pisoni, 1998; Vitevitch and Luce, 1999).
- An effect of phonological neighborhood density was observed with Japanese native listeners (Yoneyama, 2002).
- This talk examines this topic in Japanese speakers learning English as a second language.

L2 perception research

- L2 speech perception research has focused on how L1 and L2 listeners differ in their processing of the sound structure of a language
- L2 phonemes that can be assimilated to a single phonemic category in the L2 listener's L1 present the greatest discrimination difficulty.
  - /i/ and /ɪ/ as instances of Spanish /i/ (Best, 1995)

L2 perception research

- Consequently, adult L2 learners are often not sensitive to all of the relevant contrasts in the L2
  - Word initial /p/, /t/, and /k/ are produced with more aspiration in English than Italian, and word-final stops are less likely to be released in English.
  - The difference between word-initial and word-final /p/, /t/, /k/ is difficult for Italian listeners (MacKay, et al., 2001).

L2 perception research

- Spanish-dominant bilinguals require more phonetic information to identify Catalan-specific phonemes than Catalan-dominant bilinguals did in a gating task (Sebastian-Galles and Soto-Faraco, 1999).

L2 Word Recognition Research

- Other research has examined the extent to which phoneme-level perception affects word-recognition accuracy.

Meador, Flege, and MacKay (2000)

- Early and late Italian-English bilinguals
  - Measured repetition accuracy for low-predictability English sentences presented in the noise
  - Found that perception of English vowels consonants predicted a significant proportion of variance in word recognition, beyond what was accounted for by demographic variables
- A close relationship between L2 learners' segmental perception and word recognition

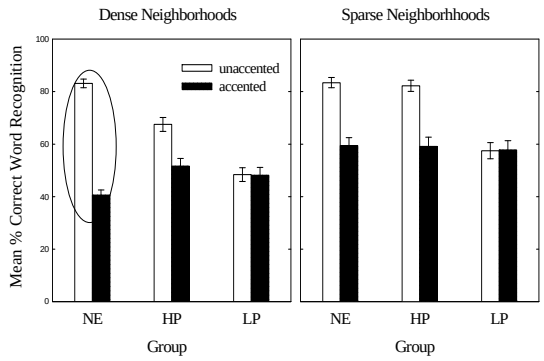
Bradlow and Pisoni (1999)

- There was a bigger discrepancy in recognition accuracy between easy English words (i.e., those with high frequency of use and low neighborhood density) and difficult English words (those with low frequency and high neighborhood density) by non-native listeners when compared with native listeners.
- Lexical factors influence L2 word recognition disproportionately.

Imai, Walley and Flege (2005)

- Extension of Bradlow and Pisoni (1999). Examined high- and low-proficiency L1 Spanish/L2 English listeners' perception of easy and hard words in noise. Words were produced either with Spanish-accented English or with unaccented English.
- Low-proficiency listeners were equally (in)accurate for accented and unaccented words. High-proficiency L2 listeners and native listeners were less accurate for the accented words; however, the L2 listeners showed a smaller discrepancy than the native listeners.
- Everyone was less-accurate for hard words than for easy ones.

Imai, Walley and Flege (2005)



Imai, Walley and Flege (2005)

- English language proficiency of Spanish listeners is a key to determine whether a Spanish-accented English representation or a more native-like English representation is used by non-native speakers.
  - Highly proficient L2 listeners' representations were flexible enough to accommodate accented and unaccented pronunciations

Imai, Walley and Flege (2005)

- In L2, over learning, holistic representations change to → more detailed segmentally better-specified representations (Charles-Luce and Luce, 1990).
- With greater L2 exposure and additional word learning, L2 learners' lexical representations may also become more fine-grained or fully specified vis-a-vis their new language.
- Good L2 pronunciation corresponds to more native-like L2 lexical representation (Flege et al., 1995)

Current Study: Research questions

- **Question 1**: Can Imai et al. (2005)'s findings be extended to a second case of L2 acquisition, that of Japanese L1 speakers' acquisition of English as an L2?
  - Do Neighborhood density, word frequency, and/or word familiarity affect recognition?

Current Study: Research questions

- **Question 2:** Can apparent effects of phonological neighborhood density be attributed to the phonetic content of the stimuli rather than to lexical characteristics?
- The lists of words in Imai et al. were not balanced phonetically. Can apparent effects of lexical difficulty be attributed to difficulties perceiving specific phonemes?

Why a Replication?

- Though we will argue that Imai et al.'s methodology is not perfectly suited to the specific questions that we have about Japanese L2 learners of English, we feel that it's important to investigate whether we can replicate their findings using a typologically different L1.

Methods: Stimuli

- 80 test words used in Imai et al. (2005)
- Highly familiar 3- or 4-phoneme monosyllabic words
- Four lists of 20 words that varied orthogonally in word frequency (WF) and neighborhood density (ND)

	High WF		Low WF	
ND	Dense	Sparse	Dense	Sparse
Example	"bed"	"bring"	"bell"	"boss"
WF	169.5	177.8	18	22.4
ND	23.5	10	23.8	10.4

Methods: Stimuli

- A series of χ^2 tests showed that the distribution of vowels, onsets, and rimes differed significantly as a function of density
  - More /ʃ/, singleton /s/, and bilabial stops in high-density words; more /m/, /f/, and /l/ in low-density words
  - More /e/, /ɛ/, /ɝ/ in high-density words; more /ɔɪ/ and /aʊ/ in low-density words
  - More /n/, /k/, and /d/ in high-frequency words; more /ʃ/, /ŋ/, and /s/ in low-frequency words
- Rimes also differed as a function of frequency

Methods: Recordings

- Two speakers' recordings
  - Native stimuli recorded by a female speaker of North-Central English.
    - Is a doctoral student and a certified speech-language pathologist
    - Was born in Detroit, Michigan, but has lived in Minnesota for 22 years
  - Japanese-accented stimuli recorded by a female speaker of Fukushima Japanese.
    - Earned a M.A. in English language
    - Is currently working as a junior high school English teacher

Methods: Recordings

- The duration of the stimuli were equated (to facilitate measuring response times, not reported here) (Boersma and Weenink, 2008).
- There were numerous phonetic differences between the talkers, the most (perceptually) salient of which were…
  - The L2 talker had a non-native /r/ and /l/, and strongly derhoticized vowels.
  - The L1 talker had strongly back-round /u/ and /o/ and monophthongal /o/ and /e/ (characteristically Minnesotan, but different from most other US dialects)

Methods: Procedures

- Two blocks: A and B (40 test words each, 10 randomly selected from each of the four lexical sets)
- Two stimulus types: native or Japanese-accented
- 4 presentation conditions: half Japanese-accented and half native stimuli
  - A-native/B-Japanese-accented
  - A-Japanese-accented/B-native
  - B-native/A-Japanese-accented
  - B-Japanese-accented/B-native

Methods: Procedures

- In each trial, the listeners heard a word which they repeated, then typed the response they repeated
  - This differed from Imai et al., who only used written responses.
  - The spoken response was used to measure vocal response time, not analyzed here.
  - A subset of the responses were checked to ensure that typed responses were the same as spoken ones. There was very high agreement. Hence, spoken responses are analyzed here.

Methods: Participants

- All participants:
  - Aged between 19 and 50
  - No hearing difficulties
- Three groups of participants:
  - 22 speakers of English living in twin cities, MN (E-US)
    - Undergraduate students, graduate students, and staff at University of Minnesota
  - 24 Native speakers of Japanese with high proficiency in English living in twin cities, MN (E-JP)
    - Mainly graduate students, postdocs, staff and faculty members at University of Minnesota
  - 23 native speakers of Japanese living in Tokyo or Saitama (T-JP)
    - Undergraduates at Daito Bunka University
    - Intermediate English proficiency (all but one students, TOEFL < 475)
    - No experience living in the English speaking countries more than three months

Methods: Analysis

- Spoken responses were corrected for misspellings or alternative spellings (*cheeze* for cheese, *kart* for cart) then were converted to phoneme representations.
  - Postvocalic /r/ were treated as a single phoneme (i.e., the vowel in *corn* was treated as /or/).
- A word was counted as correct only if the response was identical to the target (or was merely misspelled)
- Phonemes were counted as correct if they matched those in the target in the correct word position.
  - "Cart" for "car" would get a 0% for word accuracy but 100% for phoneme accurate
  - "Cone" for "corn" would get 66% phoneme accuracy and 0% word accuracy

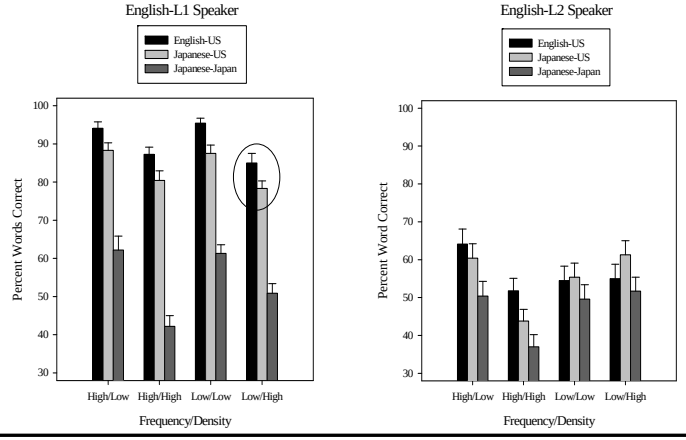
Methods: Analysis

- Percent words and percent phonemes correct were submitted to separate four-factor ANOVAs, with talker, frequency, and density as within-subjects variable, and group as a better-subject variable.

Percent Words Correct: Results

- Significant main effects of
  - Density (LD > HD)
  - Talker (English > Japanese)
  - Listener Group ([E-US = E-JP] > T-JP)
- Significant two-way interactions between
  - Talker language and Listener language
  - Frequency and Listener Language
  - Density and Frequency
  - Density and Talker Language
- Significant three-way interactions among
  - Frequency, density, and talker language

Percent Words Correct: Results



Percent Words Correct: Results

- Significant main effects of
  - Density (LD > HD)
  - Talker (English > Japanese)
  - Listener Group ([E-US = E-JP] > T-JP)
- Significant two-way interactions between
  - Talker language and Listener language
  - Frequency and Listener Language
  - Density and Frequency
  - Density and Talker Language
- Significant three-way interactions among
  - Frequency, density, and talker language

A larger effect of frequency and a smaller effect of talker language for Tokyo listeners

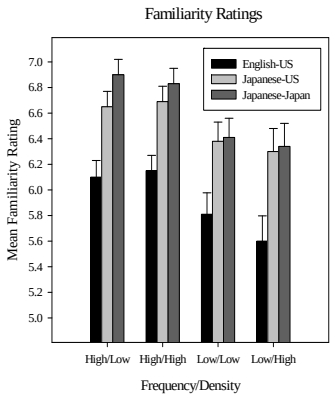
The LF/LD words were recognized less accurately when spoken by the native Japanese speaker than we would predict given the accuracy of LF/HD words

Word by word

- Words which the L1 speaker was more than 10% less intelligible (averaged over all three groups of listeners) than the L2 speaker: ***soup, noon, date, face, lake, save, job, list, duck***
- Words for which the L1 speaker was more than 50% more intelligible (average over all three groups of listeners) than the L2 speaker: ***smell, teach, cute, match, mouth, note, corn, cart, part, cheese, boss, fall, fast, five, lost, park***

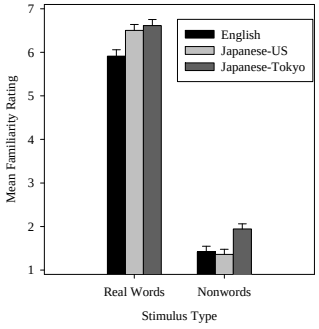
Familiarity Ratings

- Imai et al. did a separate analysis based on participants' familiarity ratings.
- High-frequency words were rated as more familiar than low-frequency words.
- The less-proficient speakers rated words as more familiar than the native speakers.



Familiarity Ratings

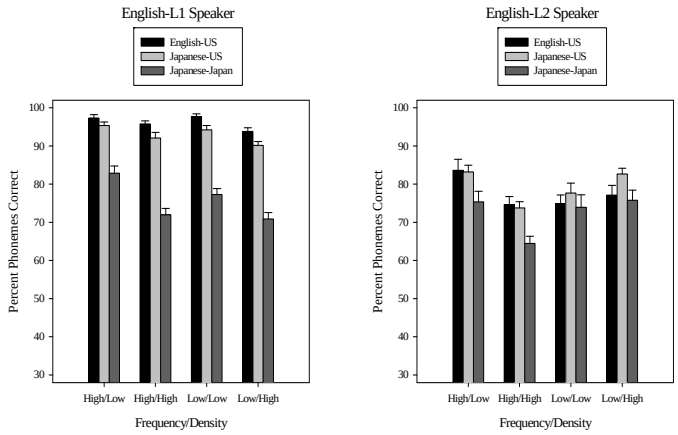
- The less-proficient speakers also rated nonword fillers as more familiar than the native speakers.
- Familiarity ratings were too compressed to allow us to examine the influence of familiarity on word recognition accuracy.



Percent Phones Correct: Results

- Significant main effects of
  - Density (LD > HD)
  - Talker (English > Japanese)
  - Listener Group ([E-US = E-JP] > T-JP)
- Significant two-way interactions between
  - Talker language and Listener language
  - Frequency and Talker Language
  - Density and Frequency
- Significant three-way interactions among
  - Frequency, density, and talker language

Percent Phones Correct: Results



Percent Phones Correct: Results

- Significant main effects of
  - Density (LD > HD)
  - Talker (English > Japanese)
  - Listener Group ([E-US = E-JP] > T-JP)
- Significant two-way interactions between
  - Talker language and Listener language
  - Frequency and Talker Language
  - Density and Frequency
- Significant three-way interactions among
  - Frequency, density, and talker language

A smaller effect of talker language for Tokyo listeners

Phones in the LF/LD words were recognized less accurately when spoken by the native Japanese speaker than we would predict given the accuracy of LF/HD words

Percent Phones Correct: Some Surprises

- The L2 English speaker's production of low-frequency, high-density words are perceived more accurate than we would predict given the overall poor intelligibility of this speaker.
- Why?
- We examined confusion matrixes for vowels and onset consonants to examine this.

Vowel Confusion Matrix: English L1 hearing English L1

	ɑ	æ	ʌ	aɪ	ɑr	aʊ	eɪ	ɛ	i	ɪ	o	ɔ	ɔɪ	or	u	ʊ	ɝ	other
ɑ	86	9	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
æ	0	99	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
ʌ	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
aɪ	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ar	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
aʊ	0	0	0	0	0	98	0	0	0	0	0	3	0	0	0	0	0	0
eɪ	0	0	0	0	0	0	98	0	0	0	0	0	0	0	0	0	0	2
ɛ	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
i	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
ɪ	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
o	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
ɔ	1	0	0	0	0	0	0	0	0	0	0	99	0	0	0	0	0	0
ɔɪ	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
or	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
u	0	0	0	0	0	0	0	0	5	0	0	0	0	0	95	0	0	0
ʊ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	91	0	9
ɝ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	2

Vowel Confusion Matrix: English L1 hearing English L2

	ɑ	æ	ʌ	aɪ	ɑr	aʊ	eɪ	ɛ	i	ɪ	o	ɔ	ɔɪ	or	u	ʊ	ɝ	other
ɑ	59	0	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
æ	1	82	12	1	3	1	0	0	0	0	0	0	0	0	0	0	0	0
ʌ	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
aɪ	0	0	0	98	0	0	0	0	0	0	0	0	0	0	0	0	0	2
ar	31	0	3	0	19	0	0	0	0	0	3	44	0	0	0	0	0	0
aʊ	0	0	2	2	2	88	0	2	0	0	0	4	0	0	0	0	0	0
eɪ	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
ɛ	0	12	0	0	0	0	3	62	0	12	0	0	0	0	1	0	0	10
i	0	0	0	0	0	0	0	0	84	16	0	0	0	0	0	0	0	0
ɪ	0	0	0	0	0	0	2	0	2	95	0	0	0	0	0	0	0	0
o	2	0	6	0	0	0	0	0	0	0	71	6	6	4	0	6	0	0
ɔ	6	3	19	0	0	0	0	0	0	0	15	47	1	4	0	3	0	1
ɔɪ	0	0	0	0	0	0	0	0	0	0	0	0	98	0	2	0	0	0
or	13	0	0	0	0	0	0	0	0	0	46	17	4	8	0	13	0	0
u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	0	2	2
ʊ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
ɝ	5	0	0	0	26	0	0	0	0	0	0	2	0	0	0	2	62	2

English listeners' confusions

- They cannot recover a postvocalic /r/ from the Japanese talker's productions of derhoticized /or/ and /ar/.
  - Not surprising, as the listeners were from a rhotic dialect.
- They had confusions among /ɔ/ and /ʌ/

Vowel Confusion Matrix: Tokyo Japanese hearing English L1

	ɑ	æ	ʌ	aɪ	ɑr	aʊ	eɪ	ɛ	i	ɪ	o	ɔ	ɔɪ	or	u	ʊ	ɝ	other
ɑ	48	39	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0
æ	0	76	6	0	0	1	1	6	1	0	0	0	0	0	0	0	6	1
ʌ	17	9	69	0	0	3	0	0	0	0	0	3	0	0	0	0	0	0
aɪ	0	0	0	87	0	0	10	0	0	3	0	0	0	0	0	0	0	0
ar	3	6	18	0	55	0	0	0	0	0	0	12	0	6	0	0	0	0
aʊ	0	0	0	0	0	89	0	0	0	0	5	2	0	5	0	0	0	0
eɪ	0	3	0	0	0	0	87	0	7	3	0	0	0	0	0	0	0	0
ɛ	0	11	0	0	0	1	1	76	0	0	10	0	0	0	0	0	1	0
i	0	0	0	0	0	0	2	0	84	11	0	0	0	0	0	0	0	0
ɪ	0	0	0	0	0	0	0	26	0	70	2	0	0	0	0	2	0	0
o	2	0	2	0	0	10	0	0	0	0	78	5	0	2	2	0	0	0
ɔ	9	4	9	0	5	0	0	0	0	0	14	51	0	3	3	3	1	0
ɔɪ	0	0	0	0	0	0	2	0	0	0	0	0	98	0	0	0	0	0
or	18	0	0	0	0	0	0	0	0	0	5	9	0	45	0	23	0	0
u	0	0	4	0	0	0	0	0	0	11	0	0	0	0	85	0	0	0
ʊ	13	0	0	0	0	0	0	0	0	0	0	4	0	0	4	78	0	0
ɝ	2	9	0	0	19	0	0	0	0	0	0	0	0	2	2	9	57	0

Vowel Confusion Matrix: Tokyo Japanese hearing English L2

	ɑ	æ	ʌ	aɪ	ɑr	aʊ	eɪ	ɛ	i	ɪ	o	ɔ	ɔɪ	or	u	ʊ	ɝ	other
ɑ	57	9	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
æ	4	72	12	0	0	1	0	0	0	0	2	1	0	0	0	0	7	0
ʌ	0	6	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
aɪ	0	0	0	88	0	0	4	1	0	4	0	0	1	0	0	0	0	0
ar	22	22	17	0	6	3	3	0	0	0	8	14	0	3	0	0	3	0
aʊ	0	0	2	0	2	90	0	0	0	0	0	4	0	2	0	0	0	0
eɪ	0	0	0	0	0	0	95	0	4	0	1	0	0	0	0	0	0	0
ɛ	0	7	0	0	0	0	2	70	1	5	2	0	0	0	5	1	1	4
i	0	0	0	0	0	0	0	0	87	13	0	0	0	0	0	0	0	0
ɪ	0	0	0	0	0	0	0	2	4	93	0	0	0	0	0	0	0	0
o	5	0	0	0	0	7	0	0	0	0	63	14	2	7	2	0	0	0
ɔ	9	17	7	0	0	1	0	0	0	0	19	43	0	1	1	1	0	0
ɔɪ	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
or	4	0	4	0	0	0	0	0	0	0	25	21	4	42	4	0	0	0
u	0	0	4	0	0	2	0	0	0	0	4	0	0	0	83	4	2	0
ʊ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
ɝ	7	4	4	0	20	0	0	0	0	0	9	9	0	0	2	9	43	2

Which vowels are poorly recognized?

- The vowels /ɑ/, /ɑr/, /ɔ/, /or/, and /ɝ/ are hardest to perceive.
  - /ɑ/: ***job*** (HFHD), ***sock*** (LFHD),
  - /ɑr/: ***park, part*** (HFHD), ***cart*** (LFLD)
  - /ɔ/: ***call, fall,*** (HFHD) ***lost, wrong*** (HFLD), ***boss, frog, wash*** (LFLD)
  - /or/: ***corn*** (LFHD), ***fork*** (LFLD)
  - /ɝ/: ***heard*** (HFHD), ***bird, burn, hurt*** (LFHD)
- 72% of the words with poorly identified vowels are from dense neighborhoods

Which vowels are accurately recognized?

- /aɪ/, /aʊ/, /i/, /eɪ/, /ɔɪ/, /u/ are the most accurately recognized vowels.
  - /aɪ/: ***five, kind, smile, white*** (HFLD), ***hide, shine*** (LFHD)
  - /aʊ/: ***house, mouth*** (HFLD), ***loud, mouse*** (LFLD)
  - /i/: ***peace, sheep*** (LFHD), ***cheese, teach*** (LFLD)
  - /eɪ/: ***date, face, lake, rate, save*** (HFHD), ***faith, safe*** (HFLD), ***cake, nail, shake*** (LFHD)
  - /ɔɪ/: ***choice, join, voice***, (HFLD) ***coin, noise*** (LFLD)
  - /u/: ***foot, move*** (HFLD), ***noon, soup*** (LFHD)
- 58% of the words with well identified vowels are from sparse neighborhoods (74% excluding /eɪ/)

Vowels and Lexicality

- A χ^2 test shows that well-recognized and poorly recognized vowels are not distributed randomly among high- and low-density words
  - $\chi^2_{[df=1]}=5.88$, p = 0.015
- It's difficult to tease the density effects apart from the vowel effects

Onset confusion matrix: English L1 hearing English L1

	J	b	br	ch	d	f	fr	h	k	kj	l	m	n	p	r	s	sh	sm	t	v	w	other
J	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	88	0	0	0	8	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
br	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	0	0	0	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33
f	0	0	0	0	0	96	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0	0
fr	0	0	10	0	0	0	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
h	0	4	0	0	0	2	0	91	0	0	0	0	0	4	0	0	0	0	0	0	0	0
k	0	0	0	0	0	0	0	0	98	0	0	0	0	0	0	0	0	0	2	0	0	0
kj	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	0	0	3	98	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
p	0	0	0	0	0	0	0	2	0	0	0	0	0	93	0	0	0	0	2	0	0	2
r	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88	0	0	0	0	0	0	12
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	2
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
sm	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
t	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	96	0	0	0
v	0	0	0	0	0	4	0	0	0	0	4	0	0	0	0	0	0	0	0	92	0	0
w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0

Onset confusion matrix: English L1 hearing English L2

	J	b	br	ch	d	f	fr	h	k	kj	l	m	n	p	r	s	sh	sm	t	v	w	other
J	97	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	86	0	0	2	5	0	2	0	0	0	2	0	0	0	0	0	0	0	0	0	3
br	0	0	67	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	8
ch	0	0	0	82	0	0	0	0	0	0	0	0	0	0	0	5	0	0	14	0	0	0
d	0	0	0	0	85	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10
f	0	1	0	0	0	52	0	7	6	0	0	0	0	13	1	0	0	0	7	1	7	4
fr	8	0	0	0	0	0	83	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0
h	0	0	0	0	0	0	0	68	18	0	0	0	0	4	0	0	0	0	11	0	0	0
k	0	0	0	0	0	0	0	2	89	0	0	0	0	7	0	0	0	0	0	0	0	2
kj	0	0	0	8	0	0	0	0	0	50	0	0	0	0	0	0	0	0	42	0	0	0
l	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	0	0	0	88	2	0	0	0	0	8	0	0	0	2
n	0	0	0	0	0	0	0	0	0	0	2	4	94	0	0	0	0	0	0	0	0	0
p	0	2	0	0	0	2	0	16	8	0	0	0	0	51	0	0	0	0	20	0	0	1
r	0	0	6	0	0	0	0	0	0	0	0	0	0	0	94	0	0	0	0	0	0	0
s	0	0	0	0	2	8	0	0	0	0	0	0	0	0	0	90	0	0	0	0	0	0
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	97	0	0	0	0	0
sm	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	75	0	0	0	4
t	0	0	0	0	0	10	0	5	20	0	0	0	0	20	0	0	0	0	45	0	0	0
v	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0
w	0	3	0	0	0	0	0	3	0	0	0	0	0	0	23	0	0	0	0	0	70	0

Onset Confusion Matrix: Tokyo Japanese hearing English L1

	J	b	br	ch	d	f	fr	h	k	kj	l	m	n	p	r	s	sh	sm	t	v	w	other
J	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	82	0	0	2	0	0	0	0	0	2	0	0	0	0	4	0	0	0	4	0	7
br	0	9	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
ch	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	4	0	0	25	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	67
f	0	13	0	0	0	44	3	3	1	0	0	0	0	9	1	10	0	1	2	2	0	12
fr	0	0	18	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	73
h	0	2	0	0	0	7	0	49	11	0	0	0	0	23	0	0	0	0	9	0	0	0
k	0	0	0	0	0	1	0	0	92	0	0	0	0	3	0	0	0	0	1	0	0	2
kj	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0	1	0	0	0	0	0	0	0	68	4	0	0	17	0	0	0	0	0	0	9
m	0	0	0	0	0	0	0	0	0	0	0	86	9	0	5	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	2	26	72	0	0	0	0	0	0	0	0	0
p	0	0	0	1	0	0	0	4	11	0	0	0	0	77	1	0	0	0	4	0	0	1
r	0	0	0	0	0	0	0	0	0	0	0	54	0	0	34	0	0	0	0	0	11	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	2
sh	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	11	80	0	0	0	0	0
sm	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86	0	0	0	14
t	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	88	0	0	8
v	0	38	4	0	0	0	0	4	0	0	8	0	0	0	0	0	0	0	0	46	0	0
w	0	0	0	0	0	0	0	0	0	0	17	0	0	0	19	3	0	0	0	0	58	3

Onset Confusion Matrix: Tokyo Japanese hearing English L2

	J	b	br	ch	d	f	fr	h	k	kj	l	m	n	p	r	s	sh	sm	t	v	w	other
J	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	86	2	0	0	0	0	2	0	0	2	0	0	0	0	0	0	0	0	0	3	5
br	0	0	75	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	8
ch	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	32	0	0	45	0	0	0	0	0	14	0	0	0	5	0	0	0	0	0	0	5
f	0	0	0	0	0	46	0	11	9	0	0	0	0	19	0	2	0	0	14	0	0	0
fr	0	0	0	0	25	0	17	0	0	0	17	0	0	0	0	0	0	0	0	0	0	42
h	0	0	0	0	0	2	0	41	17	0	0	0	0	21	0	2	0	0	16	0	0	2
k	0	0	0	0	0	2	0	0	92	0	0	0	0	2	0	0	0	0	2	0	0	1
kj	0	0	0	50	0	0	0	0	8	25	0	0	0	0	0	0	0	0	0	0	0	17
l	0	0	0	0	0	0	0	0	0	0	81	1	0	0	17	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	0	0	6	90	2	0	0	0	0	0	0	0	2	0
n	0	0	0	0	0	0	0	0	0	0	0	4	95	0	2	0	0	0	0	0	0	0
p	0	0	0	0	0	0	0	4	7	0	0	0	0	67	0	1	0	0	19	0	0	1
r	0	0	0	0	0	0	0	0	0	0	32	0	0	0	44	0	0	0	0	0	18	6
s	0	2	0	0	0	5	0	0	0	0	0	0	0	2	0	84	2	0	0	0	0	5
sh	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	9	85	0	0	0	0	0
sm	0	0	0	0	0	0	0	0	0	0	8	8	0	0	13	0	0	63	0	0	0	8
t	0	0	0	0	0	0	0	0	9	0	0	0	0	45	0	0	0	0	41	0	0	5
v	0	55	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	41	0	0
w	0	0	0	0	0	0	0	0	0	0	21	0	0	0	24	0	0	0	0	0	55	0

Consonants and Lexicality

- In contrast to vowels, there was no systematic relationship between consonant errors and the lexical characteristics of the stimuli
  - The consonant errors appeared to follow well-known differences between the languages' phoneme inventories

General Discussion

- There was no clear evidence for an emerging 'neighborhood competition' effect in the Japanese learners of English.
  - I.e., we didn't see a reduced competition effect in the Tokyo Japanese listeners relative to the Minneapolis Japanese listeners and the Minneapolis English listeners.
  - Unlike in Imai et al., who did show a bigger discrepancy between high- and low-density words in more- and less-proficient learners of English whose L1 was Spanish.

General Discussion

- However, there was at least one confound in the stimuli that may have mediated these results
- Moreover, we separated our groups based on location, not on proficiency
  - This was done for expedience—we are currently measuring the proficiency and we re-analyze the data when we have these findings.

General Discussion

- The phoneme-by-phoneme results suggest that we need a new methodology for examining the relative contribution of word- and phoneme-level representations on L2 performance.
- Log-linear modeling of the relationship between phoneme and word recognition might provide a clearer picture of the influence of lexical factors on word recognition when the perceptability of phonemes is controlled.

Acknowledgements

- This study was partly supported by JSPS Grand-in-Aid For Young Scientists (B) 18720110 to Kiyoko Yoneyama, and by a grant from the US National Science Foundation (BCS-0729277) to Benjamin Munson.
- Kari Urberg-Carlson and Celina Marnie, for their testing in Twin Cities
- We thank people who participated in the study in Twin Cities and Tokyo.

References

Best, C. T. (**1995**). "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Baltimore), pp. 171-206.

Bradlow, A. R., and Pisoni, D. B. (**1999**). "Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors," *Journal of the Acoustical Society of America* **112**, 272-284.

Charles-Luce, J., and Luce, P. A. (**1990**). "An examination of similarity neighborhoods in young children's receptive vocabularies," *Journal of Child Language* 22, 727-735.

Flege, J. E., Munro, M. J., and MacKay, I. R. A. (**1995**). "Factors affecting degree of perceived foreign accent in a second language" *Journal of Acoustical Society of America* **97**, 3125-3134.

Imai, S., Walley, A. C., & Flege, J. E. (**2005**). "Lexical frequency and neighborhood density effect on the recognition of native and Spanish-accented words by native English and Spanish listeners," *Journal of the Acoustical Society of America* **117**, 896-907.

Luce, P. A., and Pisoni, D. B. (**1998**). "Recognizing spoken words: The neighborhood activation model," *Ear and Hearing* **19**, 1-36.

Meador, D., Flege, J. E., and MacKay, I. R. A. (**2000**). "Factors affecting the recognition of words in a second language," *Bilingualism: Language and Cognition*, **3**, 55-67.

Sebastian-Galles, N., and Soto-Faraco, S. (**1999**). "Online processing of native and non-native phonemic contrasts in early bilinguals," *Cognition*, **72**, 111-123.

Vitevitch, M., and Luce, P. A. (**1999**). "Probabilistic phonotactics and neighborhood activation in spoken-word recognition," *Journal of Memory and Language* **40**, 374-408.

Yoneyama, K. (**2002**). *Phonological neighborhoods and phonetic similarity in Japanese word recognition*. Doctoral dissertation, Department of Linguistics, Ohio State University.