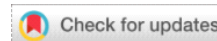




pISSN 2005-8063
eISSN 2586-5854
2026. 9. 30.
Vol.18 No.3
pp. 53-64



The effect of talker accent on perceptual cue weighting in English tense-lax vowels: Native speakers and Korean L2 learners*

Hyunjung So**

Department of English Language and Literature, Pusan National University, Busan, Korea

Abstract

Drawing on Exemplar Theory, this study investigates whether a talker's accent (American versus Korean) alone can activate their social background and associated linguistic features—specifically, Korean learners' reliance on duration cues when distinguishing English tense and lax vowels (/i/ and /ɪ/). Sixteen native English speakers and fourteen advanced Korean learners completed the two experiments. Experiment 1 established baseline cue weightings, whereas Experiment 2 primed the participants with accented carrier sentences to test their perceptual adjustments. In Experiment 1, native listeners relied on both spectral and duration cues, whereas the learners relied predominantly on duration. In Experiment 2, native listeners significantly increased their reliance on duration cues when hearing Korean-accented English compared with American-accented English, while their spectral weighting was unchanged. Learners showed reliable main effects of both spectrum and duration in this task, but adjusted neither cue across accent conditions. These results are consistent with native listeners drawing on accent as social-indexical information and adjusting their listening strategy accordingly, as predicted by Exemplar Theory, although alternative accounts cannot be ruled out. Learners' lack of adjustment may reflect the demands of second-language processing, but their already near-exclusive reliance on duration also left limited room for further increase.

Keywords: speech perception, perceptual cue weighting, Exemplar Theory, talker accent, Korean English learners

1. Introduction

The speech sounds we hear are not interpreted solely by relying on the primary cues of the physical signals entering our ears. As Labov (1966), a pioneer in sociolinguistics, revealed, linguistic

variation is not a random error but a systematic signal containing the talker's social identity. Because of this systematicity, the same sound can be perceived completely differently depending on who is speaking. This means that listeners actively utilize social information (e.g., age, gender, and region) embedded in phonetic

* This article is based on a part of the author's Master's thesis (So, 2025).

** minny980@pusan.ac.kr, Corresponding author

Received 11 August 2026; Revised 9 September 2026; Accepted 10 September 2026

© Copyright 2026 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

cues, such as a talker's accent, during the speech perception process.

Through what cognitive mechanism, then, are speech signals and social information integrated in the listener's mind? Exemplar Theory (Goldinger, 1998; Hay & Foulkes, 2016; Hay, 2018; Johnson, 1997) provides a convincing answer to this question. According to this theory, whenever we experience a specific utterance, we store it as a single vivid memory, or an 'exemplar.' This memory includes not only the acoustic features of the sound but also the talker's social information and the context in which the utterance occurred. These countless accumulated experiences form clusters based on similarity. Simply hearing a specific accent automatically activates the characteristics and entire language use patterns of the associated talker group, allowing us to flexibly adjust our perception.

This argument also has important implications for second language (L2) speech perception research. Zhang (2022), who compared the speech perception of native English speakers and Chinese learners of English, confirmed that native speakers' speech perception systems can be flexibly adjusted solely by the implicit cue of a 'foreign accent,' without explicit information. However, because this study only partially demonstrated native speakers' perceptual adjustment patterns, follow-up research is required to generalize the effect. Furthermore, empirical studies on how sociolinguistic cues, such as a talker's accent, function in the speech perception process of L2 learners rather than native speakers are still lacking. The present study aims to fill this research gap. By examining whether Zhang's (2022) findings can be generalized to a new linguistic environment involving Korean learners of English and Korean-accented English¹, this study seeks to support Zhang's (2022) results and further reveal the differences in perceptual mechanisms between native speakers and L2 learners. Specifically, this study addresses the following two questions:

1. Do native English listeners systematically adjust their cue weighting for tense-lax vowel perception depending on the talker's accent (American English vs. Korean English)?
2. Can highly proficient Korean learners demonstrate a level of perceptual flexibility similar to that of native speakers?

To answer these questions, this study followed the basic research design of Zhang (2022) while attempting a more in-depth analysis through several methodological improvements. In the main experiment (Experiment 2), a secondary task was added alongside the primary task to encourage participants to pay more attention to the talker's social information (accent). Additionally, the participants' 'amount of episodic memory' regarding Korean-accented English and their 'English listening proficiency' were measured to explore whether these variables moderated the results.

2. Theoretical Background and Research Hypotheses

2.1. The Effect of a Talker's Social Information on Speech Perception

When exploring the effect of a talker's social information on speech perception, many previous studies have presented "high-level explicit cues," such as facial photographs or regional origin labels (Drager, 2011; Hay et al., 2006; Koops et al., 2008; Niedzielski, 1999). In these types of studies, listeners clearly receive social categories that have already been conceptually processed. Therefore, without going through a separate cognitive step of inferring the talker's characteristics themselves, listeners activate immediate stereotypes or expectations about the talker based on the given information, which in turn influences their speech perception. From the perspective of Exemplar Theory, this explicit information directly activates exemplars related to a specific talker cluster (e.g., 'elderly talkers') stored in the listener's long-term memory. As a result, it has been extensively proven that a talker's age, ethnicity, and dialect cues induce perceptual biases in listeners or adjust their perceptual cue-weighting (Babel & Russell, 2015; Schertz et al., 2019).

Conversely, recent studies have demonstrated that a listener's perception can change based solely on low-level implicit cues inherent in the voice itself, without the presentation of explicit talker information (Kim, 2016; Kim & Drager, 2017, 2018; So & Kim, 2024; Zhang, 2022). So & Kim (2024) confirmed that cue-weighting during the perception of Korean stops changes immediately based solely on voices reflecting the talker's age, without visual cues. This suggests that auditory input alone can prime episodic memories where a talker's social information and linguistic features are bundled together, thereby influencing perception.

Zhang (2022), a foundational study for the present research, investigated how listeners' perception of tense-lax vowels changes depending on a talker's English accent. While native English speakers primarily use spectral cues to distinguish the two vowels, Chinese lacks this vowel contrast, causing Chinese learners of English to rely heavily on vowel duration cues. This study hypothesized that listeners store the social information of a 'Chinese learner of English' and the linguistic information that they 'primarily rely on vowel duration when distinguishing tense-lax vowels' as a single representation through numerous exemplars. The experimental results showed that while no changes in cue-weighting based on accent were observed in the Chinese learner group, native listeners significantly decreased their reliance on spectral cues when listening to Chinese-accented English. This empirically demonstrated that the implicit cue of a 'Chinese accent' alone, without explicit information about the talker's nationality, can prime relevant exemplars in native listeners' long-term memory and directly adjust the perceptual cue-weighting they rely on during vowel discrimination.

In this context, the present study aims to present low-level accent cues, similar to Zhang (2022), and expand the research scope to Korean learners of English and Korean-accented English to simultaneously compare native speakers and L2 learners. This study empirically investigates how a talker's accent information activates

¹ In this study, 'Korean-accented English' is defined as speech that systematically reflects Korean-specific phonological features caused by L1 interference such as vowel insertion and phoneme substitution (e.g., pronouncing 'find' as [p^haɪndi]).

linguistic knowledge within the cognitive mechanisms of both native speaker and learner groups, and what differences it brings to the adjustment of perceptual cue-weightings for tense–lax vowels.

2.2. Speech Perception of Second Language Learners

The English vowel system possesses a [tense] feature that distinguishes two phonemes, and tense and lax vowels are categorized based on the presence or absence of this feature. Phonetically, according to Yavas (2020), tense vowels are produced with a higher tongue position, a longer duration, and greater muscular tension than their corresponding lax counterparts. The English vowel inventory includes a total of four tense–lax pairs: the high front /i–ɪ/, the mid front /e–ɛ/, the high back /u–ʊ/, and the mid back /o–ɔ/. Native English speakers primarily rely on the first and second formant frequencies (spectral cues) when distinguishing between tense and lax vowels, utilizing vowel duration as a secondary cue (Hillenbrand et al., 2000).

Unlike English, Korean lacks tense–lax vowel pairs and thus does not require the 'tense' feature to distinguish between two vowels. Consequently, Korean learners of English generally experience difficulty distinguishing English tense–lax vowels using acoustic features like the formant frequencies used by native speakers (Kang, 2022; Yang, 2010). To compensate for this, lower-proficiency English learners depend excessively on vowel duration cues (Flege et al., 1997; Lee & Rhee, 2019).

In Zhang (2022), when the talker's regional background information was provided as an accent cue, native speakers flexibly adjusted their cue-weighting by co-activating the talker's social characteristics and relevant linguistic information. In contrast, no such change in perceptual patterns according to accent was observed among Chinese learners. Zhang (2022) explained the causes of this difference in two main ways.

First, within the cognitive system of L2 learners, the core information of exemplars is only activated to a limited extent (Drager & Kirtley, 2016). Learners tend to focus solely on the acoustic cues (spectrum and duration) essential for task performance, failing to sufficiently activate secondary cues such as the talker's region of origin or social information.

Second, previous work has attributed this to the processing demands of listening in a second language (Heald & Nusbaum, 2014). According to Nijveld et al. (2022), L2 speakers face substantial processing demands, making it difficult for them to utilize social information, such as a talker's voice, as exemplars. Additionally, Lee & Kong (2024) found that in environments with divided attention, native speakers display a compensatory mechanism by flexibly shifting cue-weightings (e.g., increasing F0 reliance when VOT reliance decreases), whereas L2 learners exhibit a simultaneous drop in reliance on all cues. This suggests that when cognitive capacity is lacking, learners are unable to shift cue-weightings to suit the context.

In summary, previous work has attributed this pattern to a combination of factors—limited exemplar activation stemming from insufficient language experience, the processing demands of foreign-language tasks, and a consequent difficulty in shifting cue weights—such that L2 learners have been reported not to show the flexible, socio-linguistic context-based perceptual adjustment exhibited by native speakers.

2.3. Research Hypotheses

The specific research hypotheses for this study, derived based on the previously discussed theoretical background and the results of previous studies, are as follows.

First, native English listeners are predicted to exhibit perceptual flexibility based on the talker's accent information. According to Exemplar Theory, the 'Korean-accented English' stimulus will simultaneously activate the social category of 'Korean talkers' and the linguistic information that they 'primarily rely on duration cues when distinguishing vowels.' Therefore, it is hypothesized that listeners will modify their perceptual strategy by decreasing their reliance on spectral cues and increasing their reliance on duration cues when listening to Korean-accented English compared to the American-accented English condition.

Second, following Zhang (2022), the advanced Korean learner group is not expected to show a distinct change in cue-weighting according to the talker's accent. Previous work has attributed this pattern to restricted exemplar activation and to the processing demands of second-language listening.

To verify this, all participants sequentially completed Experiment 1 to establish a baseline and Experiment 2 to test the core hypotheses. After completing Experiment 2, a talker nationality evaluation task was commonly conducted to be used as a primary variable in the results analysis. Subsequently, the native speaker group completed a post-survey, while the Korean learner group completed a dictation task to secure additional variables before responding to the survey. All experimental procedures were conducted in a single session lasting approximately 40 to 50 minutes per participant, without any separate breaks.

3. Experiment 1

3.1. Experiment 1 Design

Experiment 1 is a preliminary experiment conducted to establish a baseline for the main experiment (Experiment 2). It was conducted as a forced-choice identification task to determine the extent to which the two groups (native English speakers and advanced Korean learners) rely on spectral and duration cues, respectively, when distinguishing English tense–lax vowels. In line with previous research, it was predicted that along with the native speaker group, the highly proficient learner group would also primarily rely on spectral cues (Lee & Rhee, 2019; Zhang, 2022).

As for the participants, 16 native speakers of English (3 males, 13 females) aged 35.6 years on average ($SD=13.6$, range 21–69) were recruited. To limit variation across English varieties, recruitment was restricted to speakers from the US ($n=10$), the UK ($n=3$), and Canada ($n=3$); speakers of other varieties were not included. Native listeners had lived in Korea for 6.4 years on average ($SD=5.9$, range 0–15). Fourteen advanced Korean learners of English (9 males, 5 females), aged 25.7 years on average ($SD=5.4$, range 20–43), also took part. The Korean participants consisted of individuals with TOEIC scores of 950 or higher (12 participants) or Korean-English bilinguals (2 participants) to ensure a high level of proficiency.

For the experimental stimuli, utterances of 'seat (/i/)–sit (/ɪ/)' produced by a female native English speaker in her 20s were recorded. Subsequently, only the spectral and durational variables of the vowel segment were manipulated using Praat (Boersma & Weenink, 2025). The spectrum was synthesized into a 7-step

continuum utilizing a Praat script (Winn, 2023), and the duration was manipulated into 7 steps at 30 ms intervals within the same range (80 to 260 ms) as Zhang (2022) using the PSOLA algorithm. By crossing these variables, a total of 49 stimuli (7-step spectrum×7-step duration) were created.

All experimental procedures were conducted using E-Prime (3.0). To minimize researcher influence, all instructions were presented on-screen in English. Prior to the main experiment, practice trials were conducted using unmanipulated natural speech from a different talker, allowing participants to acclimate to the task. During the main experimental trials, the two words 'seat' and 'sit' were visually presented on the screen. After listening to the acoustically manipulated speech stimulus, participants identified one of the two words by pressing designated keys using their index fingers. The inter-stimulus interval was set to 1,000 ms, and the 49 stimuli were randomly presented twice, completing a total of 98 trials per participant.

3.2. Experiment 1 Results

The experimental results were analyzed with mixed-effects logistic regression using the lme4 package in R (version 4.6.1). To compare the two groups, the fixed effects included the main effects of spectral step, duration step, and group, together with all interactions among them. Continuous predictors (spectral step, duration step) were mean-centered, and group was sum-coded, -0.5 (Korean learners) and $+0.5$ (native listeners). The random-effects structure included by-participant random intercepts and random slopes for spectral step and duration step, with correlations excluded. To identify detailed patterns within each group, the data were then separated into subsets and independent models with the identical structure—excluding only the 'group' variable—were fitted.

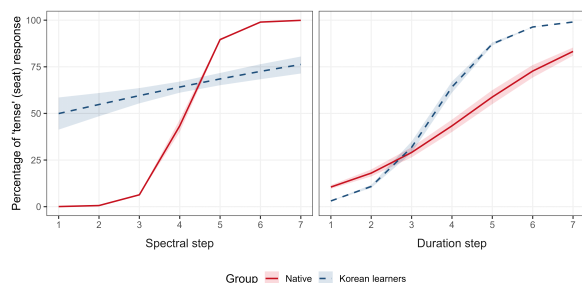


Figure 1. Cue weighting in Experiment 1. Model-predicted probability of a 'tense' (seat) response by continuum step and listener group in Experiment 1. Shaded bands show 95% confidence intervals.

Table 1 reports the combined model. The main effects of both

spectral step ($\beta=1.308$, $p<.001$) and duration step ($\beta=0.985$, $p<.001$) were significant, indicating that as the steps of both cues increased, the proportion of tense vowel responses increased. The main effect of group was also significant ($\beta=-0.850$, $p=.035$), with native listeners responding 'lax' more often than learners.

Looking at the interaction between the cues, a significant 'spectral step×duration step' interaction ($\beta=0.060$, $p=.023$) was confirmed, indicating that tense vowel responses increased as both spectral step and duration step simultaneously increased. Furthermore, differences in cue weighting by group were clearly evident. The native speaker group relied significantly more on spectral cues ($\beta=2.227$, $p<.001$) than the Korean learner group, while their reliance on duration cues was lower ($\beta=-0.721$, $p=.004$). These differences in cue weighting between groups are prominent in Figure 1: native speakers exhibited a steep categorical perception pattern in response to spectral changes, whereas Korean learners showed a steeper slope in the duration graph. Moreover, the three-way interaction of 'spectral step×duration step×group' ($\beta=0.152$, $p=.004$) was also significant, demonstrating that the effect of the interaction between the two cues was manifested more strongly in the native speaker group.

Additionally, follow-up analyses of each group separately revealed that for the native speaker group, the main effect of spectral step ($\beta=2.333$, $p<.001$), the main effect of duration step ($\beta=0.602$, $p<.001$), and the interaction between the two cues ($\beta=0.130$, $p=.005$) were all significant. In contrast, for the Korean learner group, only the main effect of duration step was significant ($\beta=1.440$, $p<.001$), while the main effect of spectral step did not reach statistical significance.

3.3. Discussion of Experiment 1

The results of Experiment 1 revealed a distinct difference in the utilization of cues for tense-lax vowel discrimination between the native English speakers and the advanced Korean learners. While native speakers utilized both spectrum and duration, the learner group relied heavily only on the duration cue. This contrasts with the findings of Zhang (2022), where advanced learners also primarily relied on spectral cues, and three potential causes can be considered for this.

First is the relative salience of the duration cue. Due to differences in stimulus creation methods (i.e., the use of a Praat script), this study employed a narrower range of the spectral continuum compared to Zhang (2022). Since the duration was manipulated identically, the discriminability of the spectral cue might have been relatively weakened. This is supported by the fact that, unlike the native speaker group in Zhang (2022), even the native speaker group in this study significantly utilized the duration cue.

Table 1. Fixed effects for the combined model (Experiment 1)

Term	β	SE	z-value	p-value	95% CI
(Intercept)	0.157	0.201	0.78	.436	[-0.238, 0.552]
Spectral step	1.308	0.162	8.08	<.001	[0.991, 1.626]
Duration step	0.985	0.126	7.81	<.001	[0.737, 1.232]
Group=Native	-0.850	0.403	-2.11	.035	[-1.640, -0.060]
Spectral step×Duration step	0.060	0.026	2.28	.023	[0.008, 0.112]
Spectral step×Group=Native	2.227	0.320	6.95	<.001	[1.599, 2.855]
Duration step×Group=Native	-0.721	0.249	-2.90	.004	[-1.209, -0.233]
Spectral step×Duration step×Group=Native	0.152	0.053	2.90	.004	[0.049, 0.256]

Second is the limitation of the participant selection metric. Although this study selected advanced learners with "TOEIC scores of 950 or higher," high TOEIC scores—which heavily involve contextual understanding and test-taking strategies—may not fully represent the low-level perceptual ability required to capture the fine acoustic cues of tense-lax vowels.

Third, despite their high English proficiency, Korean learners may insufficiently utilize spectral cues compared to native speakers when discriminating tense-lax vowels. According to several previous studies, even when Korean learners have resided in the US for an extended period or possess top-tier proficiency, they fail to produce and perceive spectral cues as adequately as native speakers do during tense-lax vowel discrimination, and they find it difficult to overcome their dominant reliance on duration cues (Flege et al., 1997; Lee, 2008; Yang, 2010).

In conclusion, since the advanced learners in this study demonstrated a remarkably low utilization of spectral cues, it is difficult to expect a significant change in spectral reliance according to accent in Experiment 2. We note one further possibility, formulated after the Experiment 1 data had been collected and therefore departing from the hypothesis stated in Section 2.3: duration, as the cue learners rely on most, might respond to accent information even where spectral reliance does not. If hearing Korean-accented English activates the linguistic information associated with Korean talkers, learners too might weight duration more heavily in that condition than under American-accented English.

4. Experiment 2

4.1. Experiment 2 Design

Experiment 2 is the core part of this study, designed to verify whether listeners activate exemplars associated with a talker's social information solely through accent cues, and consequently adjust their cue weightings for spectrum and vowel duration.

4.1.1. Stimuli recording and acoustic manipulation

The experimental stimuli consisted of the 'seat-sit' continuum, identical to Experiment 1, recorded in American English by four female Korean-English bilinguals in their 20s and 30s. A '7-step spectrum×7-step duration' continuum was created in the same manner as in Experiment 1. The endpoints of the spectral continuum were set separately for each talker so that all seven steps would sound natural rather than synthetic (Talker 1: F1 592–218 Hz, F2 2,047–2,694 Hz; Talker 2: F1 299–224 Hz, F2 2,121–2,908 Hz; Talker 3: F1 526–248 Hz, F2 2,105–2,764 Hz; Talker 4: F1 451–244 Hz, F2 2,184–2,823 Hz), and intermediate steps were interpolated in equal intervals on the Bark scale using a Praat script (Winn, 2023). The duration range was standardized across all talkers (80–260 ms), and the final /t/ was released in all stimuli, with its closure duration fixed at 100 ms to control for the intervention of other acoustic variables.

The carrier sentences preceding the target stimuli ('Find the word', 'Focus on the word', 'Go with the word') were recorded by the four talkers in two accents: American English and Korean-accented English. While the American English recordings were made to closely imitate native speakers' utterances, the Korean-accented English deliberately reflected typical first language (L1) interference

phenomena of Korean learners at the syllable structure, phoneme, and prosody levels as follows.

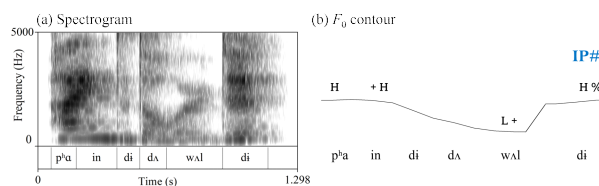


Figure 2. Talker 1's production of 'Find the word' reflecting features of Korean-accented English. (2) Spectrogram(b) F0 contour.

4.1.1.1. Syllable structure and phoneme levels

Unlike English, which allows complex consonant clusters (e.g., CCCVCCCC), the maximum syllable structure in Korean is CVC, with a constraint that fricatives cannot appear in the coda position. Reflecting this, when pronouncing consonant clusters, an epenthetic vowel (/i/) was inserted, or coda fricatives were resyllabified into the onset (e.g., find→[pʰaɪndi], focus→[pʰokasi]). In addition, target English phonemes were substituted with similar Korean phonemes. The voiceless labiodental fricative /f/ was pronounced as the bilabial aspirated stop /pʰ/, and the voiced dental fricative /ð/ was pronounced as the allophone [d] of the alveolar stop. Reflecting the complementary distribution of Korean liquids, the retroflex /ɻ/ in the coda was substituted with the lateral approximant /l/ (e.g., word → [wʌldi]). Figure 2 shows Talker 1's actual production of "Find the word" reflecting these features, where the previously described phonetic changes can be seen in the spectrogram (a).

4.1.1.2. Prosody level

According to Jun's (2014) prosodic typology, English is a head-prominence and stress-timed language that indicates prominence through nuclear pitch accents in a phrase. In contrast, Korean is an edge-prominence and syllable-timed language that forms accentual phrases (AP) through boundary tones. Because beginning learners transfer this L1 prosodic system directly to their English utterances (Choe, 2016), this study strictly reflected the following two prosodic features in the experimental stimuli.

First, utilizing the syllable-timed characteristic of Korean, the English stress rhythm was excluded, and the talkers were guided to articulate segmentally, pronouncing each syllable clearly with a consistent duration. Second, Korean-specific AP boundary tone rules were applied to the English sentences. Depending on the laryngeal feature of the onset in the first syllable of the AP, sentences beginning with a [+stiff vocal folds] onset, such as the aspirated stop /pʰ/, were realized with an HH initial boundary, while those beginning with a [-stiff vocal folds] onset, such as the lenis stop /k/, were realized with an LH boundary.

How these prosodic features are realized in actual production can be seen in the pitch contour in Figure 2(b). Specifically, this figure illustrates Talker 1's utterance of 'Find the word,' visually demonstrating that all syllables are produced with equal timing, without the application of English stress. Furthermore, it clearly confirms the realization of the HH initial boundary tone, triggered by the phrase-initial aspirated stop /pʰ/.

Additionally, all three sentences were uttered as a single

Intonational Phrase (IP), implementing a high boundary tone (H%) and IP final lengthening at the end of the sentence. A 150 ms silence was placed before the target stimulus to ensure it always appeared at a new IP initial boundary. How these prosodic features are realized in the actual production can be seen in the contour in Figure 2(b).

4.1.1.3. Acoustic comparison of the carrier sentences

The two accent conditions differed systematically in their temporal properties. Korean-accented carriers were longer overall ($M=1,265$ ms, $SD=121$) than American-accented carriers ($M=994$ ms, $SD=142$), $t(11)=11.29$, $p<.001$, $d_z=3.26$. Because vowel epenthesis increased the number of syllables from 4.0 to 6.3 on average, however, mean syllable duration was shorter in the Korean-accented condition (200 ms vs 253 ms), $t(11)=-7.33$, $p<.001$, $d_z=-2.12$, and articulation rate was correspondingly faster (5.03 vs 4.00 syll/s), $t(11)=7.82$, $p<.001$, $d_z=2.26$. All twelve talker-sentence pairs showed the same direction on every measure (Wilcoxon $p=.0005$). The instruction to produce syllable-timed rhythm therefore did not slow the Korean-accented carriers; the increase in syllable count outweighed the increase in overall duration. Since the rate of a preceding context can shift the perceived duration of a following vowel (Bosker & Reinisch, 2017), this difference constitutes a potential alternative account of any accent-related change in duration weighting and is addressed in the analyses reported in Section 4.4.

Table 2. Carrier sentence duration and articulation rate

Talker	Carrier	Syllables		Total duration (ms)		Articulation rate (syll/s)	
		Am	Ko	Am	Ko	Am	Ko
F1	Find	3	6	863	1,124	3.48	5.34
F1	Focus	5	7	1,199	1,429	4.17	4.90
F1	Go	4	6	1,033	1,325	3.87	4.53
F2	Find	3	6	789	1,218	3.80	4.93
F2	Focus	5	7	1,092	1,315	4.58	5.32
F2	Go	4	6	1,036	1,321	3.86	4.54
F3	Find	3	6	781	1,128	3.84	5.32
F3	Focus	5	7	1,061	1,291	4.71	5.42
F3	Go	4	6	971	1,058	4.12	5.67
F4	Find	3	6	848	1,183	3.54	5.07
F4	Focus	5	7	1,098	1,389	4.55	5.04
F4	Go	4	6	1,153	1,395	3.47	4.30
	Mean	4.0	6.3	994	1,265	4.00	5.03
	SD	-	-	142	121	0.43	0.41

Syllable counts reflect the vowel epenthesis implemented in the Korean-accented productions (Section 4.1.1.1). Durations were measured in Praat from the onset of the first segment to the offset of the last, excluding pre- and post-utterance silence. Am, American-accented; Ko, Korean-accented; Find, “Find the word”; Focus, “Focus on the word”; Go, “Go with the word”.

Syllable counts reflect the vowel epenthesis implemented in the Korean-accented productions (Section 4.1.1.1). Durations were measured in Praat from the onset of the first segment to the offset of the last, excluding pre- and post-utterance silence.

Am, American-accented; Ko, Korean-accented; Find, “Find the

word”; Focus, “Focus on the word”; Go, “Go with the word”.

4.1.2. Experimental procedure

Experiment 2 was designed with both a ‘target task’ and a ‘secondary task’ to encourage participants to voluntarily pay attention to the talker’s accent. Like Experiment 1, the target task was a forced-choice identification task. However, participants first listened to an accented carrier sentence before identifying the manipulated ‘seat-sit’ stimulus.

The secondary task was randomly presented once every 10 trials of the target task. It consisted of choosing one of two options regarding the phonetic and social characteristics of the talker’s voice heard in the immediately preceding target task (24 trials in total; e.g., “How fast is the talker speaking?” [Fast vs. Slow], “What is the talker’s socioeconomic status?” [Upper class vs. Lower class]). This encouraged participants to focus on the talker’s accent information, facilitating the activation of relevant exemplars, while simultaneously concealing the actual purpose of the study to minimize the demand characteristics (Orne, 1962).

Furthermore, to prevent confusion regarding the talker’s nationality that might arise if a single talker used both accents, the participants were divided into two counterbalancing lists. List 1 was presented with Talkers 1 and 2 in an American accent and Talkers 3 and 4 in a Korean accent, while List 2 received the reverse. This ensured that each participant was exposed to only a single accent for any given talker. Consequently, each participant completed a total of 220 trials (4 talkers×49 stimuli+24 secondary tasks), and the 3 carrier sentences were randomly distributed across the stimuli.

4.2. Talker Nationality Evaluation Task

This task was conducted to secure a predictor variable for analyzing the effect of the participants’ amount of episodic memory regarding Korean-accented English on the perceptual cue-weighting changes in Experiment 2. The experimental stimuli were created using an AI text-to-speech platform (‘Natural Reader’). Twelve sentences that clearly reflected major phonological phenomena of Korean-accented English (e.g., phoneme substitution, epenthetic vowel insertion, absence of flapping for intervocalic /t/, transfer of Korean nasal assimilation rules) were selected. Furthermore, to minimize the demand characteristics that occur when participants discern the study’s purpose, 36 distractor stimuli synthesized in English with accents from 8 other countries (e.g., US, UK, France, Japan) were included, presenting a total of 48 sentences randomly.

After listening to the voice in each trial, participants selected the talker’s nationality from a provided list and also rated their confidence in this choice. To evaluate participants’ sensitivity specifically to Korean-accented English, nationality-evaluation performance was scored as d' with the log-linear correction (Hautus, 1995), separating sensitivity from response bias. For this calculation, hits were based on accurately identifying the 12 Korean-accented sentences, and false alarms were based on incorrectly identifying the 36 distractor sentences as Korean. False alarms were near floor (native $M=0.12/36$; learners $M=0.00/36$). Learners’ d' was both higher and less variable than native listeners’ (3.77, $SD=0.36$ vs. 3.06, $SD=0.75$), with ten of fourteen learners at ceiling on hits. Consequently, this d' measure was utilized as a moderator solely for the native listener group. Confidence

ratings (0–3) confirmed that responses were not arbitrary: on the Korean-accented items, ratings were higher when the talker was correctly identified as Korean than when not (native listeners: 2.52 vs 1.61; learners: 2.83 vs 1.17).

4.3. Dictation Task Results for Korean English Learners

To measure the detailed listening proficiency within the advanced Korean learner group, an additional dictation task was conducted after Experiment 2 to obtain a further predictor for the learner models. Ten sentences were constructed using listening items from an English vocabulary learning application (Malhaeboka). Participants performed the dictation after listening to the audio up to two times per item. During grading, any allowance for semantic similarity was excluded, and spelling accuracy was strictly used as the criterion to calculate the ratio of the 'number of accurately written words to the total number of words.' Based on this, the average value for the 10 sentences was calculated. The results showed that the Korean learners' average dictation score was 52.33 (median: 43.8, range: 21.8 to 83.8).

4.4. Experiment 2 Results

Responses were analyzed as in Experiment 1. A combined model including both groups was fitted first to test whether cue weighting differed between them, followed by separate models for each group. The fixed effects of the combined model comprised the main effects of spectral step, duration step, accent, and group, together with all higher-order interactions among them. Talker and counterbalancing list were included as control predictors. Accent, group, and talker were sum-coded: accent was coded -0.5 (American-accented) and $+0.5$ (Korean-accented), group -0.5 (Korean learners) and $+0.5$ (native listeners), and talker with Talker 1 as the reference level. The group-specific models used the same structure with group omitted. The random-effects structure included by-participant random intercepts and random slopes for spectral step, duration step, accent, and the spectral step $\times$ accent interaction, with correlations excluded. The

by-participant slope for the duration step $\times$ accent interaction was estimated at zero variance in every model and was removed (Matuschek et al., 2017).

Individual-difference measures were tested as moderators and retained only if they improved model fit, assessed by likelihood-ratio test. For native listeners, adding the d' $\times$ cue $\times$ accent interactions did not improve fit, $\chi^2(2)=0.47$, $p=.791$, and neither did the Korean residence $\times$ duration step $\times$ accent interaction, $\chi^2(1)=0.77$, $p=.379$. We note that with the d' terms retained, the duration step $\times$ accent coefficient is estimated at the mean of d' rather than as an average effect ($\beta=0.124$, $SE=0.067$, $p=.063$); we therefore report the model without them. For learners, the dictation score $\times$ cue $\times$ accent interactions likewise did not improve fit, $\chi^2(2)=0.63$, $p=.729$. None of these terms was retained.

4.4.1. Combined model

Table 3 reports the combined model. Both cues had reliable main effects (spectral step: $\beta=1.054$, $p<.001$; duration step: $\beta=0.584$, $p<.001$), and cue weighting differed sharply between groups: native listeners relied more on spectral step ($\beta=1.092$, $p<.001$) and less on duration step ($\beta=-0.585$, $p<.001$) than learners. Listeners also responded 'lax' more often under the Korean-accented carrier ($\beta=-0.503$, $p=.012$).

The two interactions bearing on the central question were not reliable. Neither the duration step $\times$ accent $\times$ group interaction ($\beta=0.141$, $SE=0.089$, $p=.115$) nor the spectral step $\times$ accent $\times$ group interaction ($\beta=-0.017$, $SE=0.156$, $p=.913$) reached significance. Contrary to the hypothesis, then, the present sample provides no evidence that the two groups differed in how accent affected cue weighting, though with 16 and 14 listeners the design also had limited power to detect such a difference. The group-specific analyses below are accordingly reported as descriptions of each group rather than as a demonstration that the groups differ.

4.4.2. Native listeners

Table 4 and Figure 3 report the native listener model. Spectral

Table 3. Fixed effects for the combined model (Experiment 2)

Term	β	SE	z-value	p-value	95% CI
(Intercept)	0.460	0.414	1.11	.266	[-0.351, 1.270]
Spectral step	1.054	0.102	10.38	<.001	[0.855, 1.253]
Duration step	0.584	0.061	9.60	<.001	[0.465, 0.704]
Accent=Korean	-0.503	0.199	-2.52	.012	[-0.893, -0.112]
Group=Native	-0.558	0.267	-2.09	.036	[-1.081, -0.036]
Talker 2	0.403	0.114	3.53	<.001	[0.179, 0.627]
Talker 3	-1.189	0.117	-10.14	<.001	[-1.419, -0.959]
Talker 4	1.278	0.117	10.94	<.001	[1.049, 1.508]
List	-0.352	0.266	-1.32	.186	[-0.873, 0.169]
Spectral step $\times$ Duration step	0.007	0.013	0.52	.602	[-0.019, 0.033]
Spectral step $\times$ Accent=Korean	0.055	0.078	0.70	.485	[-0.098, 0.208]
Duration step $\times$ Accent=Korean	0.067	0.045	1.51	.131	[-0.020, 0.155]
Spectral step $\times$ Group=Native	1.092	0.202	5.40	<.001	[0.695, 1.488]
Duration step $\times$ Group=Native	-0.585	0.121	-4.82	<.001	[-0.823, -0.347]
Accent=Korean $\times$ Group=Native	-0.201	0.399	-0.50	.614	[-0.983, 0.581]
Spectral step $\times$ Duration step $\times$ Accent=Korean	-0.003	0.027	-0.12	.904	[-0.055, 0.049]
Spectral step $\times$ Duration step $\times$ Group=Native	-0.046	0.027	-1.73	.084	[-0.099, 0.006]
Spectral step $\times$ Accent=Korean $\times$ Group=Native	-0.017	0.156	-0.11	.913	[-0.323, 0.289]
Duration step $\times$ Accent=Korean $\times$ Group=Native	0.141	0.089	1.58	.115	[-0.034, 0.316]
Spectral step $\times$ Duration step $\times$ Accent=Korean $\times$ Group=Native	0.058	0.053	1.09	.274	[-0.046, 0.162]

Talker was sum-coded with Talker 1 as the reference level.

Table 4. Fixed effects for the Native model (Experiment 2)

Term	β	SE	z-value	p-value	95% CI
(Intercept)	0.186	0.551	0.34	.736	[-0.894, 1.265]
Spectral step	1.548	0.128	12.10	<.001	[1.298, 1.799]
Duration step	0.279	0.053	5.26	<.001	[0.175, 0.383]
Accent=Korean	-0.607	0.195	-3.12	.002	[-0.989, -0.226]
Talker 2	-0.018	0.126	-0.15	.884	[-0.266, 0.229]
Talker 3	-0.687	0.130	-5.31	<.001	[-0.941, -0.433]
Talker 4	1.432	0.132	10.84	<.001	[1.173, 1.691]
List	-0.359	0.347	-1.04	.300	[-1.039, 0.321]
Spectral step×Duration step	-0.018	0.020	-0.90	.369	[-0.058, 0.021]
Spectral step×Accent=Korean	0.050	0.152	0.33	.744	[-0.249, 0.349]
Duration step×Accent=Korean	0.135	0.061	2.21	.027	[0.015, 0.254]
Spectral step×Duration step×Accent=Korean	0.027	0.040	0.67	.501	[-0.051, 0.105]

Talker was sum-coded with Talker 1 as the reference level.

Table 5. Fixed effects for the Korean learner model (Experiment 2)

Term	β	SE	z-value	p-value	95% CI
(Intercept)	0.860	0.652	1.32	.188	[-0.419, 2.138]
Spectral step	0.527	0.161	3.27	.001	[0.211, 0.844]
Duration step	0.923	0.126	7.33	<.001	[0.676, 1.170]
Accent=Korean	-0.328	0.238	-1.38	.167	[-0.794, 0.138]
Talker 2	0.883	0.143	6.16	<.001	[0.602, 1.163]
Talker 3	-1.713	0.151	-11.38	<.001	[-2.008, -1.418]
Talker 4	1.029	0.144	7.13	<.001	[0.746, 1.312]
List	-0.419	0.429	-0.98	.329	[-1.259, 0.422]
Spectral step×Duration step	0.033	0.017	1.98	.048	[0.000, 0.066]
Spectral step×Accent=Korean	0.060	0.086	0.70	.484	[-0.108, 0.228]
Duration step×Accent=Korean	-0.015	0.065	-0.23	.821	[-0.143, 0.113]
Spectral step×Duration step×Accent=Korean	-0.032	0.033	-0.97	.333	[-0.098, 0.033]

Talker was sum-coded with Talker 1 as the reference level.

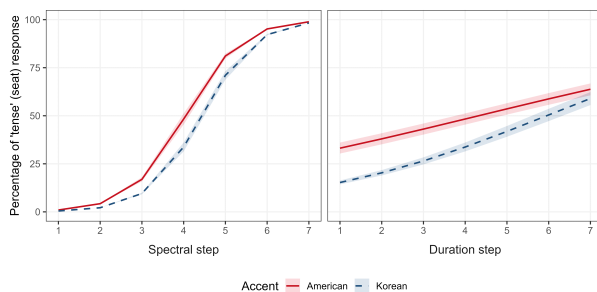


Figure 3. Native listeners' cue weighting by accent. Model-predicted probability of a 'tense' (seat) response by continuum step and accent condition, native listeners. Shaded bands show 95% confidence intervals.

step carried substantially more weight than duration step ($\beta=1.548$ vs 0.279 , both $p<.001$), and listeners responded 'lax' more often under the Korean-accented carrier ($\beta=-0.607$, $p=.002$). The main effect of talker was also reliable (Talker 3: $\beta=-0.687$, $p<.001$; Talker 4: $\beta=1.432$, $p<.001$).

Talker was sum-coded with Talker 1 as the reference level.

Turning to the interaction between accent and the acoustic cues, the spectral step×accent interaction was not reliable ($\beta=0.050$, $p=.745$), whereas the duration step×accent interaction was ($\beta=0.135$, $SE=0.061$, $p=.027$). The duration-to-spectral weight ratio rose from .14 to .22 across conditions, a 59% increase, while the spectral coefficient itself did not change; the shift was therefore selective rather than global. That is, listeners relied more heavily

on duration cues when listening to Korean-accented English compared to the American-accented English condition (Figure 3); this result supports the hypothesis of the present study and previous research (Zhang, 2022) demonstrating that perceptual weighting is adjusted upon exposure to a non-native accent.

As reported in Section 4.1.1.3, the Korean-accented carriers were faster than the American-accented carriers. To assess whether this temporal difference could account for the reweighting described above, articulation rate was included as a covariate in the model. The duration step×accent interaction remained unchanged ($\beta=0.134$, $SE=0.061$, $p=.028$). The accent main effect was likewise unaffected, and articulation rate itself was not reliable ($\beta=0.070$, $p=.661$).

4.4.3. Korean learners

Regarding the results for the Korean English learner group, the main effects of both spectral step ($\beta=0.527$, $p=.001$) and duration step ($\beta=0.923$, $p<.001$) were significant (Table 5), with duration step carrying considerably more weight than spectral step. The two cues interacted ($\beta=0.033$, $p=.048$). The main effect of accent did not reach significance ($\beta=-0.328$, $p=.167$).

As predicted, neither the spectral step×accent ($\beta=0.060$, $p=.484$) nor the duration step×accent interaction ($\beta=-0.015$, $SE=0.065$, $p=.821$) approached significance (Figure 4). The duration-to-spectrum weight ratio was 1.87 in the American-accented condition and 1.64 in the Korean-accented condition—a small change in the direction opposite to that observed for native listeners. The learner group thus showed no reliable adjustment of

cue weighting across accent conditions, replicating the learner pattern reported by Zhang (2022).

Excluding the two Korean–English bilinguals left this pattern unchanged: the duration step×accent ($\beta=-0.030$, $p=.674$) and spectral step×accent interactions ($\beta=0.044$, $p=.633$) remained non-significant, and duration step continued to outweigh spectral step ($\beta=0.950$ vs 0.560).

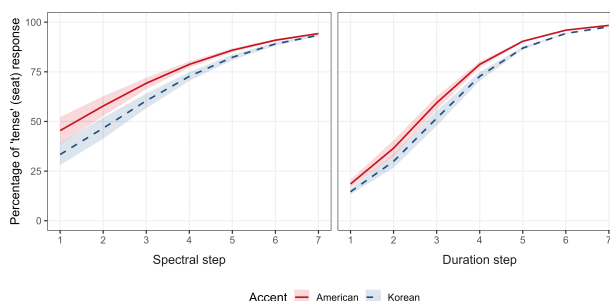


Figure 4. Korean learners' cue weighting by accent. Model-predicted probability of a 'tense' (*seat*) response by continuum step and accent condition, Korean learners. Shaded bands show 95% confidence intervals.

4.5. Discussion of Experiment 2

4.5.1. Cue-weighting adjustment in the native speaker group

The native speaker group significantly increased their reliance on duration cues when listening to Korean-accented English compared to American-accented English. The change was specific to duration: the weight given to spectral cues stayed the same across the two conditions, and the ratio of duration weight to spectral weight rose from .14 to .22. Listeners therefore did not simply become more decisive overall; they shifted how much they relied on one particular cue. As no change was observed in spectral cues, this partially supports the hypothesis. Compared to Zhang (2022), where spectral reliance decreased instead, the finding suggests that accent-based perceptual tuning can take more than one form.

This pattern is what Exemplar Theory predicts. Hearing a Korean accent activates stored exemplars of Korean talkers, which encode not only the acoustic form of past utterances but also the information, accumulated across many such encounters, that these talkers mark the tense–lax distinction mainly through vowel length. Listeners then weight duration more heavily when judging the following vowel.

Two alternatives cannot be excluded, however. The first requires no knowledge of Korean talkers at all. Vowels in accented speech are produced less precisely, so their formant values are a less reliable guide to category membership. Listeners are known to down-weight an acoustic dimension when its relationship to category membership becomes less dependable, and to do so within a few trials of exposure (Idemaru & Holt, 2011); the same mechanism would produce the pattern observed here. Distinguishing this from exemplar activation requires a control accent whose speakers do not themselves rely on duration, which the present study did not include. The second concerns where the effect is located. The accent also shifted how often listeners chose *sit* over *seat* overall, and this shift was larger and more consistent ($\beta=-0.607$, $p=.002$) than the reweighting itself ($\beta=0.135$, $p=.027$).

Part of the accent effect therefore lies in the decision criterion rather than in cue weighting.

The temporal difference between the two carrier conditions offers a further possibility, since the Korean-accented carriers were spoken faster (Section 4.1.1.3). Controlling for articulation rate, however, left the reweighting unchanged ($\beta=0.134$, $p=.028$), so the effect does not reduce to a temporal contrast effect.

The nationality-evaluation measure did not predict how much listeners adjusted their cue weighting. Two features of that task make this hard to interpret. It used synthesized speech while Experiment 2 used natural recordings, so the measure was taken on one set of materials and applied to another; and only 12 of its 48 trials presented Korean-accented English. Although a statistical moderating effect from the evaluation task was absent, this does not imply that the native speakers failed to link the Korean-accented English to a Korean talker during the actual experiment. Considering that the majority of native speakers accurately evaluated the talkers who spoke Korean-accented English in Experiment 2 as 'Korean' in the post-survey, the possibility that a specific 'Korean learner' exemplar—rather than a 'general non-native' one—was activated within their cognitive mechanisms remains open.

4.5.2. The absence of adjustment in the Korean learner group

The Korean English learner group showed no change in cue weighting across accent conditions, replicating the learner pattern reported by Zhang (2022). Zhang attributed this to limited exemplar activation in second-language listening. Learners appear to activate the acoustic properties of stored episodes without also activating the talker-related information bound up with them, so that accent does not function as a cue to how the contrast will be marked. The present results are consistent with this account, although the mechanism itself was not measured here and the interpretation is carried over from that study.

The result is open to more than one reading. Learners may simply have had little scope for further adjustment: before any accent was introduced, they already weighted duration about twice as heavily as spectrum, whereas native listeners weighted it at roughly one-seventh the weight of spectrum. A learner who relies on duration almost exclusively has little room for relying on it further. The absence of a shift may reflect a ceiling on the measure rather than an inability to adjust. These data cannot distinguish whether learners could not adapt or simply did not need to. A second possibility is that learners may not have treated the accent as informative. Korean-accented English resembles their own production errors, and they may have heard it as an obstacle to be overcome rather than as a guide to how the contrast is marked. Consistent with this, the accent did not shift learners' overall response tendency either ($\beta=-0.328$, $p=.167$), whereas it clearly did for native listeners. A further possibility is that learners' spectral representations are simply less precise, so that spectrum offers them less usable information.

The dictation measure also failed to predict adjustment, but this is difficult to interpret. Dictation asks for more than acoustic sensitivity: writing down what one hears also draws on guessing words from context, knowing how to spell them, and holding a sentence in memory. The materials came from a commercial vocabulary application and were not independently validated, and the scores were strongly skewed across only fourteen participants

($M=52.33$, range 21.8–83.8). We therefore treat these analyses as exploratory and the null result as uninformative.

Learners showed no spectral effect in Experiment 1 but a reliable one in Experiment 2. The sentence context of Experiment 2 may have helped. The two experiments differ in several other ways, however: one talker versus four, the presence of a secondary task, the number of trials, the materials, and the order, since Experiment 1 always came first. Learning, fatigue, and order effects are therefore all possible, and we draw no conclusions from comparing the two experiments.

4.5.3. Limitations

The accent manipulation changed many things at once. The Korean-accented carriers differed from the American ones in vowel epenthesis, consonant substitutions, syllable timing, and phrase-level tone, and they were also spoken at a faster articulation rate (Section 4.1.1.3). Rate was measured and controlled, and did not account for the accent effect, but the remaining differences were varied together with the social information the accent conveys. Separating them calls for a matched-guise recording in which only one or two features change while everything else stays fixed.

There was no control accent. As noted above, distinguishing exemplar activation from a general response to unreliable cues requires an accent whose speakers do not themselves rely on duration. This study included none.

The comparison between groups is inconclusive. The relevant interaction did not reach significance, so the study can neither establish nor rule out a group difference of the size hypothesized.

The two groups differed in mean age by roughly ten years. Age was not modelled, so its possible contribution to the group differences reported above cannot be ruled out.

Talker and accent were not fully crossed within participants. Each listener heard two talkers in one accent and two in the other, so that comparisons involving both talker and accent are partly comparisons between different listeners. The counterbalancing list was included as a control predictor in all models, but the design constraint remains.

The spectral continua were not equivalent across talkers. The F1 range spanned 374, 75, 278, and 207 Hz and the F2 range 647, 787, 659, and 639 Hz for Talkers 1 to 4 respectively, so a given step number does not represent the same acoustic distance for every talker. The steps were, however, equally spaced on the Bark scale within each talker's continuum, so step index is at least perceptually uniform within a talker. Because the same continuum was used in both accent conditions for a given talker, this inequality does not enter the within-talker comparisons on which the present conclusions rest; talker was also included as a fixed effect, absorbing differences in overall response tendency across talkers. The unequal ranges therefore enlarge differences between talkers rather than affecting the accent comparisons.

4.5.4. Suggestions for future research

To test how far these results generalize, stimuli could be extended to other non-native accents (e.g., Japanese-accented English) and to other tense-lax pairs (e.g., /u–ʊ/). Sampling could be refined by grouping participants by length of exposure to English-speaking environments. Finally, because the accent moved listeners' decision boundary at least as much as it changed their

cue weighting, designs that separate sensitivity from response bias would be especially informative.

5. Conclusion

This study investigated the effect of a sociolinguistic cue—specifically, a talker's accent—on the listener's speech perception process, particularly on the perceptual cue-weighting of English tense-lax vowels (/i/–ɪ/), from the perspective of Exemplar Theory.

The experimental results showed that the native English speaker group demonstrated perceptual flexibility by significantly increasing their reliance on duration cues when listening to Korean-accented English compared to American-accented English, while their reliance on spectral cues remained unchanged. This is consistent with native listeners recognizing the talker's social background through the accent alone and activating the linguistic information associated with it, as Exemplar Theory predicts, although the present design cannot rule out an adjustment to less reliable spectral information or a shift in response criterion.

In contrast, the advanced Korean learner group did not show significant changes in cue-weighting according to the talker's accent, despite their high proficiency. This is consistent with restricted exemplar activation in second-language listening, but it may equally reflect the fact that learners already relied on duration almost exclusively, leaving little room for further adjustment. The two accounts cannot be distinguished on the present data. The difference between the groups was also not statistically confirmed, so the two patterns are reported side by side rather than as a demonstration that native listeners and learners differ.

These results suggest that speech perception may involve more than the decoding of acoustic signals and may be a complex cognitive process in which a talker's social information dynamically interacts with a listener's accumulated language experience.

Therefore, this study provides implications for the field of English education. Since standardized language test scores do not guarantee native-level perceptual flexibility, pronunciation and listening instruction for advanced learners may need to train them to explicitly recognize and utilize spectral differences, not just vowel length. Furthermore, an educational approach that helps learners categorize pronunciation variants based on talkers' social characteristics and build a rich set of exemplars through consistent exposure to speech input containing diverse accents may also be worth exploring.

In conclusion, this study examined the effect of a talker's implicit social information on the speech perception system and documented the cue-weighting patterns of native speakers and Korean learners under two accent conditions. It is expected that the discussions of this study will be expanded in the future to include a wider variety of proficiency groups and accents from other language backgrounds.

Acknowledgement

The author thanks three anonymous reviewers for their detailed and constructive comments, which substantially improved this paper. Participants received a verbal explanation of the study before

the experiment, and the post-experiment survey stated that their personal information would not be disclosed and that their data would be used for research purposes only. The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

- Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, 137(5), 2823-2833.
- Boersma, P., & Weenink, D. (2025). Praat: Doing phonetics by computer (version 6.4.14) [Computer program]. Retrieved from <http://www.praat.org/>
- Bosker, H. R., & Reinisch, E. (2017). Foreign languages sound fast: Evidence from implicit rate normalization. *Frontiers in Psychology*, 8, 1063.
- Choe, W. K. (2016). Acquisition of prosodic phrasing and edge tones by Korean learners of English. *Phonetics and Speech Sciences*, 8(4), 31-38.
- Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54(1), 99-121.
- Drager, K., & Kirtley, M. J. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In A. M. Babel (Ed.), *Awareness and control in sociolinguistic research* (pp. 1-24). Cambridge, UK: Cambridge University Press.
- Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437-470.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251-279.
- Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d' . *Behavior Research Methods, Instruments, & Computers*, 27(1), 46-51.
- Hay, J. (2018). Sociophonetics: The role of words, the role of context, and the role of words in context. *Topics in Cognitive Science*, 10(4), 696-706.
- Hay, J., & Foulkes, P. (2016). The evolution of medial /t/ over real and remembered time. *Language*, 92(2), 298-330.
- Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458-484.
- Heald, S. L. M., & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience*, 8, 35.
- Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America*, 108(6), 3013-3022.
- Idemaru, K., & Holt, L. L. (2011). Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), 1939-1956.
- Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145-165). San Diego, CA: Academic Press.
- Jun, S. A. (2014). *Prosodic typology II: The phonology of intonation and phrasing*. Oxford, UK: Oxford University Press.
- Kang, H. (2022). A study of English vowel pronunciation by Korean university students with articulatory and auditory training. *Linguistics*, 30(4), 1-18.
- Kim, J. (2016). Perceptual associations between words and speaker age. *Laboratory Phonology*, 7(1), 18.
- Kim, J., & Drager, K. (2017, August). Sociophonetic realizations guide subsequent lexical access. *Proceedings of Interspeech 2017* (pp. 621-625). Stockholm, Sweden.
- Kim, J., & Drager, K. (2018). Rapid influence of word-talker associations on lexical access. *Topics in Cognitive Science*, 10(4), 775-786.
- Koops, C., Gentry, E., & Pantos, A. (2008). The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics*, 14(2), 12.
- Labov, W. (1966). The linguistic variable as a structural unit. *Washington Linguistics Review*, 3, 4-22.
- Lee, H., & Kong, E. J. (2024). Pre-service English teachers' cue-weighting in L2 stop perception. *Phonetics and Speech Sciences*, 16(4), 35-42.
- Lee, J. Y. (2008). Perception of English high vowels: Duration as a cue by Korean speakers of English. *Kansas Working Papers in Linguistics*, 30, 195-204.
- Lee, S., & Rhee, S. C. (2019). The relationship between vowel production and proficiency levels in L2 English produced by Korean EFL learners. *Phonetics and Speech Sciences*, 11(2), 1-13.
- 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.
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18(1), 62-85.
- Nijveld, A., ten Bosch, L., & Ernestus, M. (2022). The use of exemplars differs between native and non-native listening. *Bilingualism: Language and Cognition*, 25(5), 841-855.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776-783.
- Schertz, J., Kang, Y., & Han, S. (2019). Sources of variability in phonetic perception: The joint influence of listener and talker characteristics on perception of the Korean stop contrast. *Laboratory Phonology*, 10(1), 1-32.
- So, H. (2025). *The effect of talker accent on perceptual cue weighting of English tense-lax vowels: A comparison of native English speakers and Korean learners* (Master's thesis). Pusan National University, Busan, Korea.
- So, H., & Kim, J. J. (2024). Age-indexed perceptual cue-shifting of Korean stops in various prosodic positions. *Linguistic Research*, 41(3), 367-390.
- Winn, M. (2023). Make formant continuum [Computer program]. Retrieved from <http://www.mattwinn.com/praat.html#formantContinuum>
- Yang, B. (2010). Formant trajectories of English high tense and lax vowels produced by Korean and American speakers. *Korean Journal of Linguistics*, 35(2), 407-423.
- Yavas, M. (2020). *Applied English phonology* (4th ed.). Hoboken, NJ: Wiley-Blackwell.
- Zhang, Y. (2022). *Acoustic cue-weighting strategies for English*

vowel tenseness and experience-based talker adaptation
(Doctoral dissertation). Pusan National University, Busan,
Korea.

• **Hyunjung So**, Corresponding author
M.A., Dept. of English Language and Literature,
Pusan National University
2 Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan 46241,
Korea
Tel: +82-51-510-1510
Email: minny980@pusan.ac.kr
Areas of interest: Experimental phonetics, Sociophonetics