

Pre-service English teachers' cue-weighting in L2 stop perception*

Hyunjung Lee¹ · Eun Jong Kong^{2,**}

¹Department of English Language Education, Incheon National University, Incheon, Korea

²School of Liberal Arts and Sciences, Korea Aerospace University, Goyang, Korea

Abstract

This study investigates how Korean second language (L2) learners of English perceive the voicing contrast, primarily signaled by voice onset time (VOT) and fundamental frequency (F0), and how attentional load influences their use of these acoustic cues. Testing pre-service English teachers, we observed a general preference for VOT over F0 in identifying the voicing contrast, consistent with the pattern previously reported in native English speakers. At the group level, the interaction between distraction and each acoustic cue revealed that increased cognitive load reduced reliance on both VOT and F0. However, an analysis of individual correlation coefficient showed that VOT and F0 were not equally affected by the distraction. While both cues were negatively influenced, individuals whose reliance on VOT decreased more under cognitive load tended to maintain their reliance on F0. This suggests a compensatory relationship between the two acoustic cues. We discussed that Korean learners' cue-weighting strategies may be influenced by their native language (L1) and classroom learning environments, which limit exposure to naturalistic listening conditions and hinder perceptual flexibility. This study underscores the role of L1 transfer in shaping L2 speech perception and highlights the need for training that incorporates real-world listening challenges.

Keywords: English education, pre-service English teacher, cue-weighting, attentional load, Big5, cloze test

1. Introduction

A phonemic contrast is identified by multiple acoustic cues, and the relative importance of the multiple cues to a speech category varies depending on many factors including cognitive aspects such as attention, environmental noise, and individual characteristics like personality traits and language experiences (e.g., Francis et al., 2000; Yu et al., 2013). In other words, the cue-weighting variations

are explained by who processes the speech sound, and in which environment he/she processes it. Notably, the effect of these variables is even greater when processing unfamiliar or a second language (L2) speech sounds because processing unfamiliar sounds requires more effort compared to processing a mother language (e.g., Mora & Darcy, 2023). Under such high variability, individual L2 learners show varying prioritization for multiple acoustic dimensions of an L2 sound. This study investigates whether and

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2024S1A5A2A01024518).

** ekong@kau.ac.kr, Corresponding author

Received 2 December 2024; Revised 12 December 2024; Accepted 12 December 2024

© Copyright 2024 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.

how L2 learners vary their perceptual cue-weightings modulated by different attentional loads. We examine the individual differences of cue-weightings, focusing on prospective English teachers' perception of the English voicing contrast in stops (i.e., /d/ vs. /t/). This study aims to understand how attention influences L2 learners' cue-weighting of voice onset time (VOT) and fundamental frequency (F0) in the stop perception taking the listener's English proficiency and personality traits into consideration, and discuss how the pattern differs from findings from prior studies with English native speakers.

1.1. Individual Differences in the Perception of the Voicing Contrast in English Stops

English has the two-way voicing distinction of stops, voiced (/b, d, g/) and voiceless (/p, t, k/). While VOT is the primary acoustic cue to signal the voicing distinction of English stops, the VOT cue correlates with F0 at the onset of the following vowel of the stops (Abramson & Lisker, 1985; Francis et al., 2008; Whalen et al., 1993). English voiceless stops usually have longer VOT, which correlates with higher F0 at the following vowel onset, while voiced stops are preceded by relatively lower F0. English-speaking listeners/speakers have different cue-weightings between the two acoustic dimensions varying by many sources. That is, there are individual differences in the relative importance between VOT and F0.

Researchers have sought to identify the sources of individual variation in cue-weightings, examining both environmental factors and individuals' internal cognitive traits. The environmental factors affecting listeners' cue-weighting strategies are primarily related to attentional load during speech processing (e.g., Francis et al., 2000). Consequently, research in this area has investigated speech perception patterns under various distracting conditions, simulating everyday challenges such as noise or multitasking. This line of research typically uses a dual-task paradigm in which listeners respond to auditory stimuli while performing another task. The examples of the additional task include math calculation or letter recall task, which increases cognitive loads during the speech perception task (e.g., Gordon et al., 1993; Kong & Lee, 2018; Mattys & Wiget, 2011).

Gordon et al. (1993) studied how listening conditions influence the use of multiple acoustic cues in speech perception. They focused on the voicing contrast in English stops identified with the primary VOT cue and the secondary F0 cue. By using a dual-task paradigm, Gordon et al. (1993) examined how cognitive distractions manipulated by solving math problems, would affect listeners' reliance on these cues. In their experiment, participants completed a speech identification task under two attentional conditions: one with no distractions (no-distractor) and the other with a math-solving task embedded (distractor). Results showed that in the distractor condition, listeners' reliance on the primary VOT cue decreased, while their use of the secondary F0 cue increased. Gordon et al. (1993) concluded that stronger cues like VOT require more attentional resources, while weaker cues like F0 can be more relied upon when attention is limited. They suggested that under such distracting conditions, secondary cues might become more prominent to compensate the reduced role of primary cues.

Many attempts have also been made to explain the sources of the individual variability in cue-weighting by exploring the effect of cognitive control mechanism (executive functions, e.g., Francis & Nusbaum, 2002), listeners' sensitivity to talkers (e.g., Yu, 2022) or listeners' personality and autistic traits (e.g., Yu, 2010). For

example, Yu (2010) showed that less autistic women adjusted less for phonetic coarticulation compared to men and women with more autistic traits. In his later study, Yu (2022) indicated that variations in how listeners weigh perceptual cues are influenced by listeners' gender and their personal assessment of a speaker. Hutchinson (2022), for example, indicated that L2 learners who demonstrated higher levels of conscientiousness, including being more careful, attentive, and detail-oriented, tended to show less native-like patterns in their perception.

To summarize, the previous studies showed that the individual difference in cue-weighting tended to be systematic in a way that English listeners varied in the reliance of the secondary F0 cue in the perception of stops. The findings suggest that speech perception is not static but highly adaptable, depending on both environmental and individual factors. The insights from this research emphasize the importance of understanding cognitive mechanisms and personal characteristics in speech perception, which could potentially inform individual approaches in language learning and auditory processing interventions.

1.2. Individual Differences in Cue-Weighting During L2 Speech Perception

A subsequent question arises regarding L2 speech perception which is inherently more cognitively demanding than L1 speech perception due to its unfamiliarity. Extensive L2 research has explored the source of the variability ranging from learner, context and linguistic variables, aiming to identify more efficient learners and the environmental conditions that enhance L2 speech perception (e.g., Munro & Bohn, 2007). A recent study, Mora & Darcy (2023), showed the relationship between cognitive abilities and individual difference in L2 speech perception. They examined how attention control abilities measured by a separate task affected phonological processing of adult L2 learners, focusing on Spanish speakers learning English and English speakers learning Spanish. They showed that learners with more efficient attention-switching skills could discriminate L2 vowels more quickly, and the attention control was linked to better production of L2 vowels.

In addition to the learners' cognitive functions, L2 learners are also affected by the environmental variables, and they tend to be more native-like perception pattern under less distracting contexts (e.g., Asano 2018; Lee, 2014). For example, Asano (2018) showed that under higher task demands with extended inter-stimuli-intervals (ISIs), the non-native listeners exhibited a decline in the perception, indicating that L2 learners were strongly affected by increased memory load and attentional challenges. Asano (2018) concluded that L2 learners' ability to use acoustic dimensions diminishes when distracting, suggesting the difficulties that advanced learners often face in real-life listening environments with numerous distractions.

Closely relevant to the current study, using the dual-task paradigm in Gordon et al. (1993), Lee & Kong (2023) examined the attentional modulation on VOT and F0 cues in the perception of English voicing distinction by Korean-speaking learners of English. They examined how the relative importance of VOT and F0 was modulated by two different distracting conditions for the L2 learners. The results showed that while Korean learners' sensitivity to the primary VOT cue decreased under the distracting condition, the reliance on the secondary F0 cue did not compensate the reduced role of VOT and the sensitivity to F0 was also reduced when distracted. The absence of compensating effect under distracting

condition differed from English-speaking listeners' cue-weighting pattern observed in Gordon et al. (1993). Lee & Kong (2023) suggested two possible explanations for the inconsistency. One is the reflection of L1 Korean's cue-weighting pattern. Korean has three-way laryngeal contrast of stops which is distinguished both by VOT and F0 (e.g., Lee et al., 2020). The amount of attention to each of the two acoustic cues is almost equal in processing the stop contrast, that is, F0 is not secondary to compensate the reduced VOT. The other explanation is the classroom learning setting where Korean learners of English are normally exposed with English. That is, such non-naturalistic learning context has made Korean learners adverse to process L2 acoustic signals under distracting contexts.

Given that English voicing distinctions typically have a well-established prioritization between the two acoustic dimensions, VOT and F0, this discrepancy prompts further scrutiny of the learner group in Lee & Kong (2023) particularly with respect to potential variables such as varying English proficiency and non-controlled personality traits, which may have influenced the L2 learner group's reduced sensitivity to the secondary F0 cue. Therefore, the current study examines the effect of attention on the cue-weighting of L2 English voicing contrast (/ta/ vs. /da/) for Korean-speaking prospective English teachers who majored in English Education.

This study specifically tested whether and how prospective English teachers, as a homogeneous learner group, exhibit differences in their reliance on VOT and F0 under varying distracting conditions, compared to L1 English listeners as examined by Gordon et al. (1993). By limiting the observation to future English teachers, we ensure a relatively homogeneous group in terms of educational exposure, which helps control potential factors such as proficiency, psychological motivation and personality traits. This allows us to focus more closely on the target variable—attentional load. The findings offer implications for understanding speech perception differences between classroom-taught learners and native speakers exposed to naturalistic environments.

This is a preliminary study, which is a part of a larger investigation. We will later compare perceptual patterns under attentional conditions between two distinct groups with differing traits in English proficiency, personality, and autistic characteristics, including communicative skills. For instance, we aim to compare the perceptual patterns of English education majors with those of science or engineering majors to observe how these groups respond to attentional load when using multiple acoustic cues. The results of this study could inform more tailored pedagogical approaches that account for the specific cognitive and perceptual challenges faced by different learner groups under varying attentional demands.

The current study specifically questions how Korean pre-service English teachers adjust their reliance on VOT and F0 when perceiving English stop voicing contrasts under distracting conditions. We discuss the results in reference with the results from native English speakers reported in Gordon et al. (1993). We subsequently test how the patterns related to their individual traits measured by Big5.

2. Methods

2.1. Participants

Twenty-six Korean-speaking university students (four males)

participated in the perception experiment for a nominal fee. The participants' ages ranged from 19 to 26, with a mean age of 21.6 ($SD=2.06$). To examine the correlation of individual L2 learners' personality traits and attention with their speech perception patterns, this study limited participants to prospective English teachers at middle and high schools. All of the participants majored in English Education, and online recruitments were made among the English majors at Incheon National University and at Korea National University of Education. While speakers of Seoul/Gyeonggi, Chungcheong, and Jeolla dialects were included in the study, speakers of the Gyeongsang dialect were excluded to reduce the potential effect of tonal dialects. None of the participants reported any language disorders.

We conducted cloze test (Chung & Ahn, 2019) and Big5 to test the homogeneity of the present learner group. cloze test assessed participants' English proficiency, asking them to fill in the blanks made to assess reading comprehension, grammar, and vocabulary of English. Big5 measures five major personality traits: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (e.g., Goldberg, 1993, 2001; John & Srivastava, 1999). In the Big5 survey, respondents rated their agreement with 44 statements on a Likert scale, such as "I am the life of the party." The scores for each personality trait were calculated by averaging the responses for the statements associated with that specific trait, giving a comprehensive score for each personality dimension.

2.2. Auditory Stimuli

The auditory stimuli were adapted from Lee & Kong (2023). The stimuli were consonant-vowel (CV) syllables, forming a continuum from /ta/ to /da/, created using words produced by an English-speaking adult male speaker from Wisconsin, USA. A /da/ token was selected, and VOT and F0 were systematically varied. Seven log-scale VOT steps (9, 13, 19, 28, 40, 58, to 100 ms) were created by modifying the burst release/aspiration, while five F0 values (98, 106, 114, 122, to 130 Hz) were applied to the vowel portion. This resulted in 35 stimuli (7 VOT×5 F0 combinations). CV syllables were used to avoid lexical confounds and ensure consistency across tasks.

2.3. Task and Procedure

The speech perception experiment was developed and conducted using an online platform, Gorilla (gorilla.sc, Anwyl-Irvine et al., 2019). Each participant completed a set of tasks, including a language background questionnaire, the Big5 Personality Test, the cloze test, and a 2-alternative forced-choice (2AFC) dual paradigm task (with and without a math calculation prior to category decision), by clicking the experiment URL provided by Gorilla. A progress bar on the computer screen helped participants track their progress throughout the online experiment.

Specifically, the dual paradigm 2AFC task followed the experimental design of Gordon et al. (1993) and Kong & Lee (2018), focusing on attentional manipulation. Participants were asked to decide between the English stop voicing contrasts (/t/ vs. /d/) under two different attentional conditions: with a distractor (Distractor) and without a distractor (No-Distractor). In the No-Distractor condition, participants responded solely to the auditory stimuli. In the Distractor condition, participants performed a simple math calculation (e.g., "Given the three numbers 10, 20, and 30, is 20 minus 10 the same or different from 30 minus 20?")

while also completing a speech perception task (/d/ or /t/?). The auditory stimulus syllable was presented before the math question, and participants categorized the stop sound after completing the calculation. Each participant responded to 280 trials (35 auditory stimuli×2 distractor conditions×4 repetitions) on the Gorilla online platform.

Cloze test: The L2 learners' responses to 40 questions were manually graded by the authors based on whether the responses were identical to the answer keys (2 points), comparable to them (1 point), or irrelevant (0 points) (Jun, 2018, among others). The graders independently scored the responses and later consulted with each other to resolve any discrepancies (about 10 cases). The L2 learners' cloze test scores ranged from 21 to 65 out of a possible 80 points (mean=38.54, $SD=13.18$). Each learner's score was used to represent their English proficiency in the statistical model.

Big5 survey: Figure 1 shows the partial correlation coefficients and p -values for every pair of Big5 sub-components, with the rest of the components controlled. At the 0.05 significance level, no significant correlations were found between the five sub-components, except for the relationship between Openness and Agreeableness. The prospective English teachers' Openness and Agreeableness scores were positively correlated, controlling for Conscientiousness, Extraversion, and Neuroticism: $r(26)=0.572, p<0.005$. Given that the sub-components were relatively independent and represented distinct personal traits, we selected Agreeableness and Extraversion for this analysis to represent the personal traits of prospective English teachers. Participants' scores for Extraversion ranged from 1.625 to 4.875 (mean=3.375, $SD=0.674$), and for Agreeableness from 2 to 4.778 (mean=3.765, $SD=0.657$).

2.4. Analysis

2AFC with and without a distractor: A mixed-effects logistic regression model was constructed to predict stop category responses (/d/ and /t/) based on acoustic parameters in both distracted and non-distracted listening conditions (using the *lme4* and *lmerTest* packages in R via RStudio (Racine, 2012, Version 2024.04.1+748): Bates et al., 2015; Kuznetsova et al., 2017). Fixed effects included VOT and F0 in normalized units [using `scale()` in R] and the presence of distracting math tasks (DistCondition: Distractor vs. No-Distractor). Random effects included by-subject intercepts and slopes for VOT and F0 at the speaker level and at the speaker-by-distractor level. The fixed effect coefficients of VOT and F0 in the Distractor and No-Distractor conditions indicate group-averaged weights for VOT and F0 with and without the distractor. The random effect coefficients at the speaker-by-distractor level estimate individual speakers' coefficients as deviations from the group averages in both conditions.

To investigate the effect of the distractor on individual prospective English teachers' stop perception, we ran a partial correlation test between VOT and F0 coefficients from the mixed-effects model, controlling for the listeners' personal traits (Agreeableness and Extraversion) and English proficiency (the cloze test scores). The `ppcor` package was used for the analysis {Kim, 2015: `ppcor(data, method=c("pearson"))`}.

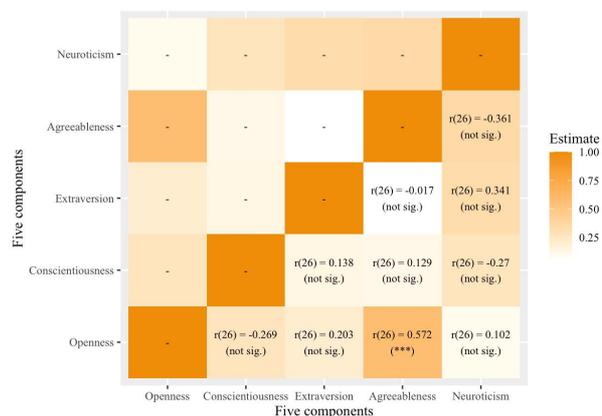


Figure 1. The partial correlation coefficients and p -values of each Big5 sub-component pair with the rest of the components controlled.

3. Results

3.1. Relative Cue-Weighting Between VOT and F0

Table 1 presents the beta coefficients of the fixed-effect variables estimated from the mixed-effects logistic regression model using a treatment coding scheme. Both the two acoustic variables, VOT and F0, were statistically meaningful factors in L2 learners' identification of /t/ over /d/ without a distractor ($\beta_{VOT}=3.79, SE=.41, p<0.0001$; $\beta_{F0}=1.65, SE=0.13, p<0.0001$), and the VOT coefficient was greater than F0 coefficient suggesting that the Korean L2 listeners, as a group, put more perceptual weight on VOT than F0 as English native listeners did (e.g., Abramson & Lisker, 1985). Importantly, both acoustic dimensions were significantly affected by the distractor such that the interaction terms of VOT and F0 with Distractor were significant ($\beta_{VOT \times Distractor}=-1.29, SE=0.33, p<0.0005$; $\beta_{F0 \times Distractor}=-0.36, SE=0.11, p<0.005$). The coefficients of the interaction terms suggest that not only VOT but F0 were used less when distracted than not. Overall, the L2 learner group weighted VOT over F0 in perceiving the L2 English voiced- voiceless contrast, and the distractor affected them in using both primary and secondary cues.

Figure 2 plots the logistic curves based on the beta coefficients of Table 1. In both Distractor and No-Distractor conditions, the VOT curves (dashed line) are steeper than the F0 curves (solid line), showing that the listeners were more sensitive to VOT than to F0 in identifying /t/ over /d/. In addition, both VOT and F0 curves were less steep in the Distractor condition (black lines) than those in the No-Distractor condition.

The left panel of Figure 3 plot the individual listeners' VOT coefficients as a function of their F0 coefficients, illustrating how the individual coefficients changed from No-Distractor condition (circles) to Distractor condition (squares). With or without distractor, most VOT coefficients were greater than F0 coefficients, located under the diagonal line. This suggests that most L2 learners used VOT primarily for the English /d-/t/ distinction. There were only four datapoints above the line who relied on F0 primarily for the contrast. In terms of the effect of distractor, the arrows connecting each individual learner's coefficients of No-Distractor condition with those of Distractor condition tend to head toward the lower and left corner of the panel. This indicates that VOT and F0 were used less when distracted than not, in general. There were a handful of listeners whose arrows head for different directions deviant from

this general tendency. That is, a single listener used F0 more with its VOT coefficient decreased when distracted than not (a red square, an arrow pointing to an upper left corner of the panel). Another six listeners used VOT more with their F0 coefficients decreased when distracted than not (six blue squares, arrows pointing to a lower right corner of the panel).

While the L2 learners did not show a reconciliation of a secondary cue under distractor except one participant (red square) because both cues VOT and F0 were reduced, they were not equally affected by the distractor. In Figure 3, the right panel repeats the coefficient distributions of the two distractor conditions, highlighting six datapoints in green whose F0 coefficients were greater than VOT coefficients when distracted, although their F0 coefficients were not greater than VOT coefficients without distractor. In other words, distracting listening condition affected these six listeners using F0 as a primary cue even though they primarily relied on VOT without distractor. They may also support Gordon et al. (1993) in the L2 perception in a way that the role of a secondary cue was boosted in a relative sense under distractor.

Table 1. Mixed-effects logistic regression summary (reference level: No-Distractor)

	Estimate	Std. error	p-value
(Intercept)	0.44	0.2	<.05
VOT	3.79	0.41	<.0001
F0	1.65	0.13	<.0001
Distractor	-0.47	0.16	<.005
VOT×Distractor	-1.29	0.33	<.0005
F0×Distractor	-0.36	0.11	<.005

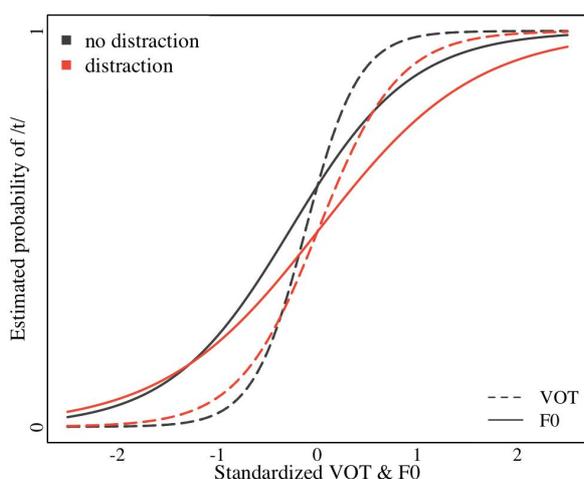


Figure 2. Logistic curves of the fixed effect variables (VOT: dashed and F0: solid) based on the mixed-effects regression model coefficients. The curves are color-coded by the listening conditions: Distractor in black and No-Distractor in red.

3.2. Distractor on VOT and F0, Big5, and English Proficiency

Finally, the partial correlation test was conducted to estimate the correlation coefficient between acoustic variables given personal traits and L2 proficiency. The two test variables of the test were individual listeners' VOT and F0 coefficient differences between Distractor and No-Distractor conditions. Two listener-related variables were set as control variables of the test: (1) Individual learners' mean scores of Big5 survey, and (2) their mean scores of

the cloze test. The test result showed that VOT coefficient differences was significantly correlated with F0 coefficient differences given their personal traits and English proficiency: $r(26)=-0.431$, $p<0.05$. The negative correlation coefficients suggest that greater VOT differences were associated with smaller F0 differences (see Figure 4). The correlation test result supports that the L2 learners who lost their attention to VOT tended to lose less attention to F0 when distracted. That is, the role of a secondary cue, F0 for the voicing contrast in English, enhanced or maintained to reconcile the damaged role of a primary cue, VOT.

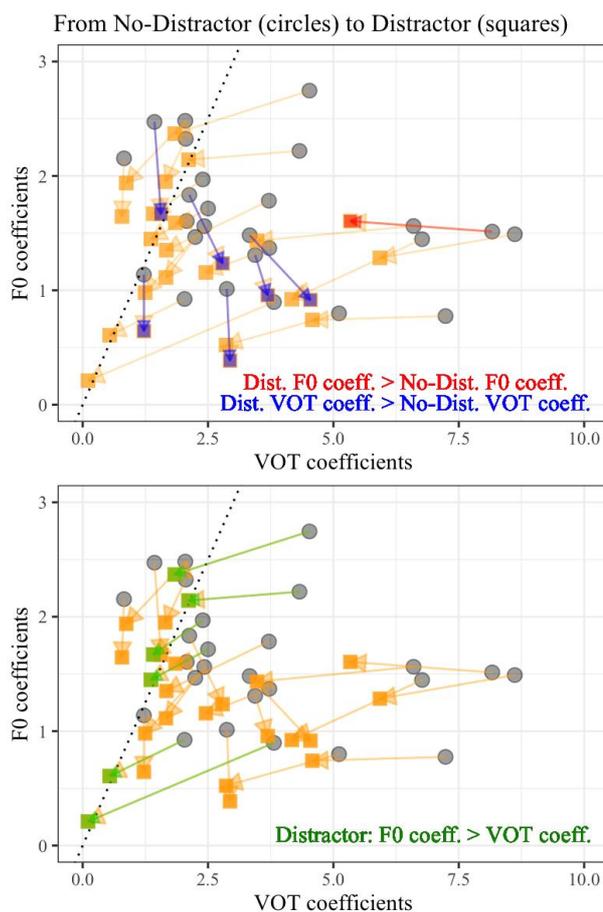


Figure 3. Individual L2 learners' F0 coefficients as a function of their VOT coefficients in the No-Distractor (grey circles) and the Distractor conditions (squares). Arrows connect the individuals' coefficients of the two listening conditions: from No-Distractor to Distractor.

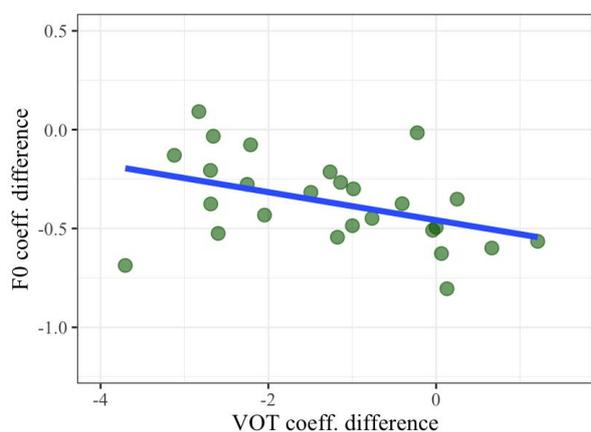


Figure 4. Individual L2 learners' F0 coefficient difference (Distractor – No-Distractor) as a function of their VOT coefficient difference.

4. Discussion

This study examined how Korean L2 learners of English perceived the English /t-/d/ voicing contrast mainly signaled by two acoustic cues, VOT and F0, and how distractors affected their use of the multiple acoustic cues in the perception of a non-native language. By testing prospective English teachers majoring in English education, we observed a homogeneous group of L2 learners with relatively similar English proficiency and personality traits, expecting a native-like perceptual flexibility reported in Gordon et al. (1993).

In the group-averaged pattern of the results, we found that in the absence of a distractor, both VOT and F0 were statistically significant in the learners' identification of /t/ over /d/, with VOT having a greater influence than F0. This suggests that Korean L2 learners, like native English listeners, prioritize VOT when distinguishing for auditory identification. However, both VOT and F0 were affected by distractors, with a significant reduction in the use of both cues. The significant interaction terms associated with the distractor indicate that learners relied less on both cues when distracted, though VOT remained the primary cue. Regarding individual variations, while most learners followed the general trend of relying more on VOT, a few exhibited different patterns. Specifically, six learners used VOT more when distracted, compensating with a decreased reliance on F0. A subset of the learners showed an enhanced use of F0 as a primary cue under distraction, consistent with Gordon et al. (1993). Finally, a partial correlation analysis showed a significantly negative correlation between changes of VOT and those of F0 coefficients across conditions. That is, learners who showed a greater drop in their use of VOT under distraction tended to sustain their reliance on F0, reinforcing the idea that F0 serves as a compensatory cue when attention to VOT is diminished. The main findings of the current study can be discussed within the context of previous research (Gordon et al., 1993; Lee & Kong, 2023) and the specific characteristics of the learner group tested.

The present results align with earlier findings (Lee & Kong, 2023) that L2 learners, like native English speakers, primarily rely on VOT as the dominant acoustic cue when distinguishing between voiced and voiceless stops. This pattern is consistent with the

widely recognized importance of VOT in marking the voicing contrast in English (e.g., Gordon et al., 1993; Lisker & Abramson, 1964). Even under distractor conditions, VOT remained the stronger cue, although its effectiveness was diminished by the increased cognitive load. This confirms that VOT requires more attentional resources, supported by Gordon et al.'s (1993) findings with native listeners. However, while F0 is generally considered a secondary cue for the voicing distinction, the present results revealed that prospective English teachers of the current study did not compensate for the reduced role of VOT by relying more on F0 when distracted, similar with the learner group with broader background observed in Lee & Kong (2023). This is different from native English speakers, who increased their reliance on F0 under similar conditions as VOT became less accessible (Gordon et al., 1993). The lack of compensatory use of F0 among the L2 learners may reflect a less flexible cue-weighting usage compared to native English speakers. Instead, it appears that both VOT and F0 are affected by the attentional load in Korean-speaking L2 learners of English, suggesting that these cues are processed together, rather than one compensating for the other under cognitive strain. However, it is important to note the negative correlation between changes in VOT and F0 coefficients. In other words, individuals who were more affected by distraction in their use of VOT tended to rely more on F0. This suggests that while both cues were negatively impacted by increased cognitive load, they were not equally affected.

In the results, the prospective teacher group of the current study did not particularly show native-like perceptual modification associated with the distracting condition as seen in Gordon et al. (1993). For the non-native-like perceptual performance, the current study suggests the influence of L1 transfer and the classroom learning environment on the lack of flexibility in using the secondary F0 cue under distraction, consistent with Lee & Kong (2023). First, Korean L2 learners may bring their native language's cue-weighting patterns into their English speech perception. In Korean, both VOT and F0 are crucial cues for distinguishing stop consonants in a three-way laryngeal contrast, unlike in English, where VOT is the primary cue and F0 plays a secondary role. This could explain why the learners in the current study did not show compensatory reliance on F0 when VOT processing became more difficult under distraction. Instead, both cues may have been perceived as equally important, leading to a simultaneous reduction in their use when cognitive demands were high. The lack of flexibility in shifting to F0 under distraction could reflect a deeper phonological reliance on both cues due to L1 influence, where the learners are accustomed to treating VOT and F0 as co-primary rather than hierarchical cues.

Second, the present findings suggest that classroom learning context may not provide appropriate L2 training in challenging listening conditions, not improving perceptual flexibility. Korean learners of English are typically educated in classroom settings where language input is less naturalistic and controlled. This could limit their exposure to the dynamic, real-world listening environments that native English speakers experience, reducing their ability to adapt cue-weighting strategies when faced with attentional distractions or competing acoustic cues. This could explain why the prospective teachers in the current study showed reduced perceptual flexibility and struggled to compensate with F0 when VOT processing was compromised under distraction. The findings highlight the importance of considering both L1

phonological systems and learning contexts when understanding L2 learners' speech perception. The results suggest that prospective English teachers may need more exposure to naturalistic listening environments and training that promotes flexibility in cue-weighting under cognitive strain. Future pedagogical approaches could incorporate dual-task listening exercises or multitasking tasks to help L2 learners better manage attentional loads and improve their adaptability in real-world communication settings. This would better prepare learners for the complex auditory demands they will face outside the classroom.

Inconsistent with Lee & Kong (2023), however, we observed more variabilities in individual learners' cue-weighting strategies in which F0 compensates the reduced VOT role under distractions. While Lee & Kong (2023) reported four listeners out of 28 whose F0 compensated the reduced VOT, the present study showed six out of 26. Although the majority of participants demonstrated a clear preference for VOT over F0, those six learners displayed different patterns, such as increased reliance on F0 when VOT was compromised. This variability may stem from the homogeneous characteristics of the learner group. That is, English proficiency or personality traits of the prospective English teacher group might explain why more learners, compared to Lee & Kong (2023), deviated from the general trend of relying primarily on VOT. The more uniform educational background of the prospective English teachers may contribute to their increased ability to shift to F0 when VOT became less reliable under distraction. There are limitations such as having small number of sample and not covering wide range of population with various background. While this study focused on a relatively homogeneous group of prospective English teachers, future research should examine how other groups of learners, such as science and engineering majors, process similar phonemic contrasts under varying attentional loads. By comparing different learner groups, we can better understand the role of individual traits in L2 speech perception and develop more personalized approaches to language teaching.

5. Conclusion

In conclusion, this preliminary study highlights the importance of attentional load in modulating cue-weighting strategies in L2 learners and points to the need for further research on how cognitive factors and individual differences influence speech perception. These findings contribute to a growing body of literature that emphasizes the dynamic nature of speech processing and the significance of both environmental and personal factors in shaping learners' perceptual strategies.

References

- Abramson, A. S., & Lisker, L. (1985). Relative power of cues: F0 shift versus voice timing. In: V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp.25-33). Cambridge, MA: Academic Press.
- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. Z., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods*, 52, 388-407.
- Asano, Y. (2018). Discriminating non-native segmental length contrasts under increased task demands. *Language and Speech*, 61(3), 409-429.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using the lme4. arXiv. <https://doi.org/10.48550/arXiv.1406.5823>
- Chung, E. S., & Ahn, S. (2019). Examining cloze tests as a measure of linguistic complexity in L2 writing. *Language Research*, 55(3), 627-649.
- Francis, A., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 349-366.
- Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics*, 62(8), 1668-1680.
- Francis, A. L., Kaganovich, N., & Driscoll-Huber, C. (2008). Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *The Journal of the Acoustical Society of America*, 124(2), 1234-1251.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26-34.
- Goldberg, L. R. (2001). An alternative "description of personality": The big-five factor structure (1st ed.). In S. Hyman (Ed.), *Personality and personality disorders* (pp. 34-47). London, UK: Routledge.
- Gordon, P. C., Eberhardt, J. L., & Rueckl, J. G. (1993). Attentional modulation of the phonetic significance of acoustic cues. *Cognitive Psychology*, 25(1), 1-42.
- Hutchinson, A. E. (2022). Individual variability and the effect of personality on non-native speech shadowing. *JASA Express Letters*, 2(6), 065203.
- John, O. P., & Srivastava, S. (1999). The big-five trait taxonomy: History, measurement, and theoretical perspectives (2nd ed.). In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102-138). New York, NY: Guilford Press.
- Jun, H. (2018). Cloze test for proficient in Korean as a second/foreign language (Master's thesis), Yonsei University, Seoul, Korea.
- Kim, S. (2015). Ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, 22(6), 665-674.
- Kong, E. J., & Lee, H. (2018). Attentional modulation and individual differences in explaining the changing role of fundamental frequency in Korean laryngeal stop perception. *Language and Speech*, 61(3), 384-408.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 26.
- Lee, H. (2014). Effects of attention on the perception of L2 phonetic contrast. *Journal of the Korean Society of Speech Sciences*, 6(4), 47-52.
- Lee, H., Holliday, J. J., & Kong, E. J. (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, 14(7), e12374.
- Lee, H., & Kong, E. J. (2023). Attentional modulation on multiple acoustic cues in phonological processing of L2

- sounds. *Phonetics and Speech Sciences*, 15(4), 11-16.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *WORD*, 20(3), 384-422.
- Mattys, S. L., & Wiget, L. (2011). Effect of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145-160.
- Mora, J. C., & Darcy, I. (2023) Individual differences in attention control and the processing of phonological contrasts in a second language. *Phonetica*, 80(3-4), 153-184.
- Munro, M. J., & Bohn, O. S. (2007). The study of second language speech. In O. S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 3-11). Amsterdam, Netherlands: John Benjamins B.V.
- Racine, J. S. (2012). RStudio: A platform-independent IDE for R and Sweave, *Journal of Applied Econometrics*, 27(1), 167-172.
- Whalen, D. H., Abramson, A. S., Lisker, L., & Mody, M. (1993). F0 gives voicing information even with unambiguous voice onset times. *The Journal of the Acoustical Society of America*, 93(4), 2152-2159.
- Yu, A. C. L. (2010). Perceptual compensation is correlated with individuals' 'autistic' traits: Implications for models of sound change. *PLOS ONE*, 5(8), e11950.
- Yu, A., Abrego-Collier, C., & Sonderegger, M. (2013) Phonetic imitation from an individual-difference perspective: Subjective attitude, personality, and "autistic" traits. *PLOS ONE*, 8(9), e74746.
- Yu, A. C. L. (2022). Perceptual cue weighting is influenced by the listener's gender and subjective evaluations of the speaker: The case of English stop voicing. *Frontiers in Psychology*, 13, 840291.

• **Hyunjung Lee**

Associate Professor, Dept. of English Language Education,
 Incheon National University
 12 Gaebeol-ro, Songdo-dong, Yeonsu-gu, Incheon 21999, Korea
 Tel: +82-32-835-8992
 Email: hjlee123@inu.ac.kr
 Areas of interest: Speech production & perception, Laboratory
 phonology, Language acquisition

• **Eun Jong Kong**, Corresponding author

Professor, School of Liberal Arts and Sciences,
 Korea Aerospace University
 76 Hanggongdaechang-ro, Deogyang-gu, Goyang 10540, Korea
 Tel: +82-2-300-0338
 Email: ekong@kau.ac.kr
 Areas of interest: Speech production & perception, Laboratory
 phonology, Language acquisition