Proceedings of the LSA

# Jack or Zack: Bilingual experience and lexical encoding of English /dʒ-z/ in Korean bilinguals

Naim Lim & Yoonjung Kang *

**Abstract**. This study examines the perceptual discrimination and lexical encoding of the English /dʒ–z/ contrast in Korean–English bilinguals, aiming to tease apart the roles of early exposure and misperception in shaping bilingual lexical representations. The English /dʒ–z/ contrast is not contrastive in Korean and is typically neutralized and adapted as the lenis affricate /c/ [t͡s] in Korean loanwords. We test heritage bilinguals with early exposure to Korean and later dominance in English in comparison to native Korean and English speakers. Participants completed a lexical decision task involving cross-modal priming followed by an auditory AX discrimination task. In critical lexical decision trials, a visual orthographic target (e.g., *Zack*) was presented following one of three types of auditory primes: Match (identical to the target: *Zack*), Mismatch (a minimal pair of the target: *Jack*), or Control (unrelated to the target: *gear*). In the auditory AX discrimination task, the same participants were tested on their ability to auditorily discriminate between the minimal pairs used in the lexical decision task. The Heritage group performed similarly to the native Korean group in the lexical decision task, even though they performed similarly to the native English group in the AX discrimination task, unlike the native Korean group. The Heritage group's performance supports the view that early exposure to a language can play a critical role in shaping lexical representations, even when language dominance shifts to another language later in life.

**Keywords**. bilingualism; heritage speakers; priming; perception; lexical access

**1. Introduction.** Different languages exhibit different phonological systems. In many cases, second language (L2) learners must learn to produce and perceive phonological contrasts that are contrastive in the L2 but non-contrastive in their first language (L1). Previous research shows that L2 listeners experience interference from their L1 phonological system during speech perception. For example, it is well known that Chinese, Korean, and Japanese learners of English have difficulty categorizing and discriminating the English /ɹ–l/ contrast, which is non-contrastive in their native languages (Aoyama et al., 2004).

These learners are not only faced with the task of learning to perceive non-native contrasts, but they must also be able to accurately encode and categorize words with the correct phonological forms into their lexicon (i.e., phonolexical forms). This raises questions about how bilinguals access and represent words containing L2-specific phonological contrasts in their mental lexicons. Studies examining bilinguals' lexical representations using lexical decision

* This work arose from the Master of Arts research paper of the first author. We thank Phil Monahan for valuable suggestions on the study design and feedback on an earlier version of this work. We are also grateful to the Department of Linguistics at the University of Toronto for the Graduate Research Award and to the University of Toronto Scarborough Office of the Vice-Principal Research and Innovation, for the Departmental Research Fund that supported this work. Special thanks to Colin Keith-Hill for recording the stimuli, and to Juhyung Cho and Laura Griffin for their assistance with participant recruitment. Authors: Naim Lim, Department of Linguistics, University of Toronto (naim.lim@mail.utoronto.ca) & Yoonjung Kang, Department of Language Studies, University of Toronto Scarborough / Department of Linguistics, University of Toronto (yoonjung.kang@utoronto.ca).

© 2026 Author(s). Published by the LSA with permission of the author(s) under a CC BY license.

tasks (Pallier et al., 2001; Sebastián-Gallés et al., 2005) have found effects of the native language, typically both the language they are initially exposed to in life as well as their dominant language, when processing words with the non-native phonological contrasts. Heritage speaker bilinguals are an understudied population in this line of research. Heritage speakers are exposed to a heritage language (i.e., the maternal or home language) from birth but generally attain greater proficiency in their second language, the majority societal language. This unique bilingual profile offers an opportunity to disentangle the confound between initial, early language exposure and language dominance in shaping bilinguals' lexical representations of perceptually difficult phonological contrasts.

**2. Background.** Pallier et al. (2001) investigated the consequences that perceptual difficulty may have on the lexical encoding of L2 minimal pairs differing only in a phoneme that constitutes a non-native phonological contrast. In an auditory repetition lexical decision task, both members of minimal pairs involving the Catalan /e–ɛ/, /o–ɔ/, and /s–z/ contrasts, which are not contrastive in Spanish and are perceptually difficult for Spanish speakers, were activated whenever either word was heard by Spanish L1–Catalan L2 bilinguals. For example, hearing /netə/ "granddaughter" earlier in the task facilitated the recognition of its minimal pair counterpart /nɛtə/ "clean, FEM.". This pattern contrasts with that of Catalan L1–Spanish L2 bilinguals, who did not show any facilitation from minimal pairs. These results were interpreted as suggesting that inaccurate perception of non-native contrasts at the time of learning has consequences for lexical encoding in L2 Catalan speakers: minimal pair words are stored with the same phonolexical representation, effectively as merged, homophonous lexical entries (Pallier et al., 2001).

Early speech development research has shown that within the first year of life, infants become attuned to the phonetic categories of their L1. Early exposure to the L1 alone can have lasting effects, with heritage speakers outperforming late L2 learners in perception and production, even when the language is not actively maintained in adulthood. (Au et al., 2002; Knightly et al., 2003; Oh et al., 2003). Language dominance, encompassing factors such as language history, proficiency, usage, attitudes (e.g., the BLP Questionnaire; Gertken et al., 2014), has been used to categorize groups in lexical decision task studies (Sebastián-Gallés, 2005; Amengual, 2016; Soo & Monahan, 2023). In most of these studies, bilinguals' L1 is also the language in which they are dominant. For example, Sebastián-Gallés, Echeverría, and Bosch (2005) used a near-word lexical decision task paradigm to test Spanish–Catalan bilinguals, including groups who had early exposure and dominance in one of the two languages, as well as simultaneous bilinguals. The results showed that even highly skilled Spanish L1–Catalan L2 bilinguals exhibited effects of early exposure to and dominance in Spanish: Catalan word representations involving the /e–ɛ/ contrast could be accessed with the incorrect vowel. In the lexical decision task, these participants accepted nonwords with the incorrect vowel more often than L1 Catalan speakers, for example treating */gəʎeðə/ as [gəʎɛðə] "bucket" or */uʎɛɾəs/ as [uʎeɾəs] "glasses." A similar pattern was found for bilinguals exposed to both languages from birth, who still showed higher error rates than native Catalan speakers, accepting items with the wrong vowel as real words. This suggests that early exposure alone is not sufficient to guarantee native-like phonolexical behaviour.

Darcy et al. (2025) point out the crucial importance of jointly considering three interdependent aspects of L2 phonolexical behaviour: perceptual processing of the input, the content and form of lexical representations, and lexical access. If it is not known how participants perceive the contrasting sounds in a task, it cannot be concluded that spurious lexical activation of either phantom (i.e., near-word) competitors (Broersma & Cutler, 2007) or minimal pairs is

taking place. Because L2 learners may have difficulty perceiving the contrast, the possibility exists that their performance in lexical tasks reflects a lexical access issue, where only one member of a minimal pair is accessed by both phonetic inputs during spoken-word recognition, rather than spurious lexical activation at the representational level. Although Pallier et al. (2001) did not separately test participants' perception of the contrast, subsequent studies have found similar non-native patterns in lexical decision tasks even when listeners could perceive the contrast. Amengual (2016) found that, even though early, proficient Spanish L1–Catalan L2 bilinguals could distinguish Catalan vowel contrasts at the perceptual level, they showed higher error rates when rejecting nonwords (e.g., /ˈbɔsk/, "forest" → */ˈbosk/). Darcy & Holliday (2019) examined L1 Mandarin–L2 Korean learners using a vowel identification task alongside a nonword rejection paradigm targeting the difficult /o–ʌ/ contrast (e.g., /sogɛ/ "introduce" → */sʌgɛ/). Once again, learners were generally able to identify the vowels accurately yet still showed high error rates in the lexical task, confirming that accurate perception does not guarantee precise phono-lexical representations. Interestingly, they also found that recently learned words were represented more accurately than earlier-learned words, suggesting that lexical updates may enter the system through new vocabulary, while older representations remain less target-like. Soo & Monahan (2023) assessed second-generation Cantonese heritage speakers on both their perceptual and lexical encoding of Cantonese tone using an AX discrimination task and a medium-term auditory repetition lexical decision task. In the discrimination task, heritage and native Cantonese speakers showed comparable sensitivity: both groups could reliably tell apart tones with very different pitch shapes (e.g., a high-level tone vs. a falling tone) but struggled when tones shared the same overall contour. This indicates that early exposure allowed heritage speakers to maintain tone perception at a near-native level. However, in the lexical decision task, the two groups diverged. While native speakers showed no priming for tone minimal pairs, consistent with storing tones as distinct lexical entries, English-dominant heritage speakers showed priming for minimal pairs with shared contours, suggesting weaker or less precise tone representations in the lexicon. This pattern mirrors previous studies showing that bilinguals often treat words containing non-native contrasts as partially overlapping in their L2. Together, these results highlight that although heritage speakers may perceive contrasts accurately, their phonolexical representations can be shaped by dominance in the majority language, making the Cantonese case an example of maintaining an inherited contrast under L2 dominance. This study shows that early exposure confers advantage in perception but early exposure without dominance does not ensure native-like phono-lexical behaviour.

Another methodological approach to testing the robustness of a non-native contrast is the cross-modal lexical decision task, which examines lexical competition between minimal pairs during spoken word recognition (Broersma, 2002, 2012). Since lexical competition is understood to arise between distinct, activated lexical entries during spoken-word recognition (Marslen-Wilson, 1987), tasks tapping into lexical competition can provide insight on the differentiation of lexical representations. Broersma (2002, 2012) conducted a version of the auditory–visual cross-modal priming task (Zwitserlood, 1996) where Dutch–English bilinguals were tested on the lexical competition between competitors forming minimal pairs with the English /æ–ɛ/ contrast, which is non-contrastive in Dutch. In this task, auditory primes (e.g., "flesh") were followed by visual targets (e.g., *flash*), allowing researchers to measure whether recognition of a target word is facilitated or inhibited by a given prime. Broersma (2012) offers two possible interpretations of facilitation by minimal pairs. First, facilitation reflects either merged or homophonous representations that were naturally activated together. Second, facilitation may reflect

nonhomophonous representations that are jointly activated but do not compete because L2 listeners postpone lexical selection until further context is available (in the experiment, until the visual target immediately resolves ambiguity while the prime word is still activated). In contrast, inhibition reflects a native-like competition process, where lexical candidates that lost the competition are deactivated almost immediately (Zwitserlood, 1989). In this study, English listeners showed a typical native-like competition process: minimal pair targets were inhibited after hearing their counterpart (e.g., "flesh" → *flash*). However, Dutch listeners experienced a combination of inhibition and facilitation, with majority of the observations showing facilitation. This pattern aligns with prior results (Pallier et al., 2001; Cutler & Otake, 2004), where a presentation of a minimal pair facilitated recognition of its counterpart for bilinguals' non-native contrasts. For our study, we focus on the English consonantal /dʒ–z/ contrast, which is non-contrastive in Korean and is anecdotally difficult for Korean speakers to produce and perceive. In English loanword adaptation, the two segments are typically neutralized to the lenis affricate /c/ [t͡s], as in *Japan* → [t͡sapan] and *zero* → [t͡seɾo] (Kang, 2013).[1]

2.1. THE PRESENT STUDY. The present study examines the perceptual discrimination and lexical encoding of the English /dʒ–z/ contrast in Korean–English bilinguals using a lexical decision task and an AX discrimination task. Participants include two groups of Korean–English bilinguals: native Korean speakers with initial exposure to and dominance in Korean, and heritage Korean speakers with initial exposure to Korean but dominance in English.

Our central research question is: How do native and heritage bilingual speakers represent the English (L2) /dʒ–z/ contrast at the lexical level? The AX discrimination task investigates listeners' ability to auditorily discriminate between words that differ only in the /dʒ–z/ phoneme. This task uses the same stimuli as the lexical decision task, with of the primary goal of assessing participants' ability to perceive the differences between minimal pairs during lexical processing. The lexical decision task assesses lexical competition using Broersma's (2012) cross-modal lexical decision paradigm. Listeners with robust lexical differentiation between the /dʒ–z/ minimal pairs should experience inhibition when retrieving a word after its minimal pair has been activated, whereas listeners with merged or jointly activated representations are expected to show little or no inhibition. Our study aims to disentangle the roles of early exposure, dominance, and perceptual ability in shaping bilingual lexical representations and their access.

Table 1 summarizes the characteristics of the three participant groups, including language dominance and L1 (the language of first exposure), along with the predicted performance in the two tasks: perception (AX) and lexical decision (LD). In the lexical decision task involving minimal pair primes, listeners may exhibit a facilitation effect (faster reaction time) or an inhibition effect (slower reaction time). Facilitation indicates joint or parallel activation between the prime and target words without competition, whereas inhibition reflects a native-like competition process, in which lexical competitors are deactivated and suppressed almost immediately (Zwitserlood, 1989).

[1] Despite this neutralization, bilinguals appear to maintain a (marginal) distinction between them in certain contexts. For instance, /dʒ/ is sometimes represented with an additional palatal glide <j> in orthography, particularly before back vowels, as in *George* → <죠지> [t͡s**j**o.t͡si], while English /z/ is adapted without <j>, as in *zone* → <존> [t͡son]. In final position, the two sources pattern differently in the quality of the epenthetic vowel, as in *George* → <죠지> [t͡sjo.t͡s**i**], vs. *cheese* → <치즈> [t͡sʰi.t͡s**i**].

| | *Early Exposure* | *Dominance* | *Perception of Contrast (AX)* | *Lexical Competition for Minimal Pairs (LD)* |
|---|---|---|---|---|
| Native English | English | English | Accurate | Inhibition (strong competition) due to robust representations |
| Native Korean | Korean | Korean | Less accurate | Lack of inhibition (weak competition) due to merged / jointly activated representations |
| Heritage Korean | Korean | English | Accurate | a. Pattern like Native English<br>b. Pattern like Native Korean |

Table 1. Participant groups and predicted performance in AX and LD tasks based on early language exposure and language dominance

We expect that the native-English control group (henceforth, "English" Group) will show good discrimination between /dʒ–z/ minimal pairs in the AX discrimination task and strong lexical competition in the LD task, resulting in inhibition. For native Korean speakers (henceforth, "Korean" Group), performance on the AX discrimination task may vary depending on English proficiency. We predict that they will demonstrate the weakest lexical competition and may even show facilitation between /dʒ–z/ minimal pairs, indicating merged or jointly activated representations. For the heritage Korean speakers (henceforth, "Heritage" Group), we predict good, native-like discrimination between /dʒ–z/ minimal pairs due to English dominance. In the lexical decision task, there are two possible outcomes. If dominance in the target language (English) is the determining factor for merged or robust representations, the Heritage group may pattern similarly to the English group. This would align with findings from Sebastián-Gallés, Echeverria and Bosch (2005) and Soo & Monahan (2023) on the effects of language dominance on lexical encoding (though in this case, unlike Soo & Monahan, the L1 is contrast-rich). Alternatively, the Heritage group may pattern closer to the native Korean group, reflecting merged or jointly activated representations despite accurate perception of the contrast. This outcome would replicate prior findings that accurate perception does not necessarily guarantee precise lexical encoding (Amengual, 2016; Darcy & Holliday, 2019; Soo & Monahan, 2023).

**3. Method**. Participants performed two tasks: a lexical decision task and an AX discrimination task. The lexical decision task was designed to test the lexical encoding and competition of /dʒ–z/ minimal pairs in Korean–English bilinguals using a cross-modal priming design. The task will assess whether Korean–English bilinguals would experience facilitation, due to merged representations, or inhibition, due to distinct representations competing in the lexicon, when retrieving a word after its minimal pair had been activated. The AX discrimination task was designed to test participants' ability to auditorily discriminate between /dʒ–z/ minimal pairs. Each participant completed the lexical decision task first, followed by the AX discrimination task, to prevent exposure to the stimuli in the AX discrimination from priming responses in the lexical decision task. The experiment was conducted on-line, implemented in the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020). Participants were restricted to completing the experiment on a computer (desktop or laptop). Participants were required to wear headphones and asked to report the type they used. Before beginning the experimental tasks, participants provided informed consent and completed a language background questionnaire. They then completed two experimental tasks: a lexical decision task and an AX discrimination task.

3.1. PARTICIPANTS. English participants were recruited through Prolific, with the criteria that they were born in and are currently residing in Canada or the USA, and that their first and primary language is English. Korean and Heritage participants were recruited in Canada and Korea via social and academic networks. Native Korean–English bilinguals included individuals who were born in Korea and lived there until at least completing elementary school education (approximately 12 years old). Heritage Korean–English bilinguals included individuals who were born in Canada or immigrated to Canada from Korea before the age of five. A total of 98 individuals participated in the study. From this total, 17 participants were excluded from the main analysis. 14 were excluded because their language background (born in Korea and immigrated to Canada between ages 6–12) did not meet the criteria for either the Heritage group or the Korean group. Two additional participants from the Heritage group were excluded due to their lexical decision task performance, in which they accepted all nonwords, suggesting that they either did not understand the task or were not paying attention to the task. One participant from the native English group was excluded as an extreme outlier on the AX discrimination task, with a markedly lower sensitivity score than the rest of the group.

After exclusions, 81 participants were included in the main analysis: 27 native Korean–English bilinguals, 27 heritage Korean–English bilinguals, and 27 native English speakers. The English group (mean age = 29.96 years) included 14 females and 13 males; 17 participants were born and residing in Canada and 10 in the United States. Five participants reported acquiring an additional language before starting school (Vietnamese, n = 1, Cantonese, n = 2, Greek, n = 1, Mandarin, n = 1). The Korean group (mean age = 25.0 years) included 18 females and 9 males. The Heritage group (mean age = 23.26 years) included 22 females and 5 males. Group averages for self-reported proficiency scores on a scale from 1–10 are reported in Table 2. Our sample sizes are comparable to those of previous studies investigating similar populations (Amengual, 2016; Soo & Monahan, 2023). All participants reported normal speech and hearing and normal or corrected-to-normal vision, provided informed consent, and received compensation. The experiment was conducted online, and participants completed a background questionnaire assessing language history, use, and proficiency.

| *Group* | *Language* | *Speaking* | *Listening* | *Reading* | *Writing* |
|---|---|---|---|---|---|
| Native English | English | 9.56 | 9.67 | 9.52 | 9.52 |
| Native Korean | English | 7.16 | 7.56 | 7.40 | 6.80 |
| | Korean | 9.85 | 9.75 | 9.63 | 9.52 |
| Heritage Korean | English | 9.93 | 10.0 | 9.93 | 9.85 |
| | Korean | 7.69 | 8.12 | 6.62 | 5.92 |

Table 2. Mean self-reported language proficiency scores

3.2. STIMULI. The crucial items consisted of ten minimal pairs of English words contrasting in /dʒ–z/, shown in Table 3. These pairs were used in both the lexical decision task and the AX discrimination task. The list included seven monosyllabic and three disyllabic English /dʒ–z/ minimal pairs. For eight pairs, the /dʒ–z/ contrast appeared in word-initial position, for one in word-medial position, and for one in word-final position.

| /dʒ/ word | /z/ word | Control prime |
|---|---|---|
| Jack | Zack | gear |
| Jen | zen | map |
| jest | zest | grass |
| juice | Zeus | stone |
| Jew | zoo | run |
| Joan | zone | book |
| jealous | zealous | ribbon |
| Joey | Zoe | data |
| region | reason | paper |
| budge | buzz | pool |

Table 3. Experimental items (English /dʒ–z/ minimal pairs)

For the lexical decision task, the primes were always real words and auditorily presented. The targets were visually presented in English orthography and included 10 real word trials containing the critical /dʒ–z/ contrasts, 30 filler real words, and 40 filler nonwords. For the /dʒ–z/ minimal pair targets, a semantically and phonologically unrelated word with the matching number of syllables was selected for each minimal pair. Both items of each minimal pair appeared as visual targets and were presented in one of three conditions. In the Match condition, the visual target was preceded by an auditory prime identical to the target (e.g., auditory "Jack" → visual *Jack*). In the Mismatch condition, the prime was the minimal pair counterpart (e.g., "Zack" → *Jack*), and in the Control condition, the prime was an unrelated word (e.g., "gear"→ *Jack*).

As for fillers, 10 real words with matching primes (e.g., "Dean" → *Dean*), 10 real words with mismatching primes differing only in a single phoneme from the target (e.g., "rice" → *dice*), and 10 real words were paired with control primes (e.g., "lazy" → *Eric*). For the mismatching fillers, the distribution of word-initial, medial, and final contrasts mirrored that of the experimental items: eight word-initial, one word-medial, and one word-final. Also included were 40 filler non-words. Because all auditory primes were real words, nonwords were never presented in the Match condition. Instead, 20 visual nonwords were paired with Mismatch primes, and 20 with unrelated control primes (Table 4). Importantly, none of the fillers with Mismatch primes involved the critical /dʒ–z/ contrast, and the same fillers appeared in all six lists.

Each participants saw one of six lists. Each list included one word from each of the 10 minimal pairs as a visual target (e.g., Either *Jack* or *Zack*, either *zest* or *jest*, etc.), presented in either the Match, Mismatch, or Control condition. In other words, each list contained only one of six possible prime-target combinations for any given minimal pair. The lists were balanced to include 3–4 trials of each prime type (Match, Mismatch, Control), totaling 10 critical target trials per list. The 30 real-word fillers and the 40 non-word fillers were identical across the six lists.

| | Number of trials | Prime Type | Prime (auditory) | Target (orthographic) |
|---|---|---|---|---|
| Real words target | 10 | Match | Jack | Jack |
| | | Mismatch | Zack | Jack |
| | | Control | gear | Jack |
| | | Match | Zack | Zack |
| | | Mismatch | Jack | Zack |
| | | Control | gear | Zack |
| Real words filler | 10 | Match | Dean | Dean |
| | 10 | Mismatch | rice | dice |
| | 10 | Control | lazy | Eric |
| Nonwords filler | 20 | Mismatch | desk | gesk |
| | 20 | Control | William | Valtun |

Table 4. Distribution of experimental items and filler words in the lexical decision task

All real words were recorded by a 23-year-old male native speaker of Canadian English with three repetitions per word. Recordings were made in a sound-attenuated booth using a DPA 4011 unidirectional cardioid microphone and a Zoom H4n digital recorder, with a sampling frequency of 48 kHz and 24-bit quantization. The two best tokens of each word were selected. A 100 ms silent interval was appended before and after the onset and offset of each target acoustic signal, and the intensity was scaled to 70 dB SPL in Praat (Boersma et al., 2025). For the AX discrimination task, only the target /dʒ–z/ minimal pairs were included. Each minimal pair was presented in four trials: two SAME trials (e.g., "Jack"–"Jack", "Zack"–"Zack") and two DIFFERENT trials (e.g., "Jack"–"Zack", "Zack"–"Jack"), for a total of 40 trials (4 trials × 10 minimal pairs). The recordings were separated by 800 ms of silence yielding an interstimulus interval of 1000 ms in Praat. For the SAME trials, two different recordings of the same word were used to ensure that listeners compared phonological representations rather than low-level acoustic cues when comparing the two tokens.

3.3. AX DISCRIMINATION TASK. Note that the discrimination task was conducted after the lexical decision task. However, for ease of exposition, we present the details of the discrimination task first.

3.3.1. **Design and Procedure.** Participants were given instructions in English, informing them that they would hear two spoken words or names in sequence. They were instructed to press 'F' on their keyboard if the two words were the same, or 'J' if the two words were different. Participants completed a training phase of four trials that are different from the main trials, with feedback provided before beginning the experiment. Each participants heard a total of 40 audio pairs presented in randomized order (2 SAME and 2 DIFFERENT trials × 10 minimal pairs). All participants were presented with the same stimuli. Their responses (same or different) for each trial were recorded.

3.3.2. **Results and analysis**. All analyses were conducted in R (R Core Team, 2021) using the lme4 package for mixed-effects modeling (Bates et al., 2015) and the emmeans package for post-hoc comparisons (Lenth, 2025); figures were produced with ggplot2 (Wickham, 2016). Responses from the AX discrimination task were used to calculate $d'$ scores for each participant. $d'$ is a bias-free measure of discriminability, where higher values reflect greater sensitivity to

differences between stimuli, and a d′ score of 0 indicates chance-level performance (Macmillan & Creelman 2004). Figure 1 displays the resulting *d′* scores by group.

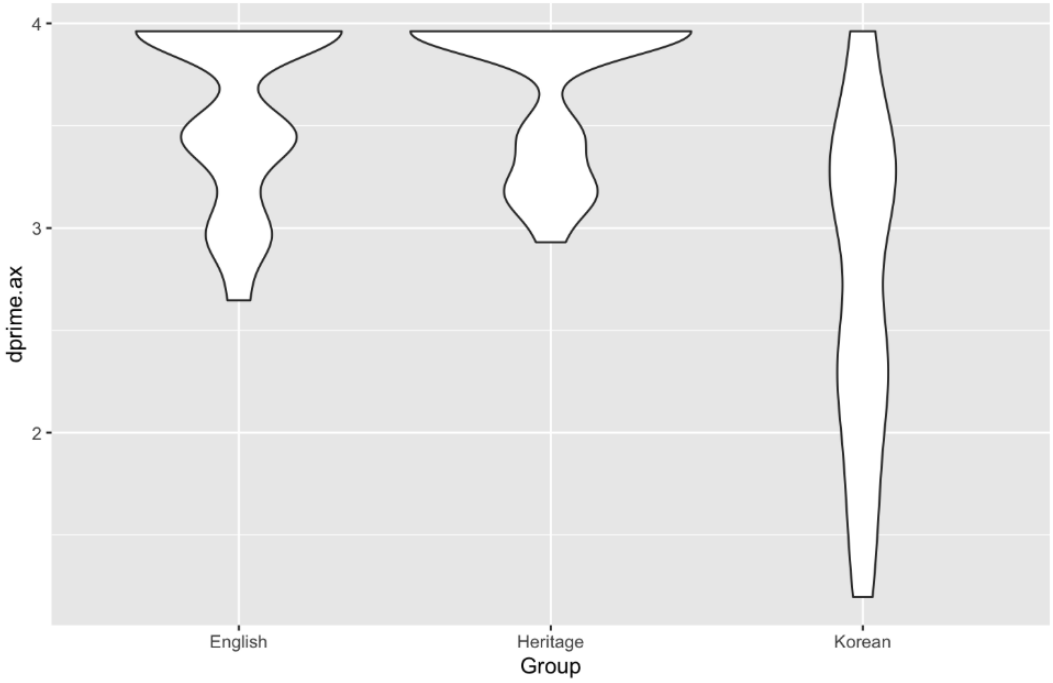


Figure 1. Violin plot of participants' *d'* scores in the AX Discrimination task by Group

As shown in Figure 1, *d′* scores for the English and Heritage groups cluster near the upper end of the scale (≈3–4). By contrast, the Korean group shows a more dispersed distribution, with fewer scores near the ceiling and a longer lower tail. For an inferential test, a linear regression model was fitted with *d′* as the outcome and Group (Heritage, English, Korean) as the predictor. Group was dummy coded, and the Heritage group was set as the reference level. The model output is summarized in Table 5. Results showed that the Korean group had significantly lower *d*' scores than the Heritage group ($\beta = -1.00$, $SE = 0.16$, $t = -6.42$, $p < .001$), while the English group did not differ significantly from the Heritage group ($\beta = -0.20$, $SE = 0.15$, $t = -1.30$, p = 0.198).

| | *β* | *SE* | *t* | *p* |
|---|---|---|---|---|
| (Intercept) | 3.72 | 0.11 | 34.43 | < .001 *** |
| Group: English − Heritage | − 0.20 | 0.15 | − 1.30 | .198 |
| Group: Korean − Heritage | − 1.00 | 0.16 | − 6.42 | < .001 *** |

Table 5. Model output for the linear mixed-effects model. The outcome variable is the *d*' score per participant in the AX discrimination task. The Heritage group was set as the reference level. Significance levels: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$

The pattern of scores indicates that both English and Heritage bilinguals showed strong discrimination of the contrast, with performance clustered near ceiling. In contrast, the Korean group demonstrated noticeably weaker discrimination, with greater variability and more scores falling well below the upper range. Thus, while English and Heritage listeners were comparable in their ability to distinguish the sounds, Korean listeners showed reduced sensitivity overall.

3.4. LEXICAL DECISION TASK. For the lexical decision task, we hypothesize that group and prime condition will interact to influence reaction times in visual word recognition following spoken word primes. Specifically, due to the perception difficulty, we predict Korean speakers' lexical decision will show a lack of inhibition.

3.4.1. **Design and Procedure.** Participants were given instructions informing them that they would hear a word over their headphones, followed by a word appearing on the screen. They were instructed to press 'F' on the keyboard if they thought the visually presented item was an English word or name, and 'J' if they thought it was not. Participants completed a five-trial training block with feedback to familiarize them with the "auditory-prime" → *visual-target* sequence and the key press response. The practice items used neutral English material (no /dʒ–z/ items) and included a mix of 'yes' and 'no' responses: auditory primes "house", "nest", "phase", "scheme", and "lake" paired with visual targets *house* (real), *nost* (nonword), *sweb* (nonword), *Jane* (real name; counted as 'real'), and *bake* (real). Feedback displayed "Correct" or "Incorrect" with the corresponding statement, e.g., " _ is a real word/name" or "_ is not a real word/name," after each trial. On each trial, an auditory stimulus was played over the headphones, and at its offset, the visual target was presented on the screen. The interstimulus interval (ISI) was 100ms, taking into account the silent interval appended at the end of recordings. The next trial began after a response was recorded. Participants' reaction times were measured from the onset of the visual target presentation.

3.4.2. **Results and Analysis**. Accuracy and reaction times (measured in milliseconds from the onset of the visual stimulus) were recorded for each trial. To normalize the distribution, a logarithmic transformation was applied to reaction times (logRTs). Trials with logRTs more than three standard deviations from the mean were removed, with the mean and standard deviation calculated separately for each combination of group, prime type, and target type (experimental item, filler real word, filler nonword).

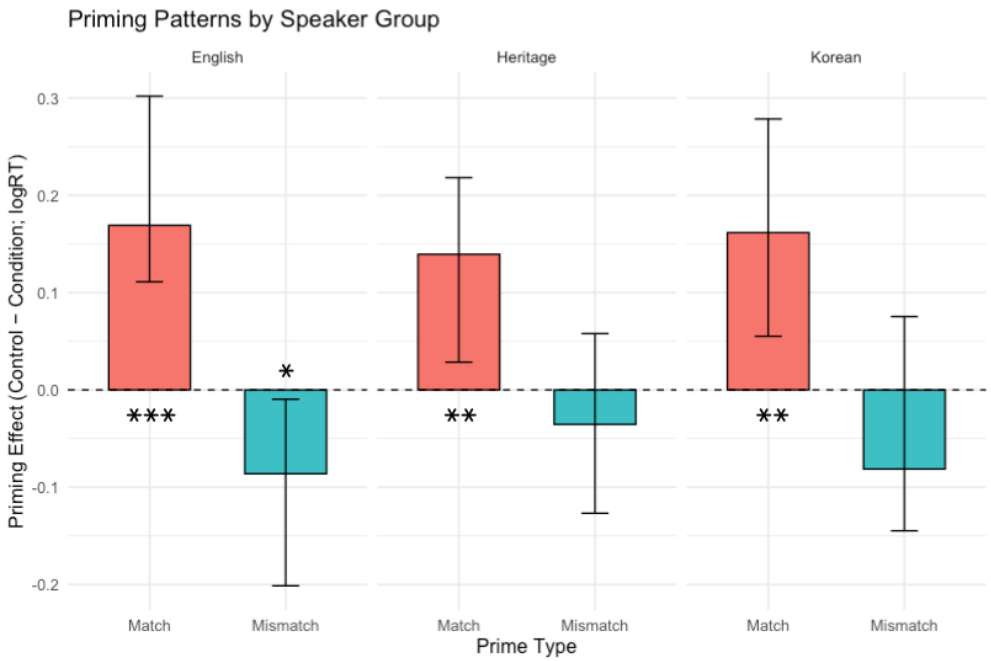


Figure 2. Average priming effects by group and prime type. Error bars represent 95% confidence intervals of the corresponding Control–Condition contrasts based on estimated marginal means from the linear mixed-effects model. *** $p \leq 0.001$,** $p \leq 0.01$,* $p \leq 0.05$

We define a priming effect as the change in reaction time when correctly recognizing a target in the Match or Mismatch condition, relative to mean reaction times for correct trials in the Control condition. Figure 2 displays the mean priming effects by GROUP and PRIMETYPE. Positive values represent faster reaction times than the Control condition indicating facilitation and negative values represent longer reaction times and indicate inhibition. As indicated by their positive values, all three groups showed facilitative priming in the Match condition. In the Mismatch condition, the English group exhibited the most negative priming effect, meaning they exhibited the strongest inhibitioncompared to the Heritage and Korean groups.The log-transformed reaction times (logRTs) were analyzed using a linear mixed-effects model with GROUP, PRIMETYPE, and their interaction as fixed effects. Random intercepts were included for PARTICIPANT**,** TARGET, and LIST. In addition, by-Target random slopes for PRIMETYPE were specified to capture item-specific variation in priming effects. The model was fit using the lmer() function with the *bobyqa* optimizer (Powell, 2009) and a maximum of 200,000 iterations. The model used dummy coding, with the English group as the reference level for GROUP and the Control condition as the reference level for PRIMETYPE.

The fixed-effects output from the model summary is presented in Table 6. The simple effects of GROUP show that in the Control condition, the Heritage group did not differ significantly from the English group ($\beta = -0.08$, $SE = 0.08$, $t(112.23) = -1.03$, $p = .306$), whereas the Korean group was significantly slower than the English group ($\beta = 0.18$, $SE = 0.08$, $t(128.60) = 2.12$, $p = .036$ ). The simple effects of PRIMETYPE show that within the English group, responses in the Mismatch condition were significantly slower than in the Control condition ($\beta = 0.10$, $SE = 0.05$, $t(554.14) = 2.13$, $p = .033$), while responses in the Match condition were significantly faster than in the Control condition ($\beta = -0.21$, $SE = 0.04$, $t(419.39) = -4.63$, $p < .001$). This pattern indicates inhibition in the Mismatch condition and facilitation in the Match condition. No significant interactions were observed (all $p > .10$). However, the direction of the interaction estimates

indicates a reduction in priming effects for both the Heritage and Korean groups relative to the English group. This reduction prevents us from concluding that the effects are equivalent across groups. At the same time, the lack of interaction does not provide evidence that these groups show reliable priming effects, as it may reflect limited statistical power. We therefore conducted post-hoc comparisons to take a closer look at each group's priming pattern using the contrast() function of the emmeans package without p-value adjustments (Table 7).

| | *β* | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| (Intercept) | 6.69 | 0.06 | 101.05 | 110.50 | < .001 *** |
| GROUP: Heritage − English | −0.08 | 0.08 | 112.23 | −1.03 | .306 |
| GROUP: Korean − English | 0.18 | 0.08 | 128.60 | 2.12 | .036 * |
| PRIMETYPE: Mismatch – Control | 0.10 | 0.05 | 554.14 | 2.13 | .033 * |
| PRIMETYPE: Match − Control | −0.21 | 0.04 | 419.39 | −4.63 | < .001 *** |
| GROUP: Heritage × PRIMETYPE: Mismatch | −0.07 | 0.06 | 567.01 | −1.02 | .309 |
| GROUP: Korean × PRIMETYPE: Mismatch | −0.06 | 0.07 | 579.53 | −0.82 | .413 |
| GROUP: Heritage × PRIMETYPE: Match | 0.08 | 0.06 | 570.87 | 1.36 | .174 |
| GROUP: Korean × PRIMETYPE: Match | 0.05 | 0.07 | 580.21 | 0.70 | .484 |

Table 6. Fixed effects output from linear mixed effects model summary

Follow-up tests of PrimeType contrasts within each group showed that all three groups showed clear facilitation in the Match condition. For English listeners, responses in the Match condition were significantly faster than in both the Control condition and the Mismatch conditions. Similar patterns were found for both Heritage and Korean groups. As for the Mismatch condition, while the English group showed a significant inhibition effect ($\beta = -0.10$, $SE = 0.05$, $t(123) = -2.12$, $p = .036$ ), as already shown by the significant simple effects of PRIMETYPE in the regression mode, neither the Heritage nor Korean groups exhibited reliable inhibition (all $p > .05$).

| GROUP | PRIMETYPE | *β* | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|---|
| English | Control − Mismatch | −0.10 | 0.05 | 123 | −2.12 | .036 |
| | Control − Match | 0.21 | 0.04 | 114 | 4.61 | <.0001 |
| | Mismatch − Match | 0.31 | 0.05 | 129 | 6.57 | <.0001 |
| Heritage | Control − Mismatch | −0.03 | 0.05 | 98 | −0.75 | .458 |
| | Control − Match | 0.12 | 0.04 | 105 | 2.74 | .007 |
| | Mismatch − Match | 0.16 | 0.04 | 98 | 3.52 | .001 |
| Korean | Control − Mismatch | −0.04 | 0.05 | 171 | −0.77 | .440 |
| | Control − Match | 0.16 | 0.05 | 166 | 2.98 | .003 |
| | Mismatch − Match | 0.20 | 0.05 | 163 | 3.99 | .0001 |

Table 7. Pairwise comparisons of estimated marginal means for PRIMETYPE within each GROUP

**4. General Discussion**. The present study investigated the perceptual discrimination and lexical competition of the English /dʒ–z/ contrast in native English speakers, heritage Korean–English bilinguals, and native Korean–English bilinguals. This study sought to separate the effects of early exposure and dominance, which were confounded in previous studies, by investigating the heritage speaker population. By employing both an AX discrimination task and a cross-modal

lexical decision task, we examined the role of early exposure, language dominance, and perception in shaping bilingual phonolexical representations.

Results from the AX discrimination task confirmed our predictions: native English and heritage Korean participants demonstrated high perceptual discriminability for the /dʒ–z/ contrast in minimal pairs, while native Korean participants performed significantly worse. This aligns with previous findings that L2 listeners whose L1 lacks the relevant contrast often have difficulty perceiving it (Pallier et al., 2001; Sebastián-Gallés et al., 2005). The finding that heritage participants matched native English listeners in perception supports the view that exposure to English from early childhood promotes native-like discrimination ability (Amengual, 2016).

In the lexical decision task, English listeners exhibited the expected native-like priming pattern: responses in the Match condition were fastest, Control was intermediate, and Mismatch was slowest, reflecting inhibition consistent with robust lexical competition. Heritage listeners, by contrast, only showed facilitation in the Match condition, with no evidence of inhibition in the Mismatch condition despite strong perceptual discriminability. The absence of inhibition points to differences in how heritage bilinguals activate competing lexical representations. According to Broersma's (2012) interpretation, this lack of inhibition suggests that either the representations of the minimal pairs are merged or were both activated without effective lexical competition. The Korean group showed a similar pattern to the heritage group: reliable facilitation in the Match condition, but no inhibition in the Mismatch condition. Together, these findings suggest that while their English dominance allows heritage bilinguals to achieve high levels of perceptual accuracy, their lexical-level processing diverges from native-like patterns, affecting how competing words are activated and suppressed during spoken word recognition. This divergence may be attributed to the Heritage group's initial exposure to Korean. These results indicated that accurate perception does not guarantee fully native-like lexical encoding, mirrors findings from Amengual (2016) and Soo & Monahan (2023).

These findings contribute to models of bilingual lexical representation by underscoring the influence of perception, dominance, and initial exposure on phonolexical behaviour. The heritage group's performance supports the view that initial, early exposure can play a critical role in shaping lexical representations, even when later dominance shifts to the L2. In this view, even limited childhood exposure to a heritage language can confer lasting influences on the representation of phonetic categories in the lexicon. In our data, heritage bilinguals' native-English-like discrimination of the /dʒ–z/ contrast reflects the ability to acquire a contrast in the dominant language, yet their reduced inhibition in the lexical task highlights that perceptual sensitivity does not necessarily translate into robust lexical competition and distinct lexical representations. Several factors may have influenced the observed effects. First, lexical frequency and familiarity of the minimal pairs varied, which may have dampened priming effects. Second, the online testing environment could have introduced variability in reaction time measurements. Third, the relatively small set of minimal pairs may limit generalizability to the broader lexicon. Future research could incorporate a larger and more frequency-controlled stimulus set, test additional contrasts, and manipulate semantic context to further probe competition processes.

## References

Amengual, Mark. 2016. The perception of language-specific phonetic categories does not guarantee accurate phonological representations in the lexicon of early bilinguals. *Applied Psycholinguistics* 37(5). 1221–1251. https://doi.org/10.1017/S0142716415000557.

Anwyl-Irvine, Alexander L., Jessica Massonnié, Adam Flitton, Natasha Kirkham & Jo K. Evershed. 2020. Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods* 52(1). 388–407. https://doi.org/10.3758/s13428-019-01237-x.

Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada & Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32(2). 233–250. https://doi.org/10.1016/S0095-4470(03)00036-6.

Au, Terry Kit-fong, Leah M. Knightly, Sun-Ah Jun & Janet S. Oh. 2002. Overhearing a language during childhood. *Psychological Science* 13(3). 238–243. https://doi.org/10.1111/1467-9280.00444.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Boersma, Paul & David Weenink. 2025. Praat: Doing phonetics by computer (Version 6.4.39) [computer program]. https://praat.org.

Broersma, Mirjam. 2002. Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP 2002)*, 261–264. Denver, Colorado: ISCA. https://doi.org/10.21437/ICSLP.2002-130.

Broersma, Mirjam. 2012. Increased lexical activation and reduced competition in second-language listening. *Language and Cognitive Processes* 27(7–8). 1205–1224. https://doi.org/10.1080/01690965.2012.660170.

Broersma, Mirjam & Anne Cutler. 2008. Phantom word activation in L2. *System* 36(1). 22–34. https://doi.org/10.1016/j.system.2007.11.003.

Cutler, Anne & Takashi Otake. 2004. Pseudo-homophony in non-native listening. *Journal of the Acoustical Society of America* 116(4). 2372–2373. https://doi.org/10.1121/1.4780547.

Darcy, Isabelle & Jeffrey J. Holliday. 2019. Teaching an old word new tricks: Phonological updates in the L2 mental lexicon. In John Levis, Charles Nagle & Erik Todey (eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, 10–26. Ames, IA: Iowa State University.

Darcy, Isabelle, Miquel Llompart, Rachel Hayes-Harb, Joan C. Mora, Miren Adrian, Svetlana Cook & Mirjam Ernestus. 2025. Phonological processing and the L2 mental lexicon: Looking back and moving forward. *Studies in Second Language Acquisition* 47(1). 361–387. https://doi.org/10.1017/S0272263124000482.

Gertken, Libby M., Mark Amengual & David Birdsong. 2014. Assessing language dominance with the bilingual language profile. In Pascale Leclercq, Aline Edmonds & Heather Hilton (eds.), *Measuring L2 proficiency: Perspectives from SLA*, 208–225. Bristol: Multilingual Matters.

Knightly, Leah M., Sun-Ah Jun, Janet S. Oh & Terry Kit-fong Au. 2003. Production benefits of childhood overhearing. *The Journal of the Acoustical Society of America* 114(1). 465–474. https://doi.org/10.1121/1.1577560.

Lenth, Russell. 2025. emmeans: Estimated marginal means, aka least-squares means (Version 1.11.0) [R package]. https://CRAN.R-project.org/package=emmeans.

Macmillan, Neil A. & C. Douglas Creelman. 2004. *Detection theory: A user's guide*. 2nd edn. New York: Psychology Press.

Marslen-Wilson, William D. 1987. Functional parallelism in spoken word recognition. *Cognition* 25(1–2). 71–102. https://doi.org/10.1016/0010-0277(87)90005-9.
Oh, Janet S., Sun-Ah Jun, Leah M. Knightly & Terry Kit-fong Au. 2003. Holding on to childhood language memory. *Cognition* 86(3). B53–B64. https://doi.org/10.1016/S0010-0277(02)00175-0.
Pallier, Christophe, Angels Colomé & Núria Sebastián-Gallés. 2001. The influence of Native-Language Phonology on Lexical Access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12(6). 445–449. https://doi.org/10.1111/1467-9280.00383
Powell, M. J. 2009. The BOBYQA algorithm for bound constrained optimization without derivatives. Report. Cambridge: University of Cambridge. 26–46.
R Core Team. 2021. R: A language and environment for statistical computing [computer software]. Vienna: R Foundation for Statistical Computing. https://www.R-project.org.
Sebastián-Gallés, Núria, Sagrario Echeverría & Laura Bosch. 2005. The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and* Language 52(2). 240–255. https://doi.org/10.1016/j.jml.2004.11.001.
Soo, Rachel & Philip J. Monahan. 2023. Phonetic and lexical encoding of tone in Cantonese heritage speakers. *Language and Speech* 66(3). 652–677. https://doi.org/10.1177/00238309221122209.
Wickham, Hadley. 2016. ggplot2: Elegant graphics for data analysis. New York: Springer. https://ggplot2.tidyverse.org.
Zwitserlood, Pienie. 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32(1). 25–64. https://doi.org/10.1016/0010-0277(89)90013-9.