말소리와 음성과학

Phonetics and Speech Sciences

한 국 음 성 학 회 지

Check for updates

English vowel production conditioned by probabilistic accessibility of words: A comparison between L1 and L2 speakers*

Jonny Jungyun Kim[1,**] · Mijung Lee[2]

[1]*Department of English Language and Literature, Pusan National University, Busan, Korea*
[2]*Independent Researcher*

Abstract

This study investigated the influences of probabilistic accessibility of the word being produced – as determined by its usage frequency and neighborhood density – on native and high-proficiency L2 speakers' realization of six English monophthong vowels. The native group hyperarticulated the vowels over an expanded acoustic space when the vowel occurred in words with low frequency and high density, supporting the claim that vowel forms are modified in accordance with the probabilistic accessibility of words. However, temporal expansion occurred in words with greater accessibility (i.e., with high frequency and low density) as an effect of low phonotactic probability in low-density words, particularly in attended speech. This suggests that temporal modification in the opposite direction may be part of the phonetic characteristics that are enhanced in communicatively driven focus realization. Conversely, none of these spectral and temporal patterns were found in the L2 group, thereby indicating that even the high-proficiency L2 speakers may not have developed experience-based sensitivity to the modulation of sub-categorical phonetic details indexed with word-level probabilistic information. The results are discussed with respect to how phonological representations are shaped in a word-specific manner for the sake of communicatively driven lexical intelligibility, and what factors may contribute to the lack of native-like sensitivity in L2 speech.

Keywords: vowel hyperarticulation, word frequency, neighborhood density, phonotactic probability, speaker attention, L2

1. Introduction

Probabilistic information about the word being produced or perceived is known to be a critical factor for word retrieval during speech processing. For example, compared to words with low usage frequency, high-frequency words are recognized faster (Dahan et al.,

2001; Luce, 1986; Marslen-Wilson, 1987; Marslen-Wilson, 1990) and produced faster in naming tasks (Forster & Chambers, 1973; Oldfield & Wingfield, 1965). To account for this sort of facilitated lexical access, frequency-related probabilistic information is coded in many influential psycholinguistic models (e.g., Marslen-Wilson 1987, 1989; McClelland & Elman, 1986; Norris, 1994) to adjust the word's resting accessibility.

Another type of probabilistic lexical information to be investigated in this study is concerned with phonological similarity among words stored in the mental lexicon, which leads to parallel activation of phonologically related words during lexical search (Luce, 1986; McClelland & Elman, 1986; Ussishkin & Wedel, 2002). For example, spoken-word recognition is impeded when the input shares acoustic similarity with a greater number of words in the lexicon (i.e., high neighborhood density) or when the neighborhood words are frequently used ones (i.e., high neighborhood frequency), due to an inhibitory process that suppresses the target lexeme from being selected (Goldinger et al., 1989; Luce & Pisoni, 1998).

There is also evidence that these probabilistic factors for lexical accessibility are used to shape phonetic forms of phonological units in a word-specific way. Munson & Solomon (2004) and Wright (2004) demonstrated that vowels in monosyllabic words are affected in such a way that the vowel space is expanded when producing infrequent words with many neighborhood words (hard words, henceforth), compared to frequent words with few neighborhood words (easy words). Similarly, Kilanski (2009) found that durational properties of consonants produced by native English speakers are also strengthened when the word is low in frequency or high in density. These effects are broadly understood in Lindblom's (1990) Hypo- and Hyper-Articulation Theory as the speaker's modification of fine phonetic details for the sake of intelligibility of words with perceptual difficulty.

As an echoing prediction for vowel length, the vowel is expected to be longer in a hard word than in an easy word. However, an opposite effect was reported in Munson & Solomon (2004) in which longer vowels were produced in easy words. As the authors posit, this may be to be due to generally low phonotactic probability for the segmental combinations in words with low density. In fact, speakers are likely to be less adept for articulatory gestures for sparse segmental sequences and thus spend greater effort as shown by Munson (2001), and this process may apply even for high-accessibility words despite their high frequency. In their lexical set, easy words were indeed significantly lower in phonotactic probability calculated using Munson's (2001) method, compared to hard words. However, the authors concluded this effect (i.e., vowel lengthening associated with low-density words) was not generalizable, because no such effect was observed in their follow-up experiment with a factorial design that teased apart the effects of word frequency and neighborhood density. This sort of vowel lengthening needs to be reinvestigated in its own right, since it was borne out in Kilanski's (2009) data, in which a new set of words were used and normalized vowel lengths were analyzed. Yun (2010) also found a similar trend from a small sample (4 native speakers).

Whether it is an artifact of a particular set of words or a generalizable effect of neighborhood density, the current study aims to replicate Munson & Solomon's (2004) findings on vowel expansion and temporal contraction using the identical lexical items produced by English native speakers. Then, the results will be compared with non-native speakers' productions to examine the extent to which

high-proficiency L2 speakers are sensitive to the implicitly statistical word-specific phonetic representations of L2 vowels. Since these changes in vowels' formants and duration are both ultimately related to the communicatively driven contextual factors (Lindblom, 1990), we also investigate whether and how the effects are further modulated by focus realizations by manipulating the speaker's attention to the target word, which has not been attested in the above-mentioned literature.

The results will be discussed in light of probabilistic views on the nature of mental representations for lexical and phonological units. Proponents of usage-based (Bybee, 2001) or exemplar-based (Goldinger, 1998; Hay & Foulkes, 2016; Pierrehumbert, 2001; Pierrehumber, 2002) approaches commonly propose that lifetime exposures to probabilistically conditioned speech forms play a significant role in shaping the phonological units in the long-term word memory. For example, observing that /t, d/ lenition occurs more frequently in high-frequency English words (Bybee, 1985; Bybee, 2000; Gahl, 2008), Bybee (2001) argues that realization of phonetic variant forms is modulated by word-specific contextual factors. In Exemplar Theory, episodic memories of word-specific forms are accumulated every time they are encountered. Over time, they are sorted as exemplar clusters, and the exemplars are activated while producing or perceiving the word.

While this sort of experience-based view is widely adopted in a variety of literature on phonological processing, the extent to which L2 speakers utilize this mechanism remains unanswered in general. Literature on the influence of English words' probabilistic accessibility is also limited but Yun (2010) explored it, deploying a production experiment with Korean L2 speakers at an intermediate or advanced level. The results showed that realizations of the first two formants were slightly strengthened for words with low frequency or words with high neighborhood density, suggesting L2-ers may also have internalized probabilistically conditioned vowel specificity. Regarding vowel lengths, however, mixed patterns across individuals were reported with no overall statistic significance.

Taken together, the experiment reported below was designed to test (1) whether words' low frequency and high density led to expanded vowel space; (2) whether the vowel was longer in the easy words Munson & Solomon (2004) used, which would test the role of phonotactic sensitivity that might override the high accessibility of high-frequency and low-density words; (3) whether the effects above were further modulated by speaker attention; and (4) whether non-native speakers behaved similarly with native speakers.

2. Method

2.1. Participants

Six native speakers of American English (the native group; F=4, M=2) and six Korean learners of English (the L2 group; F=6, M=0) participated in the experiment. We initially intended to control for speaker gender to examine females' phrase-final prosodic markings (e.g., wider pitch range for hard words), as analyzed in Lee (2021). Due to lack of female native speakers, however, two males were included in the native group, whose spectral values were standardized in the analysis (see below).

The L2 group was recruited from high-proficiency English speakers living in or near Busan, South Korea, with TOEIC scores higher than 900. They had lived in English-speaking countries

longer than 2 years, were graduate students majoring in English literature or English linguistics, or were teaching English at a college or a private institution at the time of participation.

2.2. Lexical Stimuli

The target words in the experiment were the 30 monosyllabic words used by Munson & Solomon (2004), who categorized half of them as 'lexically easy' words (i.e., with high frequency and low neighborhood density) and the other half as 'lexically hard' words (i.e., low frequency and high neighborhood density), based on the Hoosier Mental Lexicon (Pisoni et al., 1985). The target words contained six different vowel categories /ɑ, æ, ɪ, i, o, u/ as provided in Table 1. Although the number of words varied across the vowel categories, each vowel category had the same number of words across the difficulty condition.

Table 1. Lexical stimuli

Vowel	Easy words					Hard words				
	O	Word	C	Freq.	Dens.	O	Word	C	Freq.	Dens.
/i/	△	Peace	△	70	2	▲	Bead	▲	1	4
	△	Teeth	△	48	1	●	Weed	▲	11	6
/ɪ/	▲	Give	▲	1168	2	△	Hick	△	2	6
	△	Ship	△	99	3	△	Kin	●	4	7
	△	Thing	●	1089	3	△	Kit	△	17	2
/æ/	▲	Gas	△	68	5	△	Hack	△	8	3
	▲	Jack	△	252	3	△	Hash	△	3	5
	△	Path	△	25	2	△	Pat	△	18	8
/ɑ/	▲	Job	▲	413	4	△	Cod	▲	2	6
	△	Shop	△	54	2	△	Cot	△	2	4
	●	Wash	△	41	5	●	Knob	▲	2	1
	●	Watch	△	330	1	●	Wad	▲	2	4
/o/	▲	Both	△	295	2	▲	Goat	△	10	4
	▲	Vote	△	34	1	●	Moat	△	1	3
/u/	△	Food	▲	154	2	△	Hoop	△	3	7
Mean				276	2.53				5.73	4.67

The vowel's neighboring consonants (O=the onset, C=the coda) are categorized as a combination of voicing and manner of articulation (▲=voiced obstruent, △=voiceless obstruent, ●=sonorants). Also provided are the words' frequency (occurrence per 1 million words) and neighborhood density values (values obtained from Vaden, Halpin & Hickok, 2009).

2.3. Procedure

Recordings were made using a Britz BE-STM 500 microphone (mono, 32-bit, 44,100 Hz) in a quiet room. Participants produced utterances containing the target words while speech materials were orthographically presented on a computer monitor screen.

Each participant was recorded in two different experimental blocks, differing in the levels of speaker attention to the target word. In the 'unattended' block, the target word appeared in the utterance-final position of a carrier sentence, "I _____ say the word, [TARGET]." Participants were instructed to read the sentence naturally, filling in the blank by themselves with the most appropriate adverbs of frequency for their own usage frequency (among *seldom*, *sometimes*, *usually*, or *often*, presented along with the sentence). By doing so, we intended to elicit a narrow focus on the adverb while articulatory attention to the target word is reduced. Following the unattended block, the 'attended' block was conducted where the target appeared in a different carrier sentence, "*This is the word,* [TARGET].", drawing attention to the target word itself. In

case L2 speakers failed to produce the intended vowel category because of unfamiliarity with the word (e.g., vowel sound other than /ɑ/ for *cod*, *cot*, *knob*, *wad*), the experimenter (the second author) corrected it by showing an IPA symbol.

Lexical items were randomly presented for each speaker, with 4 repetitions per block. Thus, a total of 2,880 tokens (30 words×2 attentions×4 repetitions×12 participants) were obtained. Visual presentations were provided using the macro-script of the Visual Basic for Applications in the Microsoft Excel 2010 software, which enabled full randomization of the materials.

2.4. Measurement and Statistical Analysis

Two types of acoustic measures for the vowels (formant values and vowel lengths) were obtained. First, F1 and F2 values were automatically detected by Praat (ver. 6.1.14) scripting at the midpoint of each vowel token (in Hz), and then were z-scored within each speaker to compare the formant values relative to speaker specificity (e.g., difference in vocal tract length). To examine spectral expansion from the easy words to hard words, we obtained the area of the hexagon created by connecting each of the six vowels' mean x- (F2) and y- (F1) values on the coordinate plane.

Second, the vowel length (in ms) was defined as the stable vowel duration. The vowel's onset and offset were identified by the periodicity of complex wave and parallel distribution of F1 and F2 values. Note that since the vowel occurred utterance-finally, the vowel duration partially included the portion lengthened by pre-boundary effects. Then, the raw vowel duration values were divided by word duration values to factor our speech rate differences. The transformed length values (relative duration) were submitted to the statistical analysis.

Following Munson & Solomon (2004), our analysis compared the effects of lexical difficulty as a single binary factor (easy vs. hard), instead of treating frequency and neighborhood density as two separate continuous variables for the following reasons. First, the two measures covaried (i.e., high-frequency words were also lower in density and vice versa for low-frequency words). Second, we assumed our Korean participants were unfamiliar with some hard words (e.g., *wad*, *moat*), which was consistent with our preliminary analysis on their selection of adverbs of frequency, making the easy-hard binary contrast reasonable. However, all target words were assumed to be familiar to native speakers according to Munson & Solomon's (2004) evaluation.

A linear mixed effects regression model was fit to the relative duration values using the lmerTest 3.1.3 package in R 4.2.2. The three binary experimental factors were included as fixed effects in the model (with the underlined one as a reference level): lexical difficulty (easy or hard), attention (unattended or attended), and group (native or L2), all of which were sum-coded into −0.5 or 0.5. As for the effect of neighboring sounds, dummy-coded segment types of the onset and coda (voiced obstruent, voiceless obstruent, or sonorant) were added as fixed effects, as they improved the model's fit. Random effects structure was determined by evaluating the model's fit relative to the maximal structure allowed by the design (Barr et al., 2013) via a series of likelihood ratio tests (Matuschek et al., 2017). The final model reported below included by-participant random intercept and slopes for difficulty and attention, and by-item intercept and slope for group.

3. Results

3.1. Effects on Formant Realization

Figure 1 visualizes each participant group's vowel space as a function of lexical difficulty on reversed x- and y- axes. Each solid (hard words) or dashed (easy words) line segment connects the coordinate dots of vowel categories in proximity, which were defined as the mean z-scored F1 and F2 values in the respective condition. The ovals in six different colors were created using the stat_ellipse function (with level=0.5) in the ggplot2 package in R to delimit the distribution of each vowel category.
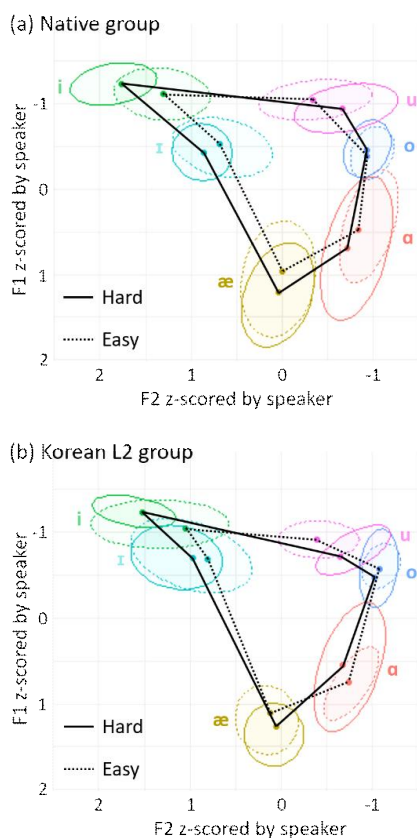
Figure 1. Vowel plots conditioned by probabilistic difficulty.

The L2 group (lower panel) exhibited some noticeable differences from the native group (upper panel). While the native group showed relatively clear separation of spectral qualities between vowel categories, the L2 group showed larger overlaps for the tense-lax contrast between /i/ and /ɪ/, and for back vowels due to their use of a smaller height range in the back cavity, both of which are known as typical patterns of Korean L2-ers' vowel production (Koo, 2000; Yun, 2010).

Importantly, as an effect of lexical difficulty, the native group's vowels appear to have been realized over a larger hexagonal area when the vowel was contained in a hard word (solid line), compared to when it was in an easy word (dashed line). Also in line with the previous studies (Munson & Solomon, 2004; Wright, 2004; Yun, 2010), the native group's articulatory expansion is less clearly observed for back vowels than front vowels, particularly for /o/, probably due to greater restrictions for retracting the tongue root.

On the other hand, the L2 group barely exhibited the expansion while their vowels in general were somewhat fronted in hard words than in easy words. The areas are compared across the groups and the difficulty levels in Table 2.

Table 2. Areas of the vowel space (in z-score²)

	Easy	Hard	Difference by difficulty
Native	2.46	3.12	0.66
Korean L2	2.67	2.71	0.04
Group difference	−0.21	0.41	0.62

Table 2 confirms that the native group produced the vowels using a larger spectral space for hard words than easy words, with the difference of 0.66 z^2 (26.8% increase). On the other hand, the L2 group showed little expansion of articulatory gestures with only 0.04 z^2 increase in the vowel space (1.5% increase). Also notable is that a greater amount of the between-group difference was induced when producing the hard words (0.41 z^2 difference), while both groups produced the vowel in the easy words using a more comparable area (−0.21 z^2 difference). This indicates that the major difference comes from the L2 group's insensitivity to the native propensity to expand the vowel space for hard words, rather than reducing vowels for easy words.

Next, F1 and F2 values are plotted as a function of the attention condition in Figure 2. No clear effects of attention on formant frequencies are visible, suggesting that increased speaker attention did not result in expanded vowel space while vowel expansion was largely led by lexical difficulty, as shown in Figure 1.
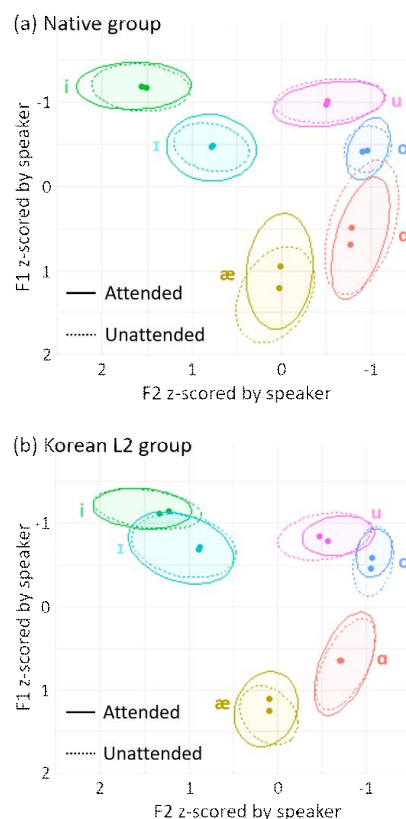
Figure 2. Vowel plots conditioned by speaker attention.

3.2. Effects on Vowel Duration

Figure 3 presents the mean relative vowel duration (in %) predicted by the effects of difficulty, attention, and group. The model output for these effects is presented in Table 3 below.
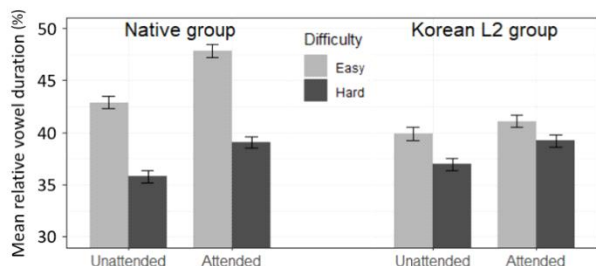
Figure 3. Relative vowel duration conditioned by experimental factors (error bars=standard error of the mean).

Apparent in Figure 3 is the main effect of difficulty: hard words were produced with significantly shorter vowel lengths in the data pooled from both groups (by about 4.4 percent points as shown in Table 3) regardless of attention and group ($p<.05$). The main effect of attention shows that the vowel was longer when it was produced in the attended condition ($p<.001$). Although the overall mean was greater for the native group than the L2 group, the non-significant group effect ($p=.317$) indicates that this group difference is in fact largely explicable by other factors, rather than by differences inherent to being native. For example, main effects of both difficulty and attention appear to be substantially greater in magnitude for the native group compared to the L2 group in Figure 3, though 2-way interactions between difficulty and group ($p=.080$) or between attention and group ($p=.065$) did not reach significance. In addition, the 3-way interaction ($p<.05$) shows that the difficulty:group interactive tendency arises largely from the attended condition, or that the attention:group interaction from the easy words.

Table 3. Summary of the model fit to relative duration of the vowel (bold = significant effects of the experimental factors)

	Estimate	s.e.	*t*-value	*p*-value
(Intercept)	0.562	0.019	30.376	<.001
Difficulty=hard	**−0.044**	**0.019**	**−2.367**	**.025**
Attention=attended	**0.029**	**0.006**	**5.115**	**<.001**
Group=L2	−0.021	0.020	−1.038	.317
Difficulty:Attention	−0.003	0.005	−0.553	.580
Difficulty:Group	0.055	0.030	1.854	.080
Attention:Group	−0.024	0.011	−2.076	.065
Difficulty:Attention:Group	**0.027**	**0.010**	**2.590**	**.010**
Onset=voiceless obstruent	−0.081	0.017	−4.807	<.001
Onset=sonorant	−0.102	0.020	−5.083	<.001
Coda=voiceless obstruent	−0.125	0.016	−7.812	<.001
Coda=sonorant	−0.170	0.030	−5.632	<.001

As for the effects of the control factors (i.e., neighboring consonants), relative vowel duration was shorter ($p<.001$) when the preceding or following consonant was either a voiceless obstruent or a sonorant, compared to the reference level (i.e., a voiced obstruent).

When the effects of difficulty and attention were reexamined by fitting the model to each group separately, the group differences were more explicitly revealed as presented in Table 4.

Table 4. Summary of the models fit to the (a) native and (b) L2 groups

(a) Group=Native	Estimate	s.e.	*t*-value	*p*-value
(Intercept)	0.554	0.023	24.340	<.001
Difficulty=hard	**−0.069**	**0.029**	**−2.418**	**.036**
Attention=attended	**0.041**	**0.004**	**10.265**	**<.001**
Difficulty:Attention	**−0.074**	**0.020**	**−3.700**	**.001**
Onset=voiceless obstruent	−0.107	0.024	−4.517	<.001
Onset=sonorant	−0.105	0.019	−5.540	<.001
Coda=voiceless obstruent	−0.127	0.036	−3.550	.002
Coda=sonorant	−0.016	0.008	−2.010	.045
(b) Group=Korean L2	Estimate	s.e.	*t*-value	*p*-value
(Intercept)	0.566	0.023	24.995	<.001
Difficulty=hard	−0.018	0.018	−1.045	.305
Attention=attended	0.017	0.010	1.817	.129
Difficulty:Attention	0.010	0.006	1.627	.104
Onset=voiceless obstruent	−0.087	0.079	−4.554	<.001
Onset=sonorant	−0.097	0.023	−4.297	<.001
Coda=voiceless obstruent	−0.141	0.018	−7.808	<.001
Coda=sonorant	−0.205	0.034	−6.010	<.001

The native group's model output in Table 4(a) shows significant effects of difficulty ($p<.05$) and attention ($p<.001$), respectively in the predicted direction. Their interaction indicates that the magnitude of the difficulty effect was enhanced when the vowel occurred in attended speech ($p<.001$). The L2 group in (b), however, showed none of these effects though both main effects trended in the same direction.

4. Discussion

As an overall summary of the results, the native group replicated Munson & Solomon's (2004) results, showing expanded vowel space and reduced vowel length when the vowel is contained in words that are probabilistically less accessible (i.e., words with low frequency and high density). However, none of these effects were present in the high-proficiency L2 group's data. We provide some implications on how sub-categorical phonetic details are represented in native speakers' word memory, and then discuss methodological and theoretical issues with respect to how the L2 system may be differentially manifested.

First, our results showing lexically induced vowel hyperarticulation highlight native speakers' implicit knowledge about words' probabilistic accessibility, as combined by frequency and neighborhood density, and its use for communicatively driven phonetic realizations. More specifically in the context of the experience-based account (Bybee, 2001; Goldinger, 1998; Hay & Foulkes, 2016; Pierrehumbert, 2001; Pierrehumber, 2002), the results build on the claim that phonetic form of the vowel in a word that is probabilistically rare and has many phonological competitors is shaped to strengthen word-internal phonemic identity, via native speakers' prior experiences with lexical-level difficulty in recalling the intended lexeme in either production or perception. Although this finding was drawn from a small data set with 6 native speakers, the effect was robust in a conservative linear model with varying intercepts and slopes, and is consistent with the previous findings (Kilanski, 2009; Munson & Solomon, 2004; Wright, 2004).

Second, there was no direct effect of speaker attention on vowel qualities. In the presence of its effect on vowel duration (discussed below), the absence of attention effect on spectral expansion

indicates that attended speech may not be fully hyperarticulated in all phonetic dimensions under the coexistence of the lexical difficulty contrast. As a more compelling possibility, we speculate that the experimental conditioning may not have been adequate enough to induce spectral expansion, because the target was located in utterance-final position where articulatory declension occurs along with lengthening (Lindblom, 1968). That also explains why speaker attention only had a significant effect on the temporal domain.

Third, our results showing temporal contraction of vowels in hard words corroborate Munson & Solomon's (2004) results obtained from the particular set of words that covaried in frequency and density, who interpreted that the covariance resulted in an overriding effect of low phonotactic probability in easy words. Notably, our native group's temporal modulation occurred even utterance-finally, in the presence of pre-boundary lengthening.

It should be reminded here that it is out of the scope of this study as to whether high-density words in general in fact cause vowel shortening. However, it is more important to note the interaction effect between lexical difficulty and speaker attention. While speaker attention generally increased vowel lengths, vowels were particularly longer in easy words when they were attended. In other words, attended speech boosted the difficulty effect on vowel duration. This finding not only bolsters the interpretation of the difficulty effect under Lindblom's (1990) Hypo- and Hyper-Articulation Theory, but also points to a possibility that the temporal fine-tuning in accordance with word-specificity may be also part of the phonetic feature that is enhanced in contextually-driven hyperarticulated speech.

In this respect, our results on vowel lengths, coupled with converging previous data obtained from different sets of words (Kilanski, 2009; Yun, 2010), call for future research on a possibly multi-layered encoding system between probabilistic information indexed at the lexical level—i.e., frequency and density that come into play in an early stage of speech production such as lexical access—and information at the phonological or articulatory level—i.e., phonotactic probability that would wield its influence in the following stage where articulatory gestures are executed.

As for notes on the L2 behavior, our L2 data revealed almost non-existent spectral and durational effects, demonstrating high-proficiency learners' limitations in acquiring vowel phonetic details. We argue that these limitations arise from L2-ers' insensitivity to lexically modified communicative functions, weaker word-level connections among phonologically similar words, and/or sheerly from sparse phonological links derived from relatively small-sized lexicon, all of which are related to experience-based phonological shaping in interaction with lexical use (Bybee, 2001; Goldinger, 1998; Hay & Foulkes, 2016; Pierrehumbert, 2001; Pierrehumber, 2002).

However, Yun's (2010) results obtained from intermediate-advanced level speakers should also be taken into account, particularly since they suggest that spectral expansion of hard words might be a characteristic that is acquired relatively easily compared to temporal contraction. The different results from the spectral analysis may have resulted from some differences in research design.

First, Yun (2010) examined a larger sample with a greater number of vowels contained the target words (10 categories vs. 6 categories). While he analyzed 7,360 L2 tokens (80 words×23 participants×4 repetitions), our L2 data had only 1,440 tokens. Thus, an L2 behavior conservatively generalizable from the two studies is

that L2 speakers may not be equally competent with all features associated with lexical accessibility in native speech. The absence of vowel-shortening in both studies may suggest that phonotactically-driven articulatory easiness is masked by L2-ers' surface unfamiliarity of hard words. Alternatively, as an anonymous reviewer suggested, it may also be related to the relatively weakened role of vowel length in the Korean vowel system, though both of these post-hoc hypotheses require further evidence.

Second, our data included utterances elicited in less prominent prosodic positions. As mentioned earlier, all tokens were made utterance-finally, and half of them were in the unattended condition, while Yun's tokens occurred in the middle of a fixed carrier sentence, "*Please say* [TARGET] *to me.*" While using this sort of carrier form is more susceptible to the observer's paradox, it may properly facilitate speakers' low-level sensitivity to implicit associations between probabilistically less accessible words and hyperarticulation. If so, a crucial difference with native speakers is deduced with respect to how lexical difficulty interacts with speech prosody. While native speakers modified phonetic forms in the spectral and temporal dimensions in prosodically weakened position, L2 speakers appear to lack the sensitivity in either dimension, given the same communicatively associated prosodic condition.

In conclusion, native English speakers in this study utilize lexical accessibility to modify vowels' sub-phonemic details in both spectral and temporal dimensions, in accordance with the communicative settings. However, all such probabilistically fine-grained shaping of vowel acoustics appears to be difficult to acquire for L2 speakers.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Bybee, J. (2001). *Phonology and language use*. Cambridge, UK: Cambridge University Press.

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Philadelphia, PA: John Benjamins.

Bybee, J. L. (2000). The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. 65-85). Stanford, CA: CSLI Publications.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*(4), 317-367.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627-635.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474-496.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251-279.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1998). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28*, 501-518.

Hay, J., & Foulkes, P. (2016). The evolution of medial /t/ over real and remembered time. *Language*, *92*(2), 298-330.

Kilanski, K. J. (2009). *The effects of token frequency and phonological neighborhood density on native and non-native English speech production* (Doctoral dissertation). University of Washington,

Seattle, WA.

Koo, H. S. (2000). Characteristics of English vowels spoken by Koreans. *Speech Sciences*, 7(3), 99-108.

Lee, M. (2021). *The effects of lexical competitor on Korean speakers' vowel articulation* (Mater's thesis). Pusan National University, Busan, Korea.

Lindblom, B. (1968). Temporal organization of syllable production. *Speech Transmission Laboratory Quarterly Progress Status Report*, 9(2-3), 1-5.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403-439). Dordrecht, The Netherlands: Kluwer.

Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon* (Doctoral dissertation). Indiana University, Bloomington, IN.

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

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.

Marslen-Wilson, W. D. (1989). *Lexical representation and process*. Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1990). Activation, competition and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305-315.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research*, 44, 778-792.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, 47(5), 1048-1058.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17(4), 273-281.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee, & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137-158). Amsterdam, The Netherlands: John Benjamins.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology 7* (pp. 101-139). Berlin, Germany: De Gruyter Mouton.

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, 4(1-3), 75-95.

Ussishkin, A., & Wedel, A. (2002). Neighborhood density and the root-affix distinction. *North East Linguistic Society*, 32(15), 15.

Vaden, K. I., Halpin, H. R., & Hickok, G. S. (2009). Irvine phonotactic online dictionary, version 2.0. Retrieved from http://www.iphod.com

Wright, R. (2004). Factors of lexical competition in vowel articulation. In J. Local, R. Ogden, R. Temple, M. E. Beckman, & J. Kingston (Eds.), *Papers in laboratory phonology VI: Phonetic interpretation*

(pp. 75-87). Cambridge, UK: Cambridge University Press.

Yun, G. (2010). The effects of neighborhood density and word frequency on the production of English-learning Korean speakers. *Studies in Phonetics, Phonology, and Morphology*, 16(2), 217-243.

• **Jonny Jungyun Kim,** Corresponding author
Assistant Professor, Dept. of English Language and Literature
Pusan National University
2 Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan 46241, Korea
Tel: +82-51-510-2038
Email: jonnykkim@gmail.com
Fields of interest: Phonetics, Phonology, Psycholinguistics, Sociophonetics

• **Mijung Lee**
Independent Researcher
Email: jude3000@naver.com
Fields of interest: Phonetics, Phonology