ENGLISH LISTENERS' PERCEPTION OF KOREAN LARYNGEAL CONTRASTS AMONG WORD-INITIAL STOPS

Eunjin Lee, University of Utah

The Perceptual Assimilation Model (PAM) predicts which novel phonological contrasts will be more challenging for adult learners by relating listeners' perceptual assimilation patterns to discriminability. In this study, we test the predictions of the PAM with English speakers and Korean three-way stop contrasts in word-initial position. Perceptual assimilation task results from 41 participants showed that Korean fortis stops were assimilated to English voiced stops and Korean lenis and aspirated stops to English voiceless stops, though the Korean aspirated velar stop was a better exemplar than was the lenis velar stop. The AX discrimination task results bore out the predictions of the PAM. The lenis–aspirated contrast was more difficult to discriminate than the lenis–fortis and aspirated–fortis contrasts. The results confirm the predictions of the PAM and suggest that the lenis–aspirated contrast may present difficulty for English learners of Korean.

INTRODUCTION

Research on cross-language speech perception of L2 speakers suggests that adults often have difficulties distinguishing L2 phonological contrasts (e.g., Best et al., 2001) due to language-specific experience that shapes sensitivities to phonetic distinctions (Kuhl et al., 1992; Werker & Tees, 1984). The Perceptual Assimilation Model (PAM: Best, 1995; see Best & Tyler, 2007 for PAM-L2) predicts relative ease and difficulty of discriminating novel L2 contrasts based on how L2 phones are perceptually assimilated to L1 phonemes. Thus, the present study examines the novice L1 English listeners' assimilation and discrimination of Korean stop contrasts to determine which Korean stop contrasts will be most difficult for L1 English listeners with no prior experience and to evaluate the predictions of the PAM.

Literature Review

Due to attunement to L1 phonological contrasts, language users become highly efficient at distinguishing phonological contrasts within their L1 (Werker & Tees, 1984). As studies of adult listeners' speech perception suggested, this results in difficulties in perceiving novel L2 phonemes when their L1 phonologies differ. Models of L2 perception, such as the Speech Learning Model (Flege, 1995; see Flege & Bohn, 2021 for SLM-r) or Perceptual Assimilation Model (Best, 1995), account for the perception of novel speech sounds by considering phonetic similarity. As the interest of the present study is the perception of non-native contrasts by inexperienced listeners, our theoretical discussion focuses on the PAM. According to the PAM, the relative ease and difficulty learners will experience with novel phonological contrasts are

Lee, E. (2024). English Listeners' Perception of Korean Laryngeal Contrasts Among Word-Initial Stops. In D. J. Olson, J. L. Sturm, O. Dmitrieva, & J. M. Levis (eds), *Proceedings of 14th Pronunciation in Second Language Learning and Teaching Conference*, (pp. 1-12). Purdue University, September 2023. https://doi.org/10.31274/psllt.17278

based on perceptual assimilation patterns. For example, when two L2 phones are assimilated to different L1 phonemes (a Two-Category assimilation pattern), learners are expected to exhibit excellent discrimination of the contrast. By contrast, when two L2 phones are mapped to a single L1 phoneme and with equal category goodness (a Single-Category assimilation pattern), the novel contrast is expected to be more difficult to discriminate. However, when one of two L2 phones, which are mapped to a single L1 phoneme, is a better exemplar of the L1 phoneme than the other one (a Category-Goodness assimilation pattern), discrimination is expected to be relatively good. Learners are also expected to show very good discrimination when one L2 phone is categorized, but the other is uncategorized. But discrimination is expected to vary from poor to excellent when both L2 phones are uncategorized (Best, 1995).

Korean and English have different word-initial stop categories (see Table 1). English stops belong to two categories, voiceless and voiced. English word-initial voiced stops are produced with short-lag voice onset time (VOT) or prevoicing, and voiceless stops are produced with longer VOT in a wider range (Chodroff & Wilson, 2017). VOT serves as a primary acoustic cue, and fundamental frequency (F0) serves as a secondary cue to separate stop categories in English (see Dmitrieva et al., 2015). Korean stops are categorized as lenis, aspirated, and fortis. In Korean, both VOT and F0 mainly serve as primary cues; Korean aspirated stops are produced with the longest VOT and highest F0, lenis stops with intermediate VOT and lowest F0, and fortis stops with the shortest VOT and intermediate F0 (Lee & Jongman, 2012). Different levels of sensitivity to different acoustic cues might contribute to difficulties in distinguishing Korean stops for English listeners with no prior experience in Korean.

Cross-language perception studies provided some evidence of the difficulty that English listeners experience with Korean stop contrasts. Schmidt (2007) investigated English listeners' perception of Korean consonants using an identification and rating task. Participants listened to a Korean CV syllable and typed the English letter(s) that most closely resembled the initial consonant of the auditory stimulus. Schmidt found that the participants labeled Korean lenis and aspirated stops as English voiceless stops and fortis stops as voiced stops. Similarly, Nagle et al. (2023) found that L1 English listeners mapped Korean lenis and aspirated stops onto English voiceless stops, while mapping of fortis stops was variable to English voiced and voiceless stops with highly variable ratings.

Table 1

Korean and English Stop Phonemes in the Word–initial Position

		Bilabial	Alveolar	Velar
Korean	Lenis	/p/ (ㅂ)	/t/ (ㄷ)	/k/ (ㄱ)
	Aspirated	/pʰ/ (ㅍ)	/tʰ/ (ㅌ)	/kʰ/ (ㅋ)
	Fortis	/p*/ (ㅃ)	/t*/ (ㄸ)	/k*/ (ㄲ)
English	Voiceless	/p/	/t/	/k/
	Voiced	/b/	/g/	/g/

Note. The corresponding Korean orthographic form of a phoneme is in parentheses.

Based on Schmidt's findings, Kwon (2014) hypothesized that Korean lenis and aspirated stops would be the most difficult for L1 English listeners to discriminate, as the lenis and aspirated stops are perceptually assimilated to English voiceless stops. Unexpectedly, the lenis-fortis contrast was least accurately discriminated in Kwon's AX discrimination task, though discriminability of the lenis-fortis contrast differed by place of articulation. The author stated that this finding contradicts the discrimination predictions based on Schmidt's identification findings. Seo et al. (2022) examined discrimination of Korean lenis-aspirated contrasts by Korean heritage speakers and compared their performance with those of English-speaking learners of Korean. It was demonstrated that the learners had difficulty in discriminating the lenis-aspirated contrasts in the AX discrimination task. Similarly, Lee-Ellis (2012) found that Korean heritage speakers and English learners of Korean less accurately discriminated the fricative lenis-fortis contrast than Korean learners of English.

The literature reviewed here has produced mixed findings with respect to the perception of Korean stops and the discrimination difficulty of English listeners. In addition, we are unaware of any existing studies that have directly examined the predictions of the PAM with Korean laryngeal contrasts involving word-initial stops and L1 English listeners. Thus, the aim of this study is (1) to provide new data regarding the perception of Korean word-initial stop contrasts by English speakers and (2) to determine whether English listeners' cross-language perceptual assimilation patterns predict their discrimination of Korean laryngeal contrasts among word-initial stops per the PAM. We hypothesized that the PAM's predictions would be confirmed.

METHODS

Participants

Participants included forty-one individuals who identified English as a native and/or childhood language and had no prior experience with Korean. They were recruited at a large public university in the United States and earned course credit for their participation. Five participants additionally identified Spanish as a native and/or childhood language, one French, one Japanese, one Catalan and one did not report a childhood language. Thirty-five participants reported knowledge of one or more additional languages. More language information is available at the Open Science Forum (https://osf.io/7yfv6).

Materials

The auditory stimuli involved CV syllables containing the nine Korean stops followed by the vowel /a/, which is a familiar phoneme in both Korean and English. A female speaker of Seoul Korean recorded the stimuli in a sound–attenuated booth using a recorder. Each token was extracted using Praat (Boersma, 2011), and the intensity was adjusted to an average of 65 dB using a scale intensity script (Vicenik, 2009). For an AX discrimination task, the CV syllables formed nine contrasts (see Table 2). Twelve keywords, shown in Figure 1 (adapted from Hayes-Harb & Barrios, 2021), were used for an English keyword familiarization phase and perceptual assimilation task. The twelve initial consonants in the keywords were included as they were the phonemes that English listeners identified as a match for Korean stop consonants in Schmidt

(2007). A female English speaker recorded forty-eight monosyllabic English words (4 words beginning with each of the initial stops in the keywords) for the keyword familiarization phase.

Table 2

Contrast Types of Korean Stops by Place of Articulation

	Bilabial	Alveolar	Velar
Lenis–aspirated	/pa/–/p^{h}a/	/ta/–/ t^{h}a/	/ka/–/k^{h}a/
Lenis–fortis	/pa/–/p*a/	/ta/–/t*a/	/ka/–/k*a/
Aspirated–fortis	/p^{h}a/–/p*a/	/t^{h}a/–/t*a/	/k^{h}a/–/k*a/

Procedure

The research was approved by the University of Utah Institutional Review Board. It consisted of the consent, AX discrimination task, English keyword familiarization phase, perceptual assimilation task, and language background questionnaire administered via PsychoPy/Pavlovia (Peirce et al., 2019) and Qualtrics (Qualtrics, Provo, UT). Participants were asked to complete the study, which lasted approximately 45 minutes, in a quiet place with their headphones on.

First, participants completed a consent form. Next, participants completed the AX task. Eight practice trials involving English words (adapted from Hayes-Harb & Barrios, 2021) were presented to familiarize participants with the task. During the AX task, they were instructed to listen to two Korean CV syllables and decide whether they were the same or different. The AX task consisted of 144 trials, which were randomized and presented in four blocks of 36 trials (9 contrast types * 4 tokens * 2 same/different types * 2 repetitions).

Next, participants completed the English familiarization phase to become familiar with the perceptual assimilation task and to ensure they associated the initial consonants in the keywords with intended English phonemes. Participants heard an English word (e.g., '*pass*' /pæs/) and selected one of the twelve keywords with the same initial consonant sound (e.g., the correct answer is *pack*). Next, they rated how good of an exemplar of the initial consonant sound is the initial consonant in the selected keyword on a 7-point Likert scale (see Figure 1). They completed forty-eight randomized trials (12 keywords * 4 English words). Next, participants completed the perceptual assimilation task. The participants followed the same procedure as in the familiarization phase, but a Korean CV syllable was presented in each trial. They completed 108 trials (9 consonants * 4 tokens * 3 repetitions) blocked by repetition; each block comprises 36 trials.

Data Analysis

The individual mean proportion correct for each initial consonant in the English keywords in the familiarization phase was computed to ensure intended associations between the initial consonants in English keywords and those of English auditory stimuli (e.g., [pæs] – *pack*). The overall mean accuracy was above .96, except for words beginning with /ð/, /θ/, and /v/. The mean

accuracy of /θ/ and /ð/ identification was .67 and .73, respectively. This was expected, as Hayes-Harb and Barrios (2021) previously found that not all L1 English listeners accurately distinguish English /θ/ and /ð/, possibly due to the identical orthography. Accordingly, we coded both *Thank* and *That* keyword responses for words beginning with /θ/ or /ð/ as correct.

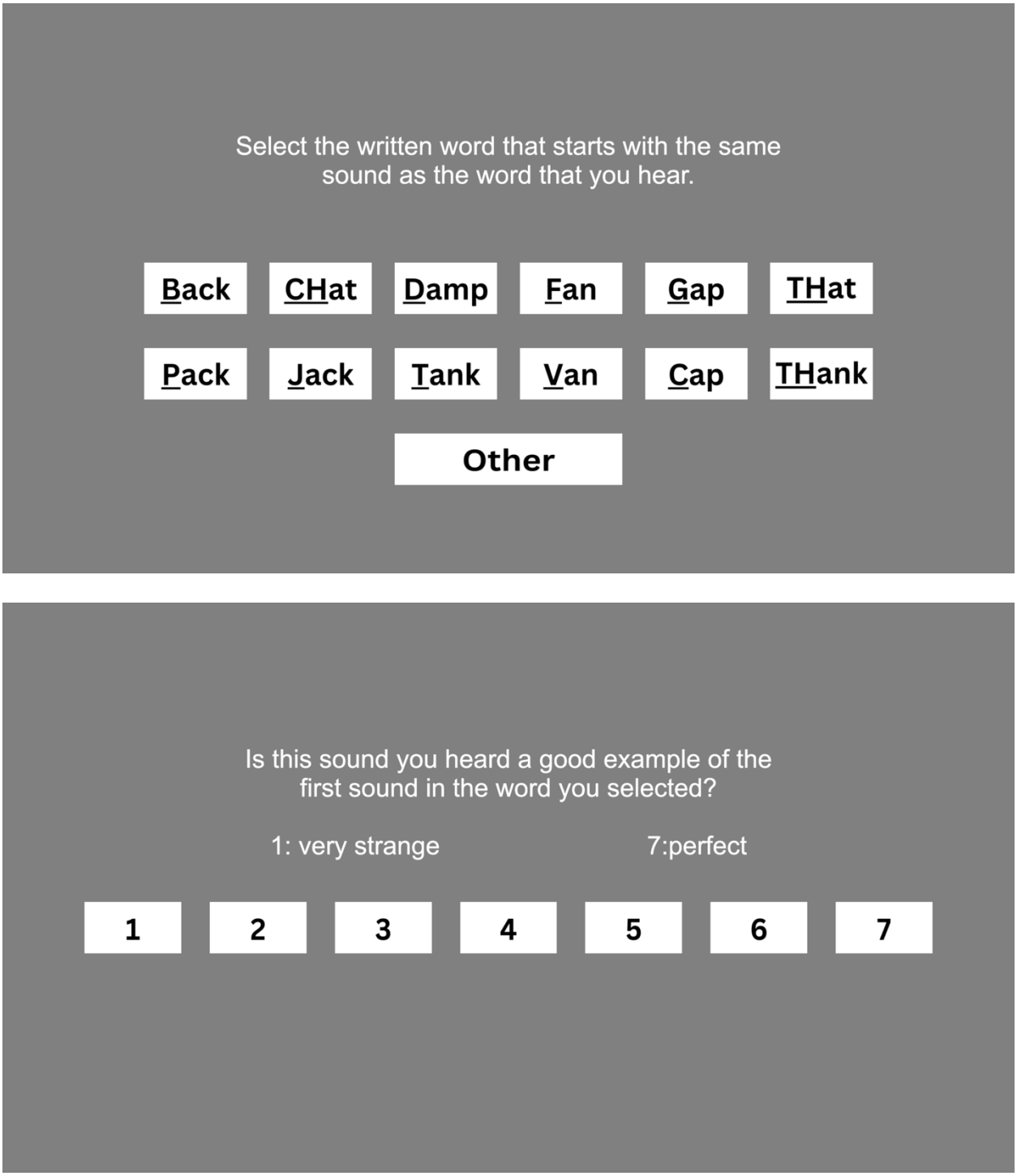

Figure 1. Screen displayed during the perceptual assimilation task (top) and rating task (bottom)

Participants whose mean accuracy was below .5 for the familiarization phase were excluded from further data analyses ($N = 4$). Data from 41 participants who met this performance-based inclusionary criteria were analyzed. To examine how English listeners perceive Korean stop consonants, the mean proportion of the English keyword selection for each Korean phoneme was computed. To determine how good of an exemplar of the Korean phoneme is the initial English consonant in the keyword, mean rating scores for the same Korean phoneme-English keyword mappings were computed.

To examine English listeners' discriminability of Korean stop contrasts, the mean accuracy for the nine contrasts was computed. Moreover, in addition to the mean proportion accuracy a *d*' score, a measure of sensitivity to contrasts, was computed for each participant and each contrast to factor out response bias. Lower *d*' scores indicate lower sensitivity to a contrast when *d*' scores are above 0, and negative *d*' scores indicate a misunderstanding of keys corresponding to 'same' and 'different' responses. A 0 *d*' score indicates no discriminability (Macmillan & Creelman, 2005). To analyze the *d*' data, we first conducted an omnibus analysis using a linear mixed effects model using the *lme4* package (version 1.1–35.1; Bates et al., 2015) in R (R Core Team, 2023). The fixed effect of ContrastType (Lenis-Aspirated, Lenis-Fortis, Aspirated-Fortis) and Place (Bilabial, Alveolar, Velar) was treatment-coded with Lenis-Aspirated and Alveolar as the reference groups, including random by-participant intercepts.

RESULTS

Perceptual Assimilation

As shown in Table 3, the mean proportions of English /p/ selection for Korean /p/ and /pʰ/ were .96 and .95, respectively. In other words, Korean /p/ and /pʰ/ were identified as an instance of /p/ in *pack* 96% and 95% of the time, respectively.

Table 3

Mean Proportions of English Stops for Each Korean Stop with Rating Scores in Parentheses

	English					
Korean	/b/	/p/	/d/	/t/	/g/	/k/
/p/		.96 (5.42)				
/pʰ/		.95 (5.55)				
/p*/	.92 (5.42)					
/t/				.88 (4.95)		
/tʰ/				.89 (5.1)		
/t*/			.75 (4.99)			
/k/						.89 (5.24)
/kʰ/						.92 (5.39)

/k*/	.88 (5.4)

Note. As a random response rate is 7.7%, mean proportions less than 7.7% were not reported.

A paired t-test revealed that the mean rating scores for Korean /p/ (5.42) and /pʰ/ (5.55) were not significantly different ($t(40) = –1.38$, $p = .18$), suggesting that Korean /p/ and /pʰ/ are equally good exemplars of English /p/. This indicates that English listeners showed Single-Category assimilation of Korean bilabial lenis and aspirated stops. The mean proportion of English /b/ selection for Korean /p*/ was .92 with a 5.42 rating score. As Korean /p*/ was identified as an instance of English /b/, bilabial lenis–fortis and aspirated–fortis contrasts exhibit a Two-Category assimilation pattern.

These patterns of assimilation also hold for alveolar stops. The mean proportions of English /t/ selection for Korean /t/ and /tʰ/ were .88 and .89 with the mean rating scores of 4.95 and 5.1, respectively. The mean rating scores were not significantly different ($t(39) = –1.8$, $p = .08$), indicating Single-Category assimilation for Korean alveolar lenis-aspirated contrast. The mean proportion of English /d/ selection for Korean /t*/ was .75 with a 4.99 rating score.

Likewise, the mean proportion of English /k/ selection for Korean /k/ and /kʰ/ was .89 and .92 with 5.24 and 5.39 rating scores, respectively. However, the mean rating scores for Korean /k/ (5.24) and for Korean /kʰ/ (5.39) differed significantly ($t(40) = –2.36$, $p = .02$), resulting in a Category–Goodness assimilation pattern for the velar lenis–aspirated contrast. This indicates that Korean /kʰ/ is a better exemplar of English /k/ than Korean /k/. The mean proportion of English /g/ selection for Korean /k*/ was .88 with a mean rating of 5.4. Accordingly, the lenis-fortis and aspirated-fortis contrasts (Two-category pattern) are expected to be discriminated more accurately than the lenis-aspirated contrast (Single-category and Category-goodness patterns).

AX Discrimination

As the PAM predicts, the lenis-aspirated contrast was least accurately discriminated (see Table 4). The mean accuracy of discriminating the bilabial lenis-aspirated contrast was .18, for the alveolar contrast was .2, and for the velar contrast was .21. The mean accuracy of the lenis-fortis and aspirated-fortis contrasts was above .93 regardless of place of articulation.

The omnibus analysis revealed that there was a significant main effect of ContrastType ($F(2, 320) = 568.12$, $p < .001$), no significant effect of Place ($F(2, 320) = 1.79$, $p = .17$), and a significant interaction of ContrastType and Place ($F(4, 320) = 4.18$, $p < .001$). To determine whether participants as a group differed in their ability to discriminate the three contrast types within each place of articulation, we conducted three separate statistical analyses of *d*' scores using the same model and package. The fixed effect of ContrastType was treatment-coded with Lenis-Aspirated as the reference group, including random by-participant intercepts. The analysis revealed that the mean *d*' score of the bilabial Lenis-Aspirated contrast was significantly different from those of the bilabial Lenis-Fortis ($\beta = 2.21$, $SE = .12$, $t = 17.82$, $p < .001$) and Aspirated-Fortis ($\beta = 2.38$, $SE = .12$, $t = 19.18$, $p < .001$), indicating that it was more challenging for participants to discriminate the bilabial Lenis-Aspirated than the other two contrasts.

Table 4

Mean Accuracy and Standard Deviation by Contrast Type and Place of Articulation
Note. Standard deviations are indicated in parentheses.

	Bilabial	Alveolar	Velar
Lenis-aspirated	0.18 (0.2)	0.2 (0.21)	0.21 (0.2)
Lenis-fortis	0.95 (0.13)	0.96 (0.08)	0.97 (0.09)
Aspirated-fortis	0.93 (0.17)	0.95 (0.1)	0.93 (0.15)

Similar to bilabial stops, the mean *d*' score of the alveolar Lenis-Aspirated contrast was significantly different from those of the alveolar Lenis-Fortis ($\beta = 2.09$, $SE = .12$, $t = 17.64$, $p < .001$) and Aspirated-Fortis ($\beta = 1.95$, $SE = .12$, $t = 16.47$, $p < .001$), indicating that it was more challenging for participants to discriminate the alveolar Lenis-Aspirated than the other two contrasts. Likewise, the mean *d*' score of the velar Lenis-Aspirated contrast was significantly different from those of the velar Lenis-Fortis ($\beta = 1.68$, $SE = .12$, $t = 13.71$, $p < .001$) and Aspirated-Fortis ($\beta = 1.92$, $SE = .12$, $t = 15.64$, $p < .001$), indicating it was more challenging for them to discriminate the velar Lenis-Aspirated than the other two contrasts.

We conducted the additional three subset analyses to better understand the significant interaction and determine whether participants discriminated the same contrast type differently across places of articulation. The fixed effect of Place was treatment-coded with Velar as the reference group, including random by-participant intercepts. The analysis revealed that the mean *d*' score of the velar lenis-aspirated contrast was significantly different from those of bilabial ($\beta = –0.29$, $SE = .1$, $t = –2.92$, $p < .01$) and alveolar ($\beta = –0.23$, $SE = .1$, $t = –2.35$, $p < .05$), indicating the velar lenis-aspirated contrast was less challenging to discriminate than the bilabial and alveolar. The analysis with lenis-fortis contrasts revealed that the mean *d*' score of the velar lenis-fortis contrast was significantly different from the bilabial ($\beta = 0.24$, $SE = .11$, $t = 2.15$, $p < .05$), but it was not significantly different from the alveolar ($\beta = 0.18$, $SE = .11$, $t = –1.54$, $p = .13$), indicating the velar lenis-fortis contrast was less accurately discriminated than the bilabial, but not than the alveolar. Lastly, the analysis with aspirated-fortis contrasts revealed that the mean *d*' score of the velar aspirated-fortis contrast was not significantly different from the bilabial ($\beta = 0.18$, $SE = .1$, $t = 1.75$, $p = .08$) and alveolar ($\beta = –0.2$, $SE = .1$, $t = –2.0$, $p = .05$), indicating that discrimination performance did not differ across aspirated-fortis contrasts of different places of articulation.

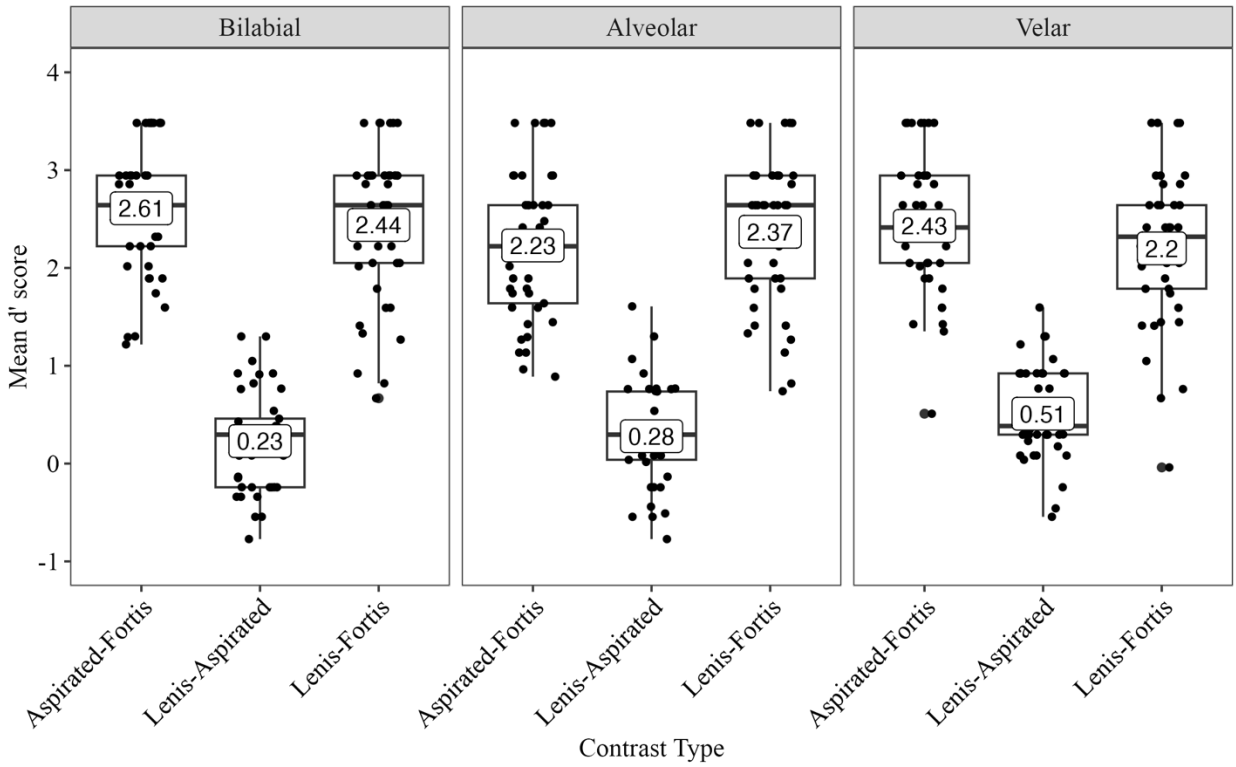

Figure 2. Boxplot of *d*' scores by Contrast Type and Place of Articulation. Each dot indicates the *d*' score of an individual.

DISCUSSION

In the present study, we examined the perceptual assimilation and discrimination of Korean stop consonants by inexperienced English listeners with the goal of determining which contrasts are expected to pose the greatest difficulty for L2 beginners. First, we observed that inexperienced English listeners perceptually assimilated Korean lenis and aspirated stops to English aspirated voiceless stops. Mean rating scores revealed a distinct perceptual assimilation pattern for Korean velar stops; Korean aspirated velar stop (i.e., /k^h/) was perceived as a better exemplar of English /k/ than Korean lenis velar stop (i.e., /k/), showing a Category-Goodness assimilation pattern. With respect to bilabial and alveolar stops, both lenis and aspirated stops were equally good exemplars of English stops, exhibiting a Single-Category assimilation pattern. This finding is consistent with previous studies of English listeners' perception of Korean stops (Nagle et al., 2023; Schmidt, 2007), except for Korean velar stops.

Korean fortis stops (i.e., /p*, t*, k*/) were perceptually assimilated to English voiced stops (i.e., /b, d, g/), resulting in a Two-Category assimilation of lenis-fortis and aspirated-fortis contrasts. Interestingly, the finding diverges from the variable mapping of Korean fortis to both English voiced and voiceless stops observed by Nagle et al. (2023). The divergent findings might be due to the fact that participants in Nagle et al. (2023) listened to productions of four Korean talkers, while those in the current study listened to productions of a single talker.

In the AX discrimination task, we found evidence for differences in discrimination difficulty among the three contrasts. In particular, the bilabial and alveolar lenis-aspirated contrasts (Single-Category) and velar contrast (Category-Goodness) were relatively more difficult to discriminate than the lenis-fortis and aspirated-fortis contrasts (Two-Category), as predicted by the PAM on the basis of the perceptual assimilation data from the same participants. While this finding stands in contrast with Kwon's finding that the lenis-fortis contrasts were the least accurately discriminated, it lends support for the PAM, while providing a new set of data on English listeners' perception of Korean word-initial stop contrasts.

Our study is not without limitations. For example, the Korean stops were examined in a single phonological context (i.e., a word–initial position), but it would be interesting to investigate how English listeners' perception changes in the word-final position, where manner neutralization of codas occurs (Kim & Jongman, 1996). Additionally, because a single Korean talker produced the stimuli, the difficulty posed here might be limited to the production of the specific talker (see Appendix A for VOT and F0 of the Korean stops of this speaker). The production of Korean talkers varies in their acoustic properties (Lee & Jongman, 2012), and this talker difference would be expected to influence the perception.

In sum, we investigated inexperienced English listeners' perceptual assimilation and discrimination of Korean word-initial stop consonants. As the PAM predicts, inexperienced English listeners exhibited greater difficulty discriminating the lenis-aspirated than the lenis-fortis and aspirated-fortis contrasts. These findings confirmed the predictions of the PAM and suggest that the Korean lenis-aspirated contrasts are likely to pose considerable difficulty in distinguishing them for English learners at early stages of learning. The current study suggests several possibilities for future work, including a closer examination of the influence of talkers on English listeners' perception of Korean stop consonants in different phonological contexts. In the context of L2 learning, examining whether various types of orthographic exposure may facilitate the learning of difficult novel contrasts may be useful for English learners of Korean.

ACKNOWLEDGMENTS

I am grateful to all the participants for their participation, members of the Speech Acquisition Lab and CLARG for providing feedback, Korean listeners, SeungKyung Kim and MinHye Joo, for their contribution to this work, and my mentors, Shannon Barrios and Rachel Hayes-Harb, who provided support and help throughout the study.

ABOUT THE AUTHOR

Eunjin Lee is a Ph.D. student and a member of the Speech Acquisition Lab at the Department of Linguistics at the University of Utah. Her primary research interests are L2 phonology, psycholinguistics, and second language acquisition. In particular, she is currently interested in a better understanding of how adult L2 learners perceive and learn novel speech sounds.

Contact information:
University of Utah,
Department of Linguistics

Languages and Communication Bldg
255 S Central Campus Dr., RM 2300
Salt Lake City, UT 84112
eunjin.lee@utah.edu

REFERENCES

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed–Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. doi:10.18637/jss.v067.i01.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775–794.

Best, C. (1995). A direct realist view of cross–language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (pp. 171–206). York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and 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. 13–34). John Benjamins.

Boersma, P. (2011). Praat: doing phonetics by computer [Computer program]. *http://www.praat.org/*. https://cir.nii.ac.jp/crid/1571417124269710464.

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, *61*, 30–47.

Dmitrieva, O., Llanos, F., Shultz, A. A., & Francis, A. L. (2015). Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English. *Journal of Phonetics*, *49*, 77–95.

Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). chapter, Cambridge: Cambridge University Press.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross–Language Research* (pp. 233–277). York Press.

Hayes-Harb, R., & Barrios, S. (2021). Native English speakers and Hindi consonants: From cross-language perception patterns to pronunciation teaching. *Foreign Language Annals*, *55*(1), 175–197.

Kim, H., & Jongman, A. (1996). Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics*, *24*(3), 295–312.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606–608.

Kwon, N. (2014). *Acoustic observation for English speakers' perception of a three–way laryngeal contrast of Korean stops*. https://minerva–access.unimelb.edu.au/handle/11343/40957.

Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *42*(2), 145–169.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Psychology Press.

Lee-Ellis, S. (2012). *Looking into bilingualism through the heritage speaker's mind.* Doctoral dissertation, University of Maryland, College Park: ProQuest Dissertations Publishing.
Nagle, C., Bruun, S., & Baese-Berk, M. (2023). Perceptual Assimilation of Spanish, Thai, and Korean Stop Consonants by Native English Speakers. *20th International Congress of Phonetic Sciences, Prague, Czech Republic*.
Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.
Qualtrics. (2005). *Qualtrics*. https://www.qualtrics.com
R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
Schmidt, A. M. (2007). Cross–language consonant identification. In O-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 185-200). John Benjamins.
Seo, Y., Dmitrieva, O., & Cuza, A. (2022). Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners. *Languages*, *7*(1), 6
Vicenik, C. (2009). *Intensity–scaler* [Praat]. http://phonetics.linguistics.ucla.edu/facilities/acoustic/IntensityScaler.txt.
Werker, J. F., & Tees, R. C. (1984). Cross–language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*(1), 49–63.

APPENDIX A

Average values of VOT (ms) and f0 (Hz) of the stimuli

		VOT	*f*0
Korean	Lenis	81.45	167.13
	Aspirated	115.8	174.26
	Fortis	12.64	179.88