

**AN ACOUSTIC ANALYSIS OF ENGLISH
LEXICAL STRESS REALIZED BY KOREAN,
JAPANESE, AND CHINESE SPEAKERS
-IN COMPARISON TO NATIVE ENGLISH
SPEAKERS**

YE-JEE JUNG (PURDUE UNIVERSITY)

WHY STRESS?

- Kang (2010): Native U.S. English speakers judging international TAs' speech samples in terms of accentedness and comprehensibility
 - Lexical stress (along with pitch range) was mostly associated with the judgement of foreign accent.
 - Suprasegmental features such as stress, pitch range, or speaking rate do have an effect on making L2 English learners' speech sound foreign-accented.
 - It is necessary for L2 English learners and ESL teachers to put an emphasis on suprasegmental features for decreasing foreign accent.

ENGLISH LEXICAL STRESS

- English has free stress system as contrary to languages which have fixed stress system (e.g. Polish, Czech)
- In free-stressed languages such as English, stress can have phonemic value and learners should learn stress as a property of lexical items
-e.g. *pérvert* (noun) vs. *pervért* (verb) (Flege & Bohn, 1989)
- English lexical stress is realized through three main acoustic correlates: duration, fundamental frequency (pitch), and intensity. Stressed syllables have longer duration, higher pitch and greater intensity than unstressed ones. (Beckman, 1986; Archibald, 1992)

KOREAN PROSODY

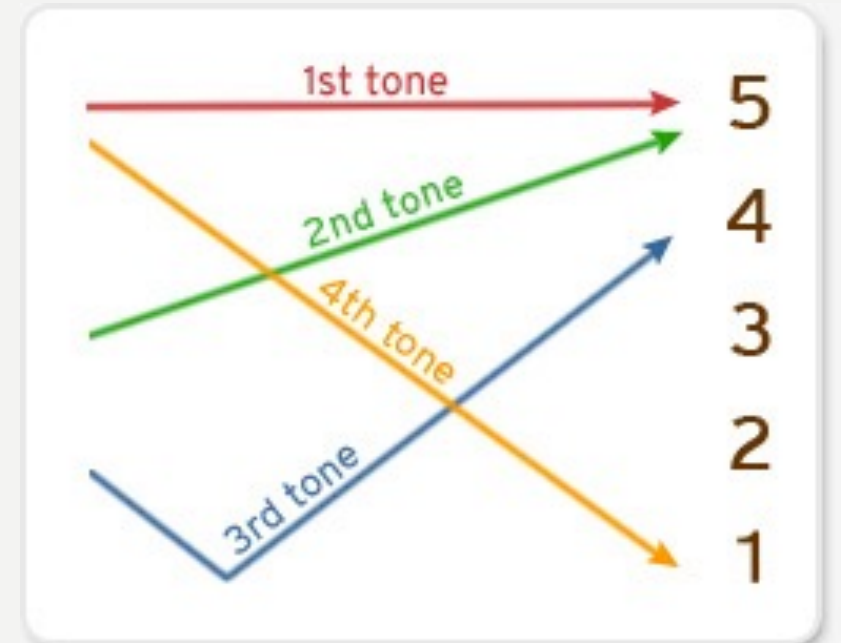
- Hong (2012)
 - Korean is a syllable-timed language in which syllables tend to have identical duration.
 - Kim & Han (1998)
 - Vowel duration used to be phonemic in Korean and some Korean dialects still use the feature (e.g. *nu:n* “snow” vs. *nun* “eye”).
- However, the feature does not exist in modern standard Korean – also known as Seoul dialect – which is the dialect of participants in the current study.

JAPANESE PROSODY

- Lee, Guion, & Harada (2006); Kondo (2009)
 - Japanese is a mora-timed language: short segments consist of one mora and long segments consist of two morae.
 - The duration of a word or phrase is determined by the number of mora it has.
 - Vowel duration has phonemic value in Japanese
 - *su* “vinegar” vs. *suu* “number”
 - *kita* “north” vs. *kiita* “listen(past)”

CHINESE (MANDARIN) PROSODY

- Zhang et al., (2008)
 - Chinese (Mandarin) is a lexical tone language where there are four lexical tones: tone 1 (high-level), tone 2 (high-rising), tone 3 (dipping) and tone 4 (high-falling).
 - Pitch is a primary acoustic cue that manifests Chinese tones.
 - Tone has phonemic value in Chinese.
 - e.g. *mā* “mother” vs. *mǎ* “horse”



PREVIOUS STUDIES

- Lots of previous studies have examined how non-native speakers of English produce English lexical stress: on Arabic learners (Zurairq & Sereno, 2007), on Chinese speakers (Zhang et al., 2008) on Spanish speakers (Flege & Bohn, 1989), on Bengali Speakers (Saha & Mandal, 2018), on Vietnamese speakers (Nguyen & Ingram 2005), on Japanese speakers (Kondo, 2009), on Korean-English bilinguals (Guion, 2005), and on Korean-and Japanese-English bilinguals (Lee et al., 2006).
- Most previous studies used English minimal pairs (e.g. *cónduct* vs. *condúct*) as stimuli and embedded these pairs in a carrier sentence such as *I'll say [target word] this time* or *I said [target word] this time* → providing a constant prosodic environment for target words but too unnatural and arbitrary (Kwon, 2007)
- It has not been much discussed what happens when English minimal pairs are embedded in context sentences: what happens in a more natural setting?

THE CURRENT STUDY

- Research Questions

- (1) How do non-native English speakers produce English lexical stress in carrier sentences?
- (2) How do non-native English speakers produce English lexical stress in context sentences?

METHODOLOGY

- 6 speakers for 4 speaker groups: English, Korean, Japanese and Chinese
 - All speakers were undergraduate, graduate or exchange students in Yonsei University, Seoul, South Korea
- Materials
 - 8 English minimal pairs (adopted from Nguyen & Ingram, 2005; Zhang et al, 2008) - *conduct, conflict, contract, insult, object, permit, record, and suspect*
 - The pairs were embedded in two types of a sentence
 - a carrier sentence “I said [target word] this time” → Experiment 1
 - context sentences “He arrested the [suspect]” vs. “I do not [suspect] him” → Experiment 2

METHODOLOGY

- Procedure
 - All recordings were conducted in a linguistic laboratory at Yonsei University, using a TASCAM HD-P2 recorder at a sampling frequency of 44.1K Hz
- Measurement & Acoustic Analysis
 - Duration, mean pitch, and mean intensity of vowels of the minimal pairs were analyzed by using Praat software (version 6.0.14; Boersma & Weenink, 2016)
 - However, as for a dependent variable, ratios of each acoustic cue were used instead of absolute values, following the methodology of previous studies (Nguyen & Ingram, 2005; Zuraig & Sereno, 2007; Kim, 2011; Na, 2013; Yang, 2016)
 - e.g. an unstressed vowel = 0.05 (s), a stressed vowel = 0.10 (s) → ratio: 2.00
 - Why is it better to use ratio as a dependent variable?
 - Ratios can show to what extent native and non-native speakers vary acoustic correlates to manifest stress
 - Using ratio is effective in controlling variation in speaker and speaking rate (Zuraig & Sereno, 2007)

RESULT – EXPERIMENT 1 (CARRIER)

- Ratios of each Acoustic Cue of Nouns in Carrier Sentences (penultimate stress)

	Duration (s)			Pitch (Hz)			Intensity (dB)		
	SV	USV	Ratio	SV	USV	Ratio	SV	USV	Ratio
Native	0.121	0.063	1.91	217.79	147.41	1.43	60.35	56.51	1.07
Korean	0.079	0.065	1.21	187.18	153.49	1.22	60.84	58.09	1.05
Japanese	0.079	0.068	1.16	233.67	174.38	1.34	58.12	53.81	1.08
Chinese	0.092	0.074	1.25	244.72	199.12	1.23	54.50	50.45	1.08

Note. SV = Stressed Vowel, USV = Unstressed Vowel

RESULT – EXPERIMENT 1 (CARRIER)

- Native English speakers showed the biggest difference between stressed and unstressed vowels in duration (D ratio = 1.91, P ratio = 1.43, I ratio = 1.07).
- Non-native English speakers also made the difference in vowel duration, but all of their ratios were significantly lower than that of native speakers ((F (3, 368) = 162.408, $p < 0.05$)).
- Native English speakers showed the stronger preference for using duration as stress cue – their duration ratio was significantly higher than pitch and intensity ratio (F (2, 276) = 117.885, $p < 0.05$). In other words, there was a significant difference between the strongest stress cue and the other ones, which was not the case in non-native speaker groups.

RESULT – EXPERIMENT 1 (CARRIER)

- Ratios of each Acoustic Cue of Verbs in Carrier Sentences (ultimate stress)

	Duration (s)			Pitch (Hz)			Intensity (dB)		
	SV	USV	Ratio	SV	USV	Ratio	SV	USV	Ratio
Native	0.105	0.061	1.73	206.49	182.90	1.13	61.24	54.10	1.13
Korean	0.084	0.062	1.35	177.48	161.34	1.10	60.90	54.42	1.12
Japanese	0.081	0.069	1.18	221.44	188.94	1.17	57.84	52.94	1.09
Chinese	0.089	0.074	1.21	231.17	210.54	1.10	53.93	48.02	1.12

Note. SV = Stressed Vowel, USV = Unstressed Vowel

RESULT – EXPERIMENT 1 (CARRIER)

- For all speaker groups, duration was the strongest stress cue. However, duration ratio of native speakers was significantly higher than that of non-native speakers ((F (3, 374) = 102.658, $p < 0.05$)
- Again, native English speakers showed the stronger preference for using duration to indicate lexical stress – their duration ratio was significantly higher than pitch and intensity ratio ((F (2, 279) = 143.904, $p < 0.05$).
- Duration ratios of Korean and Chinese speakers were also significantly higher than the other ratios ((F (2, 276) = 43.902, $p < 0.05$ and (F (2, 285) = 13.073, $p < 0.05$), respectively). Japanese speakers did not show a significant difference between duration and pitch ratio (D ratio = 1.18, P ratio = 1.17).
- It was still native speakers who showed the biggest ratio difference between the strongest and the second-strongest stress cue. → stronger preference towards duration

RESULT – EXPERIMENT 2 (CONTEXT)

- Ratios of each Acoustic Cue of Nouns in Context Sentences (penultimate stress)

	Duration (s)			Pitch (Hz)			Intensity (dB)		
	SV	USV	Ratio	SV	USV	Ratio	SV	USV	Ratio
Native	0.090	0.048	1.87	183.41	129.07	1.42	57.84	54.62	1.06
Korean	0.068	0.061	1.11	158.63	141.26	1.12	57.71	56.69	1.02
Japanese	0.072	0.067	1.07	198.62	164.01	1.21	53.23	52.65	1.01
Chinese	0.080	0.071	1.12	219.21	172.20	1.27	49.95	48.40	1.03

Note. SV = Stressed Vowel, USV = Unstressed Vowel

RESULT – EXPERIMENT 2 (CONTEXT)

- The strongest stress cue for native speakers was duration while it was pitch for non-native speaker groups.
- Except for Korean speakers (P ratio = 1.12, D ratio = 1.11), all speaker groups showed a significant difference between the strongest cue and the second-strongest or the weakest stress cue.
- However, native speakers demonstrated more clear preference about the strongest cue (duration) – they made a bigger difference between the strongest cue and the other cues than non-native speakers did.

RESULT – EXPERIMENT 2 (CONTEXT)

- Ratios of each Acoustic Cue of Verbs in Context Sentences (ultimate stress)

	Duration (s)			Pitch (Hz)			Intensity (dB)		
	SV	USV	Ratio	SV	USV	Ratio	SV	USV	Ratio
Native	0.085	0.050	1.71	183.56	197.29	0.93	58.68	54.89	1.07
Korean	0.068	0.060	1.16	170.80	160.98	1.06	59.15	56.44	1.05
Japanese	0.075	0.067	1.12	199.18	197.01	1.01	56.11	54.42	1.03
Chinese	0.081	0.072	1.13	218.54	214.46	1.02	52.36	49.02	1.07

Note. SV = Stressed Vowel, USV = Unstressed Vowel

RESULT – EXPERIMENT 2 (CONTEXT)

- As they did in all other cases, native speakers made the biggest difference between stressed and unstressed vowels in duration. In addition, it was unexpected that their pitch ratio was 0.93, indicating that they produced unstressed vowels with higher pitch. Their duration ratio was significantly higher than both pitch and intensity ratio ($F(2, 282) = 354.842, p < 0.05$).
- Duration ratios of non-native speakers were also significantly higher than the other two ratios. Nevertheless, they made a less difference between duration and the other ratios than native speakers did.

CONCLUSION

- Irrespective of sentence types (carrier or context) and word types (noun or verb), native speakers always demonstrated a consistent stress pattern – they made the biggest difference in duration between stressed and unstressed vowels.
- As native speakers did, non-native speakers succeeded in making a difference in duration between stressed and unstressed vowels in both carrier and context sentences. However, their duration ratios were always significantly lower than those of native speakers.

	Duration (s)			Pitch (Hz)			Intensity (dB)		
	SV	USV	Ratio	SV	USV	Ratio	SV	USV	Ratio
Native	0.085	0.050	1.71	183.56	197.29	0.93	58.68	54.89	1.07
Korean	0.068	0.060	1.16	170.80	160.98	1.06	59.15	56.44	1.05
Japanese	0.075	0.067	1.12	199.18	197.01	1.01	56.11	54.42	1.03
Chinese	0.081	0.072	1.13	218.54	214.46	1.02	52.36	49.02	1.07



a significant effect of speaker group on duration ratio ($F(3, 378) = 150.418, p < 0.05$)

CONCLUSION

- Non-native speakers did not show a consistent pattern regarding the strongest stress cue:
 - They relied upon pitch most in case of nouns (penult-stressed) but relied upon duration most in case of verbs (ult-stressed)
 - It is still unclear why they showed this inconsistent tendency, which needs to be clarified in future studies
- Non-native speakers did not prefer one acoustic correlate as strongly as native speakers did:
 - The ratios of the strongest stress correlate of native speakers were always significantly higher than the ratios of the other two correlates.
 - It was not the case among non-native speakers.

LIMITATION

- There were not sufficient speakers for each speaker group – it is necessary to get speech samples from more speakers for making a better generalization.
- The inherent acoustic features of a vowel were not taken into consideration and each vowel has a different phonetic environment, which can have an effect on acoustic correlates. For instance, vowels followed by voiced consonants usually have longer duration than vowels followed by voiceless consonants – this problem can be resolved by using non-real words with CVCV structure so that the effect of phonemic environments can be minimized
 - McClean & Tiffany (1973): *sasa*
 - Pollock et al. (1993): *boda, dabo, bofi, fibo*
- Non-native speakers were not strictly controlled with regard to their English proficiency – more speakers need to be recruited to divide them based on their English proficiency.

IMPLICATION

- Unlike many previous studies that examined how non-native speakers produce English lexical stress in an unnatural setting (reading carrier sentences), the current study tried to investigate how they behave in a more natural setting (reading context sentences).
- It turned out that in both conditions, non-native speakers produced English lexical stress in a different manner compared to native speakers – this difference can contribute to making their speech less native-like and foreign-accented. In particular, they should try to produce stressed English vowels with sufficiently longer duration time compared to unstressed vowels.



THANK YOU!

YEJEEJUNG@PURDUE.EDU