English Lexical Stress in Awareness and Production: Native and Non-native Speakers

Mariko Sugahara[1], Sylvain Coulange[1, 2], Tsuneo Kato[1]
[1]Doshisha University (Japan), [2]University Grenoble Alpes (France)

The Goal

To investigate **awareness** vs. **stress production of English lexical stress** among:

- Native speakers of English (**ENS**)
- Tokyo/Keihan Japanese learners of Eng (**JLE**)
- Seoul Korean learners of English (**KLE**)

and see if/how their L1 lexical prosody systems affect their stress awareness and production.

L1 lexical prosody systems

English
- a stress accent language
- word-initial primary stress is predominant[1], but the wd-final foot usually receives primary stress in multi-feet words.
- secondary stress is present.

Japanese
- a pitch accent language
- accent plays a distinctive role[2]
- not as heavily biased toward word-initial accent as English, e.g., morphologically complex and longer words[3]
- no subsidiary accent

Seoul Korean
- lacking lexically distinctive accent[4]

The current study

Stress awareness *(Paper-based stress assignment task)*	**Stress production** *(Production task)*
Participants assigned an accent mark on the syllable that they considered primarily stressed.	Participants read aloud the words below, embedded in *"Please say __ for me"* *"I said __ this time"*

Words used in the two tasks

3 groups of morphologically related words: **19 triplets (= 57 words)** were adopted.

Group1: ***X***	Group2: ***X-ing***	Group3: ***X-ion***
Prescriptive stress = **σ1**	Prescriptive stress = **σ1**	Prescriptive stress = **σ3**
AC.ti.vate	***AC***.ti.va.ting	ac.ti.***VA***.tion
DO.mi.nate	***DO***.mi.na.ting	do.mi.***NA***.tion
E.du.cate	***E***.du.ca.ting	e.du.***CA***.tion
CON.sti.tute	***CON***.sti.tu.ting	con.sti.***TU***.tion
⋮	⋮	⋮

Participants in the two tasks

- 12 ENS (only English spoken at home)
- 14 JLE (CEFR: 9 B1s, 5 B2s)
- 11 KLE (CEFR: 6 B1s, 3 B2s, 2 C1s)

Acoustic analysis of the production task

① Forced segmental alignment using **MFA**[6]

② Acoustic analysis using **PLSPP**[7]

- For each vowel, (a)-(c) were obtained
  (a) the F0 mean across 10ms frames
  (b) the max intensity value
  (c) the duration

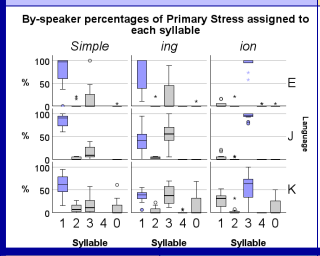
pipeline GitLab

- The **speaker-normalized percentile values** were obtained for (a), (b), (c) separately.

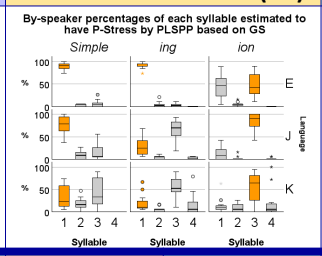
  = **F0 score (FS), intensity score (IS), duration score(DS)**
- **FS+IS+DS averaged = GS (Global Score)**

③ Estimation: The syllable containing the vowel with **the greatest GS was estimated to carry primary stress.**

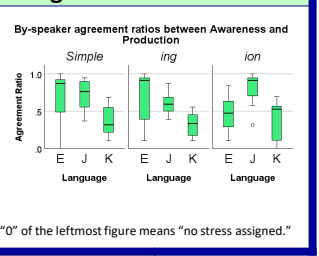
A: Awareness results



B: Production results (GS)

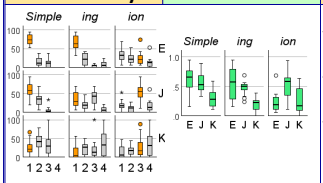


Agreement bet. A & B

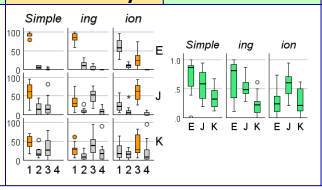


"0" of the leftmost figure means "no stress assigned."

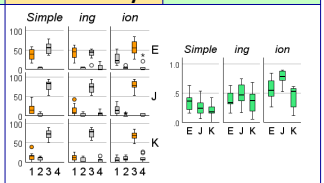
B': Estimation with FS only / Agreement bet. A & B'



B'': Estimation with IS only / Agreement bet. A & B''



B''': Estimation with DS only / Agreement bet. A & B'''



→ PLSPP's estimation based on **GS** gave the best agreement ratios with the awareness results. (None of the PLSPP's analyses with individual acoustic variables, i.e., FS, IS, or DS, by itself gave better results than those with GS.)

ENS

Simple & -ing

Awareness results	Production results
• mostly prescriptive • but, **inter-speaker variability**, especially notable for ***-ing***	• PLSPP's estimation uniformly prescriptive • **no inter-speaker variability**

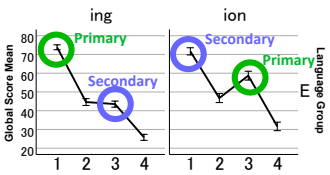
Some ENS:
- They are not always aware of the prescriptive primary stress location (= σ1).
- Nonetheless, they uniformly made σ1 acoustically most prominent.

→**skewed toward σ1 acoustic prominence**

-ion

• uniformly prescriptive • **no inter-speaker variability**	• PLSPP's estimation divided into σ1 & σ3. • **inter-spkr variability**

- ***-ion* and *-ing* were still distinguished by σ3's GS** even when they were estimated to have σ1-primary stress by PLSPP.



<Implications>
- Syntagmatic analyses of "the greatest GS" do not necessarily give a correct picture:
- **Paradigmatic relations** should be also taken into consideration when detecting non-initial primary stress.

Why was the prescriptive **σ3-stress of *-ion*** more strongly entrenched in ENS's awareness than the prescriptive **σ1-stress of *-ing***?

A possible account:
The default stress pattern of multi-feet words = primary stress **on the left σ of the wd-final foot**, e.g. ($mas_{σ1}$ $sa_{σ2}$)($CHU_{σ3}$ $setts_{σ4}$)

→ easier to be entrenched in consciousness

JLE

Simple & -ion

Awareness results	Production results
• mostly prescriptive • smaller inter-speaker variability • high degree of agreement between awareness and production	

- For ***-ion***, unlike ENS, σ3 consistently obtained the greatest GS.
- not surprising given that **only one accented syllable** is acoustically highlighted in their L1

-ing

divided into σ1- & σ3-stress	σ3-stress more dominant than σ1

lower agreement ratio between awareness and production than *Simple* and *-ion*

- The prescriptive σ1-primary stress was not fully acquired because it deviates from not only the default stress pattern of English two-feet words but also from **L1's accent pattern**:
- Their L1 (Japanese) prefers **word-medial accent in morph-complex words.**

KLE

	Awareness results	Production results
Simple	Prescriptive σ1-stress was dominant.	mostly divided into σ1- & σ3-stress
	greater inter-speaker variability and **more diverse patterns**	
-ing	mainly divided into σ1- & σ3-stress	σ3-stress more dominant than σ1
	more diverse patterns than JLE's	
-ion	• Prescriptive σ3-stress was dominant, • but there was **greater inter-speaker variability** and **more diverse stress patterns**.	

Agreement ratios between awareness and production were low across the 3 forms.

A possible account:
- Some of the KLE have set up the **lexical accent parameter**[8] so that accent is not encoded in the phonological representation,
- due to their L1 lacking distinctive accent.

Summary of major findings and comments

- For the two-feet words with ***-ing***
  - ➢ not only the E learners but also some of the ENS had difficulty in entrenching the prescriptive σ1-stress pattern in their awareness, which deviates from the default stress pattern of two-feet words.
  - ➢ JLE was more skewed toward σ3-stress in awareness and production than ENS and KLE, which implies that they are influenced by their L1 accent pattern.
- The greatest acoustic prominence in ENS's speech was skewed toward σ1 regardless of their awareness.
- JLE showed less inter-speaker variability and more prescriptive patterns in both awareness and production, and higher agreement ratios between awareness and production than KLE. This is because JLE have a distinctive lexical accent in their L1, whereas KLE do not.
- More participants are required before making a solid conclusion.

example	d	o	m	i	n	a	t	i	ng
FS		75		91		20		13	
IS		93		64		76		15	
DS		71		19		90		22	
GS		**79.7**		**58.0**		**62.0**		**16.7**	
Estimated primary stress		✓							

References
[1] Cutler & Carter. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, 2: 133-142.
[2] Sibata & Sibata. (1990). To what extent can accents distinguish homophones? *Keiryoo Kokugogaku* 17(7): 311–323.
[3] Kubozono. (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua*, 116, 1140-1170.
[4] Jun. (1996). *The phonetics and phonology of Korean prosody*. Garland Publishing.
[5] Sugahara. (2019). The effect of native languages on the judgement of primary stress locations in English. *Doshisha Studies in English*, 100: 165-221.
[6] McAuliffe, et al. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. The *18th ISCA*.
[7] Coulange. (n.d.). Pauses & lexical stress processing pipeline. https://gricad-gitlab.univ-grenoble-alpes.fr/lidilem/plspp
[8] Peperkamp & Dupoux. (2002). A typological study of stress 'deafness'. *LabPhon 7* (pp.203-240). Mouton de Gruyter.

Acknowledgment: This work was supported by JSPS KAKENHI Grant Number JP22K00627.