



L2 Processing of Words Containing English /æ/-/ɛ/ and /l/-/ɹ/ Contrasts, and the Uses and Limits of the Auditory Lexical Decision Task for Understanding the Locus of Difficulty

Shannon Barrios* and Rachel Hayes-Harb

Department of Linguistics, University of Utah, Salt Lake City, UT, United States

OPEN ACCESS

Edited by:

Svetlana V. Cook,
University of Maryland, College Park,
United States

Reviewed by:

Jeong-Im Han,
Konkuk University, South Korea
Hanyong Park,
University of Wisconsin–Milwaukee,
United States
Linda Polka,
McGill University, Canada

*Correspondence:

Shannon Barrios
s.barrios@utah.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 31 March 2021

Accepted: 28 June 2021

Published: 20 July 2021

Citation:

Barrios S and Hayes-Harb R (2021) L2
Processing of Words Containing
English /æ/-/ɛ/ and /l/-/ɹ/ Contrasts,
and the Uses and Limits of the Auditory
Lexical Decision Task for
Understanding the Locus of Difficulty.
Front. Commun. 6:689470.
doi: 10.3389/fcomm.2021.689470

Second language (L2) learners often exhibit difficulty perceiving novel phonological contrasts and/or using them to distinguish similar-sounding words. The auditory lexical decision (LD) task has emerged as a promising method to elicit the asymmetries in lexical processing performance that help to identify the locus of learners' difficulty. However, LD tasks have been implemented and interpreted variably in the literature, complicating their utility in distinguishing between cases where learners' difficulty lies at the level of perceptual and/or lexical coding. Building on previous work, we elaborate a set of LD ordinal accuracy predictions associated with various logically possible scenarios concerning the locus of learner difficulty, and provide new LD data involving multiple contrasts and native language (L1) groups. The inclusion of a native speaker control group allows us to isolate which patterns are unique to L2 learners, and the combination of multiple contrasts and L1 groups allows us to elicit evidence of various scenarios. We present findings of an experiment where native English, Korean, and Mandarin speakers completed an LD task that probed the robustness of listeners' phonological representations of the English /æ/-/ɛ/ and /l/-/ɹ/ contrasts. Words contained the target phonemes, and nonwords were created by replacing the target phoneme with its counterpart (e.g., *lecture*/*[j]ecture, *battle*/*b[ɛ]ttle). For the /æ/-/ɛ/ contrast, all three groups exhibited the same pattern of accuracy: near-ceiling acceptance of words and an asymmetric pattern of responses to nonwords, with higher accuracy for nonwords containing [æ] than [ɛ]. For the /l/-/ɹ/ contrast, we found three distinct accuracy patterns: native English speakers' performance was highly accurate and symmetric for words and nonwords (interpreted as evidence that they experienced difficulty at the perceptual coding level), and native Korean speakers exhibited asymmetries in opposite directions for words (favoring [l]) and nonwords (favoring [ɹ]; evidence of difficulty at the lexical coding level). Our findings suggest that the auditory LD task holds promise for determining the locus of learners' difficulty with L2 contrasts; however, we raise several issues requiring attention to maximize its utility in investigating L2 phonolexical processing.

Keywords: second language learning, second language phonology, phonolexical representation, speech perception, Korean, Mandarin, English, lexical decision

INTRODUCTION

Second language (L2) learners are typically faced with the challenge of learning to perceive and produce novel phonemic contrasts, as well as to build a lexicon that effectively encodes the phonetic and phonological information associated with these contrasts. A growing body of research has highlighted the role that representation at the phonolexical level may play in the perseverance of learners' difficulty with novel phonological contrasts, independent of the contributions of perceptual and/or production difficulty alone (e.g., Pallier et al., 2001; Weber and Cutler, 2004; Sebastián-Gallés et al., 2005; Escudero et al., 2008; Hayes-Harb and Masuda, 2008; Broersma, 2012; Amengual, 2016b).

Perceptual and/or lexical encoding difficulty can lead to the activation of inappropriate candidates during spoken word recognition (often referred to as spurious lexical activation), as well as to less efficient competition among competitors (see Broersma and Cutler, 2011 for review). Some evidence of spurious lexical activation comes from auditory lexical decision (LD) tasks in which a participant is required to judge whether an auditory stimulus is a word or not. A ubiquitous finding is that nonnative listeners have difficulty rejecting nonwords derived from real words (e.g., *deaf*/*d[æ]f and *lamp*/*l[ɛ]mp) when these involve confusable L2 phonemes (e.g., Sebastián-Gallés and Baus, 2005; Sebastián-Gallés et al., 2006; Broersma and Cutler, 2011; Díaz et al., 2012; Darcy et al., 2013; Darcy and Thomas, 2019; Melnik and Peperkamp, 2019). Spurious lexical activation is also known to produce priming or facilitation effects for minimal pairs (Pallier et al., 2001), near-words (Broersma and Cutler, 2011; Broersma, 2012), phonologically-related primes (Cook and Gor, 2015), and semantic associates of phonological neighbors (Cook et al., 2016). Indeed, it remains of much interest how and to what extent L2 listeners utilize various sources of contextual information for coping with phonolexical ambiguity (Chrabaszcz and Gor, 2014; Chrabaszcz and Gor, 2017).

Existing neural evidence from ERP corroborates the behavioral findings reviewed above, with nonnative listeners failing to show typical N400 effects (larger N400 responses to nonwords than for words) for nonwords involving confusable phonemes (Sebastián-Gallés et al., 2006; White et al., 2017). Moreover, incorrect lexical decisions by nonnative speakers may not result in error-related negativity which has been observed in native speakers (Sebastián-Gallés et al., 2006). These studies have been important in documenting the difficulties that even highly proficient bilinguals experience with L2 lexical processing. However, such findings typically remain ambiguous as to the locus of the effects. Indeed, such patterns of spurious lexical activation may result from challenges at the perceptual (phonetic coding) and/or phonolexical (lexical coding) levels.

Asymmetries abound in L2 speech perception and lexical processing and have been helpful in shedding light on these issues. They are often associated with situations where an L2 contrast involves two target language phonemes that map to a

single L1 category but with differing degrees of “goodness” (category goodness assimilation according to Best’s Perceptual Assimilation Model; Best, 1995). The better fitting category is often referred to as the dominant category and the other as the non-dominant (unfamiliar or new) category. The latter is typically thought to be less robustly encoded than the former. Studies employing the visual world paradigm have provided evidence of perceptual representations that are neutralized in favor of the dominant category contacting differentiated phonolexical representations (e.g., Weber and Cutler, 2004; Cutler et al., 2006; Escudero et al., 2008). For example, Weber and Cutler (2004) demonstrated that Dutch-English bilinguals experienced spurious activation of English words containing underlying /ɛ/ (e.g., looks to a picture of a ‘pencil’) in response to auditory forms containing [æ] (e.g., “panda”) but not the reverse, suggesting that these bilinguals had established differentiated lexical representations for /ɛ/ and /æ/ words, but that their ability to differentially contact these representations was undermined by neutralization of [ɛ] and [æ] to [ɛ] at the level of speech perception. Escudero et al. (2008) replicated this finding with an artificial lexicon study, demonstrating that learners infer the lexical contrast from the written forms of newly-learned words, and Cutler et al. (2006) similarly provide evidence for differentiated lexical representations for English /l/ and /ɹ/ in native Japanese speakers who perceptually neutralize the contrast. In a similar study involving native German learners of English, Llompart and Reinisch (2017) showed that the English /æ/-/ɛ/ lexical contrast can be inferred from seeing the words articulated even though they are perceptually neutralized in favor of [ɛ]. Llompart and Reinisch (2020) demonstrated that German learners of English can establish distinct lexical representations for /æ/ and /ɛ/ following exposure to minimal pairs during word learning, but that in this case, neutralization at the level of perception unexpectedly favored [æ] (rather than [ɛ]).

In other cases, studies employing auditory LD tasks have uncovered asymmetries in performance that are suggestive of the reverse scenario: differentiated perceptual representations contacting imprecise (i.e., fuzzy) lexical representations of the new category (e.g., Darcy et al., 2013; Melnik and Peperkamp, 2019). In these studies, adult learners are presented with L2 words and nonwords where the nonwords are identical to the words except that one phoneme is replaced with a confusable phoneme. In one experiment, Darcy et al. (2013) presented native English speakers at two levels of L2 German language experience German words (e.g., [honiç] “honey” containing the dominant (i.e., familiar) vowel /o/ and [koniç] “king” containing the non-dominant (i.e., new) vowel /ø/), as well as nonwords created by replacing [o] with [ø] and vice-versa (e.g., *[høniç] and *[koniç]). Darcy et al. (2013) predicted that if participants neutralized the contrast at the level of perception (while maintaining a contrast in the lexicon¹), they would show the following ordinal accuracy pattern:

¹Darcy et al. (2013) remain agnostic as to whether differentiated phonolexical representations are target-like or simply robust enough to distinguish the contrast (i.e., the non-dominant /ø/ category might be stored as “not /o/”).

1. **Word [Dominant]:** Words containing the dominant phoneme will be easy to accept because, e.g., the input [o] is perceived as [o], which matches the underlying phonolexical representation /o/. (Input [honiç] is perceived as [honiç] which matches /honiç/).
2. **Nonword [Dominant]:** Nonwords containing the dominant phoneme will be easy to reject because, e.g., the input [o] is perceived as [o], which does not match the underlying phonolexical representation containing the new phoneme. (Input *[koniç] is perceived as [koniç] which does not match /køniç/).
3. **Word [Non-dominant]:** Words containing the non-dominant phoneme will be difficult to accept because, e.g., the input [ø] is perceived as [o], which does not match the underlying phonolexical representation containing the new phoneme. (Input [køniç] is perceived as [koniç] which does not match /køniç/).
4. **Nonword [Non-dominant]:** Nonwords containing the non-dominant phoneme will be difficult to reject because, e.g., the input [ø] is perceived as [o], which matches the underlying phonolexical representation /o/. (Input *[høniç] is perceived as [honiç] which matches /honiç/).

In this scenario, where the locus of the difficulty is at the level of perceptual coding, both words and nonwords containing the dominant category should be easy to accept and reject, respectively, because learners use accurate perceptual representations of the dominant category to contact lexical representations that encode the contrast. On the other hand, words and nonwords containing the non-dominant category should be more difficult to accept and reject, respectively, due to perceptual neutralization in favor of the dominant category. Darcy et al. (2013) further assume that it will generally be easier to accept words than to reject nonwords (thus 1>2 and 3>4 above). As a result, this scenario, which we will call the perceptual coding scenario “is not expected to yield an interaction between lexical status (word vs. non-word) and category type (old vs. new)” (pp. 379–380). They proposed a second possible scenario, where learners’ perceptual coding of the input preserves the contrast, as does the lexicon; however, the phonolexical representation of the non-dominant category is imprecise, or fuzzy, such that it is activated by inputs containing either member of the contrast. Following Darcy et al. (2013) we use /?/ to indicate that a category is represented imprecisely in the phonolexical representation. According to Darcy et al. (2013), such a scenario should result in the following ordinal accuracy pattern:

1. **Word [Dominant]:** Easy to accept because, e.g., the input [o] is perceived as [o], which matches the phonolexical representation /o/. (Input [honiç] is perceived as [honiç] which matches /honiç/).
2. **Word [Non-dominant]:** Less easy to accept because, e.g., the input [ø] is perceived as [ø], and does not perfectly match the fuzzy phonolexical representation containing the new phoneme. (Input [køniç] is perceived as [køniç] which matches /k?niç/).
3. **Nonword [Non-dominant]:** Easy to reject because, e.g., the input and percept [ø] does not match the phonolexical representation /o/. (Input *[høniç] is perceived as [høniç] which does not match /honiç/).

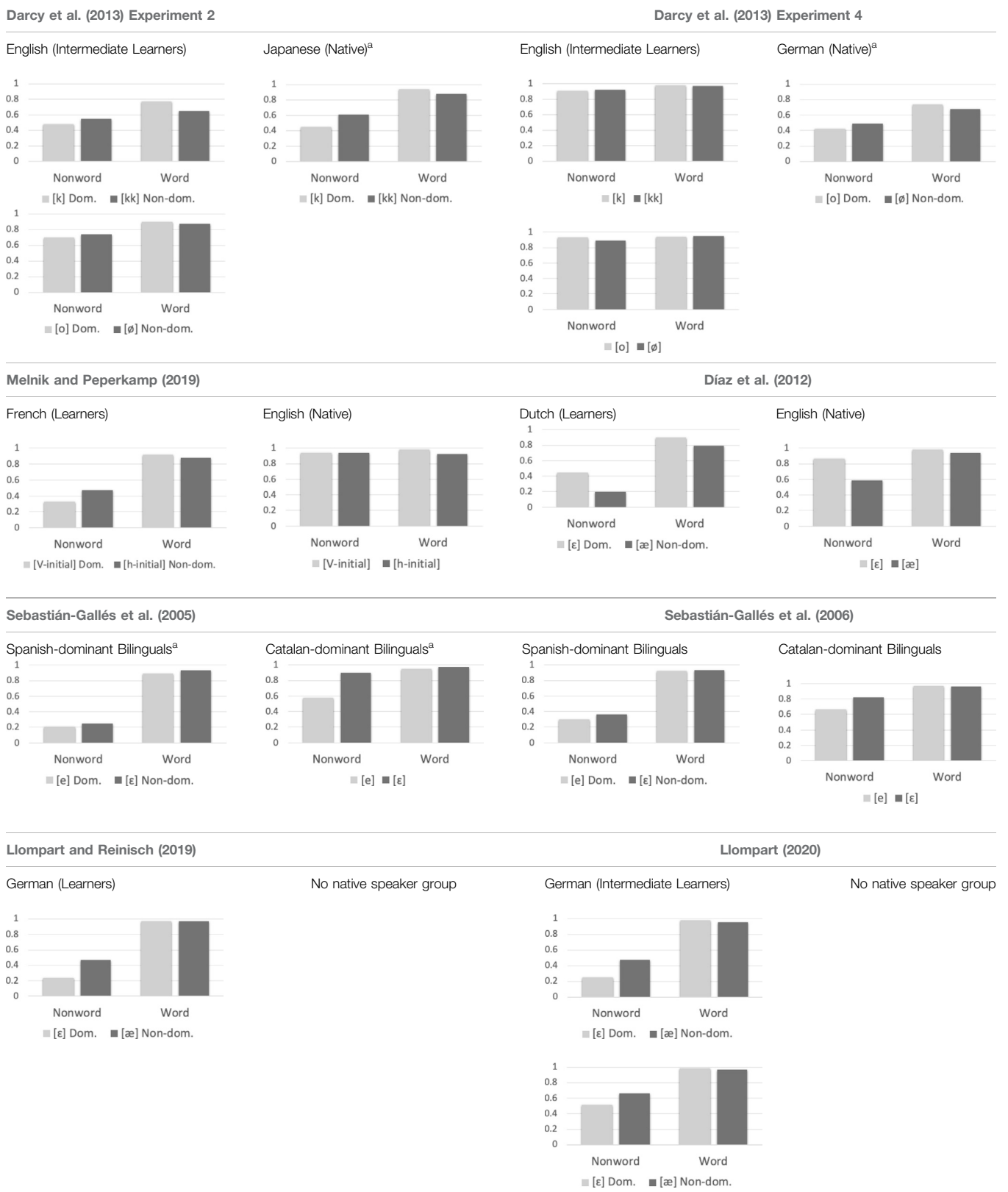
4. **Nonword [Dominant]:** Difficult to reject because, e.g., the input and percept [o] does not mismatch the fuzzy phonolexical representation containing the new phoneme. (Input *[koniç] is perceived as [koniç] which does not mismatch /k?niç/).

In this scenario, where the learner exhibits difficulty at the level of lexical coding, an interaction is expected between lexical status (word, nonword) and segment (dominant, non-dominant), with more accurate performance on words containing the dominant category than words containing the non-dominant category, but more accurate performance on nonwords containing the non-dominant category than nonwords containing the dominant one.

Darcy et al.’s (2013) LD results are summarized **Table 1**, together with the results of several additional studies which are reviewed below. They found that the intermediate-level L1 English learners of German exhibited the interaction of lexical status and segment associated with the lexical coding scenario, with more accurate rejection of nonwords containing [ø] than [o] but more accurate acceptance of words containing [o] than [ø]. The advanced-level learners exhibited a non-significant but descriptively similar pattern. In a separate LD experiment, L1 English learners of Japanese (also at two levels of experience) responded to words containing either singleton (e.g., /k/) or geminate (e.g., /kk/) consonants, in addition to nonwords that were created by replacing singleton consonants with geminates or vice-versa (e.g., [akeru] “to open” / *[akkeru] and [kippu] “ticket” / *[kipu]). Both groups of learners exhibited an interaction of lexical status and segment, with a descriptive pattern of more accurate rejection of nonwords containing [kk] than [k] but more accurate acceptance of words containing [k] than [kk]. Darcy et al. (2013) interpreted this response pattern as evidence that differentiated perception of the Japanese singleton-geminate contrast contacted fuzzy phonolexical representations of the non-dominant geminate consonants. As expected, the native Japanese-speaking control group exhibited high accuracy in all conditions, and no interaction of lexical status and segment. Curiously, however, the native German group exhibited a marginally significant interaction of lexical status and segment that was in the opposite direction of that exhibited by the German learners. This latter finding will be taken up below when we discuss asymmetries in native LD performance.

Additional scenarios beyond those presented by Darcy et al. (2013) are logically possible, depending on the degree of precision associated with lexical encoding. We spell out the full set of predictions for the English /æ/-/ɛ/ contrast in **Table 2**. For this purpose, we treat /ɛ/ as “dominant” (old/familiar) category and /æ/ as “non-dominant” (new/unfamiliar). Native speakers’ phonolexical representations are typically (and often implicitly) assumed to be both distinctive and precise (referred to henceforth as “precise”) in that both members of a given contrast will be encoded such that differentiated perceptual representations will clearly match (e.g., [æ] = /æ/ and [ɛ] = /ɛ/) or mismatch (e.g., [æ] ≠ /ɛ/ and [ɛ] ≠ /æ/). On the other hand, L2 phonolexical representations of non-dominant categories are sometimes characterized as fuzzy or imprecise, though these terms have been used somewhat variably in the literature.

TABLE 1 | Summary of LD findings of previous studies (vertical axis represents mean proportion correct).

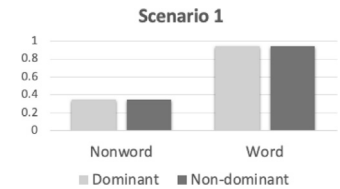


^aData estimated from figures in the article; otherwise, means are specified in the article or were provided by the authors (Díaz, 2021; Llompарт, 2021).

TABLE 2 | Lexical Decision Scenarios and Ordinal Accuracy Predictions. Summary of the eight scenarios created by crossing neutralized and precise perceptual representations with neutralized, ambiguous, “not X,” and precise phonolexical representations. Each condition (Dom = Dominant; Non-dom = Non-dominant) is followed by the corresponding input, percept, phonolexical representation (LexRep) and ordinal accuracy prediction; 1 = highest/high accuracy, 4 = lowest/low accuracy. Each scenario is accompanied by a stylized figure illustrating the ordinal accuracy predictions (note that accuracy predictions are relative, not absolute). Relative predicted accuracy in the following scenarios is computed using the assumption that words are always easier to accept than nonwords are to reject. To create the stylized figures representing the predictions, we imposed a range of 0.75–1.0 for performance on words, and a range of 0.0–0.75 for performance on nonwords (based on trends in the literature). We assigned the following accuracy proportions for words: easy to accept = 0.95, easy-ish to accept = 0.90, difficult to accept = 0.80, and nonwords: easy to reject = 0.70, difficult-ish to reject = 0.45, difficult to reject = 0.35.

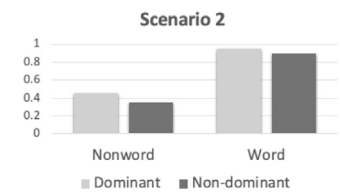
Scenario 1: Neutralized perceptual representation, neutralized phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[lɛmp]	/lɛmp/	Easy to accept (1)
Nonword (Dom)	[lɛmp]	[lɛmp]	/lɛmp/	Difficult to reject (2)
Nonword (Non-dom)	[dæsk]	[dɛsk]	/dɛsk/	Difficult to reject (2)



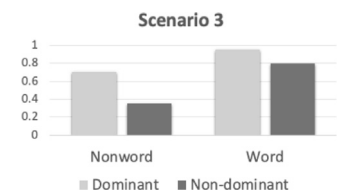
Scenario 2: Neutralized perceptual representation, ambiguous phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[lɛmp]	/l?mp/	Easy-ish to accept (2)
Nonword (Dom)	[lɛmp]	[lɛmp]	/l?mp/	Difficult-ish to reject (3)
Nonword (Non-dom)	[dæsk]	[dɛsk]	/dɛsk/	Difficult to reject (4)



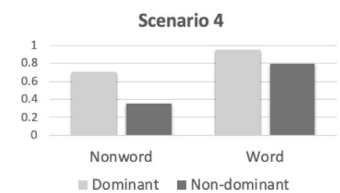
Scenario 3: Neutralized perceptual representation, “not X” phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[lɛmp]	/!{not ɛ}mp/	Difficult to accept (2)
Nonword (Dom)	[lɛmp]	[lɛmp]	/!{not ɛ}mp/	Easy to reject (3)
Nonword (Non-dom)	[dæsk]	[dɛsk]	/dɛsk/	Difficult to reject (4)



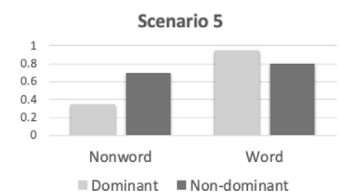
Scenario 4: Neutralized perceptual representation, precise phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[lɛmp]	/læmp/	Difficult to accept (2)
Nonword (Dom)	[lɛmp]	[lɛmp]	/læmp/	Easy to reject (3)
Nonword (Non-dom)	[dæsk]	[dɛsk]	/dɛsk/	Difficult to reject (4)



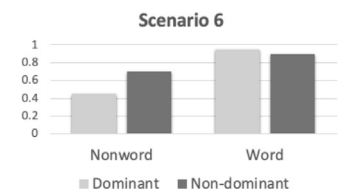
Scenario 5: Precise perceptual representation, neutralized phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[læmp]	/læmp/	Difficult to accept (2)
Nonword (Dom)	[lɛmp]	[lɛmp]	/lɛmp/	Difficult to reject (4)
Nonword (Non-dom)	[dæsk]	[dæsk]	/dɛsk/	Easy to reject (3)



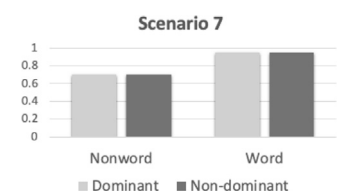
Scenario 6: Precise perceptual representation, ambiguous phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[læmp]	/!mp/	Easy-ish to accept (2)
Nonword (Dom)	[lɛmp]	[lɛmp]	/!mp/	Difficult-ish to reject (4)
Nonword (Non-dom)	[dæsk]	[dæsk]	/dɛsk/	Easy to reject (3)



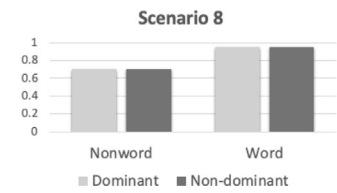
Scenario 7: Precise perceptual representation, “not X” phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[dɛsk]	[dɛsk]	/dɛsk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[læmp]	/!{not ɛ}mp/	Easy to accept (1)
Nonword (Dom)	[lɛmp]	[lɛmp]	/!{not ɛ}mp/	Easy to reject (2)
Nonword (Non-dom)	[dæsk]	[dæsk]	/dɛsk/	Easy to reject (2)



Scenario 8: Precise perceptual representation, precise phonolexical representation

Condition	Input	Percept	LexRep	Ordinal Prediction
Word (Dom)	[desk]	[desk]	/desk/	Easy to accept (1)
Word (Non-dom)	[læmp]	[læmp]	/læmp/	Easy to accept (1)
Nonword (Dom)	[lɛmp]	[lɛmp]	/læmp/	Easy to reject (2)
Nonword (Non-dom)	[dæsk]	[dæsk]	/desk/	Easy to reject (2)



Indeed, there may be multiple types of phonolexical imprecision: the representation of non-dominant categories might be neutralized to the dominant category (e.g., /æ/ encoded as /ɛ/; “neutralized”), ambiguous (/æ/ encoded as /?/, which neither matches nor mismatches [æ] or [ɛ] (“ambiguous”), or differentiated but imprecise (e.g., /æ/ encoded as /not ɛ/; see Hayes-Harb and Masuda, 2008; “not X”). Importantly, these various types of imprecision might produce different predictions regarding LD accuracy patterns. The eight scenarios (along with ordinal accuracy predictions) that result from crossing the four types of phonolexical encoding of non-dominant categories just described with either neutralized (e.g., [æ] perceived as [ɛ]) or precise (e.g., [æ] perceived as [æ]) perceptual representations are elaborated in **Table 2**. In these scenarios, the dominant phoneme is always assumed to be perceived and phonolexically encoded in a distinctive and precise manner. Given a particular set of assumptions about how relative accuracy is computed (detailed in **Table 2**), each scenario produces a prediction regarding the ordinal accuracy associated with words/nonwords and dominant/non-dominant segments. At one extreme (scenario 1), where an individual’s perceptual and phonolexical representations are neutralized to the dominant category, the learner will exhibit a bias towards YES responses symmetrically for both words and nonwords, resulting in highly accurate performance on words and inaccurate performance on nonwords. At the other extreme (scenario 8), where both perceptual and phonolexical representations are distinctive and precise, individuals will accept words and reject nonwords accurately and symmetrically. It is interesting to note this same pattern is observed when the perceptual representations are precise but phonolexical representations are “not X” (scenario 7). Scenarios 2, 3 and 4, where perceptual representations are neutralized and phonolexical representations are ambiguous, “not X,” or precise, respectively, listeners’ performance will be asymmetric with more accurate performance on stimuli containing the dominant phoneme in both nonword and word conditions (we will refer to these as perceptual coding scenarios).²

²It is worth noting that the ordinal accuracy shown in scenario 3 and 4 of **Table 2** differs slightly from the ordinal accuracy predictions made by Darcy et al. (2013). To our knowledge, the prediction associated with Darcy et al.’s (2013) perceptual coding pattern that nonwords containing the dominant category will be more accurate than words containing the non-dominant category is unattested; however, robust word > nonword accuracy patterns are reported across the literature (Sebastián-Gallés et al., 2005; Sebastián-Gallés et al., 2006; Díaz et al., 2012; Darcy et al., 2013; Llompарт and Reinisch, 2019; Melnik and Peperkamp, 2019; Llompарт, 2020, summarized in **Table 1**). For this reason, we make the assumption that words in both dominance conditions will elicit higher accuracy than will nonwords.

Scenarios 5 and 6, where perceptual representations distinguish the two phonemes and phonolexical representations are neutralized or ambiguous, respectively, listeners’ performance will be asymmetric but with an opposite directional pattern of the asymmetry for nonwords and words (these will be collectively referred to as lexical coding scenarios). It is instructive to compute this full set of scenarios, as doing so reinforces the essential difference between response patterns attributable to perceptual processing difficulty (scenarios 2, 3 and 4/ perceptual coding), and those which may be uniquely attributed to challenges at the phonolexical level (scenarios 5 and 6/lexical coding). Doing so further demonstrates that, given the present assumptions, neutralized, ambiguous, and precise phonolexical representations produce the same ordinal accuracy predictions when perceptual representations are neutralized, and when perceptual representations are distinctive and precise, both neutralized and ambiguous phonolexical representations produce the same predictions.

Equipped with these scenarios and their predictions, we now turn to several other studies that have also reported asymmetries in L2 lexical decision performance.³ In understanding the findings of these studies, it is imperative to clarify how nonword stimuli are coded: in some studies they are coded according to the underlying form (e.g., Díaz et al., 2012; *[lɛmp] “lamp” is coded as an /æ/ nonword), while in others they are coded according to the surface form (e.g., Darcy et al., 2013; *[læmp] “lamp” is coded as an [ɛ] nonword). In the following discussion, for ease of interpretation and consistency and whenever possible, we present studies’ findings using the surface form coding scheme so that performance relative to the ordinal accuracy predictions can be evaluated. In addition, given differences between the goals of the studies reviewed below and the present study, and thus analyses focused on different types of patterns, we discuss the findings of these studies in terms of the descriptive patterns (where they are presented or can be inferred) with respect to the LD scenarios.

Melnik and Peperkamp (2019) provide data that is consistent with the predictions of the lexical coding scenarios (scenarios 5 and 6). They investigated the lexical processing of /h/-initial and vowel-initial English words by Intermediate-Advanced L1 French learners of English and native speakers of English (a combination of American, British, and Canadian varieties). Given that French learners of English will often

³A number of other studies have reported asymmetries in child learners (e.g., Simon et al., 2014), adult learners (e.g., Simonchuk and Darcy, 2017; Hayes-Harb and Barrios, 2019), and even early bilinguals (e.g., Amengual, 2016b) using auditory word picture matching tasks with familiar and newly-learned words.

produce English /h/-initial words such as “husband” without the initial /h/ and accept nonwords such as “usband” and “[h]officer” as the real words “husband” and “officer,” the authors hypothesized that L1 French learners of English would exhibit a pattern of lexical decision accuracy consistent with them having what they called “fuzzy” phonolexical representations of /h/-initial English words (that is, they would exhibit more accurate performance for vowel-initial words than for [h]-initial words, but less accurate performance for vowel-initial nonwords than [h]-initial nonwords). This is indeed what they observed. Unexpectedly, they also observed a difference in word performance for the native English speakers similar to that reported for the learners, with [h]-initial words less accurate than vowel-initial words. However, no difference was observed for nonwords. Melnik and Peperkamp’s (2019) LD findings are summarized in **Table 1**.

Patterns of LD asymmetries consistent with difficulty at the level of perceptual coding (scenarios 2, 3, and 4) have also been reported. Díaz et al. (2012) examined the lexical processing of the /æ/-/ɛ/ contrast by native Dutch late-learners of English (self-rated “high” proficiency) and a control group of native speakers of British English. Participants completed a lexical decision task involving monosyllabic English words containing /ɛ/ and /æ/, and an equal number of nonwords created by substituting /ɛ/ for /æ/, and vice versa (which they refer to as /ɛ/-type stimuli and /æ/-type, respectively). Analyses of A’ scores revealed that native English speakers demonstrated greater sensitivity to the contrast than Dutch participants and that both groups were more sensitive to what they call /æ/-type than /ɛ/-type stimuli. However, the use of signal detection measures, such as A’, makes it impossible to assess whether the interaction for surface segment by lexical status predicted the lexical coding scenarios was observed. Nonetheless, the mean proportion correct data presented in figure 3 of Díaz et al. (2012); recoded for surface as opposed to underlying segment, and summarized in **Table 1**) suggests that both groups of participants were more likely to correctly accept words containing [ɛ] “desk” than words containing [æ] “lamp” and more likely to incorrectly accept nonwords containing [æ] “d[æ]sk” than nonwords containing [ɛ] “l[ɛ]mp.” While not the focus of this study, the descriptive pattern of performance of the Dutch (and the native English) speakers reported by Díaz et al. (2012) is compatible with the phonetic coding scenarios, and corroborates findings from eye tracking studies involving this contrast and learner population (discussed above).

Several studies involving Spanish-Catalan bilinguals have examined the robustness of Catalan-dominant and Spanish-dominant bilinguals’ lexical encoding of the Catalan-specific /e/-/ɛ/ vowel contrast, which is known to be particularly difficult for the Spanish-dominant group (Sebastián-Gallés and Baus, 2005; Sebastián-Gallés et al., 2005; Sebastián-Gallés et al., 2006). These studies have employed lexical decision tasks involving words with /e/ or /ɛ/ and nonwords counterparts created by substituting the other member of the contrast (e.g., [finestrə] “window”/*[finestrə] and [gəʎedə] “bucket”/*[gəʎedə]). In these experiments participants were warned that nonwords would involve a single vowel change. The nonword acceptance for [e] and [ɛ] were high in both studies, particularly

for the Spanish-dominant group. Moreover, mean proportion correct showed an asymmetrical pattern for nonwords in both groups. Sebastián-Gallés et al. (2006) report lower proportion correct for [e] than [ɛ] nonwords in addition to high accuracy on words for both groups of bilinguals. Sebastián-Gallés et al. (2005) report very similar findings in their figure 1 (see **Table 1** for a summary of Sebastián-Gallés et al. (2005) and Sebastián-Gallés et al. (2006)). In both studies, the behavioral data was coded for the underlying segment (rather than surface segment) and statistical analyses were conducted with A’ scores as the dependent variable. As a result, it is not possible to know whether the asymmetries they observed resulted in a significant interaction of lexical status and surface segment. However, assuming that the nonword difference was robust, and that [e] is the dominant category and [ɛ] is the non-dominant category for Spanish-dominant Spanish-Catalan bilinguals whose native language has /e/ but lacks /ɛ/, this direction of the nonword effect (more accurate nonword performance for [ɛ] than [e]) would be more compatible with the lexical coding than perceptual coding scenarios. The authors attribute effects in Catalan-dominant bilinguals to exposure to variable input due to experience with Spanish-accented Catalan in the bilingual speech community where the research has been conducted.

Llompert (2020) studied the lexical decision behavior of two groups of L1 German learners of English (Intermediate and Advanced) on the /æ/-/ɛ/ contrast, with nonwords derived from real words by swapping the segment of interest (e.g., lemon/*l[æ]mon and dragon/*dr[ɛ]gon). Performance for word stimuli (not presented separately by surface segment in the manuscript) was at ceiling for the two groups, and analyses focused on nonword performance. As expected, both groups of learners were more accurate in their reject of nonwords for filler than for the test contrast and the advanced group outperformed the intermediate group in nonword performance for the /æ/-/ɛ/ contrast. While the study’s focus was on relating categorization and vocabulary knowledge to lexical decision performance, the authors did provide data and an additional analysis of these nonword effects broken down by surface segment. We are told that “both groups showed numerically lower accuracies for items in which /æ/ was mispronounced as [ɛ] (e.g., *dr[ɛ]gon; Advanced: 51.03% correct (SD = 50.05); Intermediate: 25.44% (SD = 43.59)) than when the substitution pattern was the opposite (e.g., *l[æ]mon; Advanced: 66.15% correct (SD = 47.37); Intermediate: 47.74% (SD = 49.65)).” (p. 6). Assuming that /ɛ/ is the dominant category for German learners of English (e.g., Llompert and Reinisch, 2019), this result would appear to be most compatible with the lexical coding scenarios. This pattern of performance in German learners of English is consistent with LD data presented in Llompert and Reinisch (2019), though that study reported only d’ scores. The Llompert and Reinisch (2019) and Llompert (2020) LD findings are summarized in **Table 1**.

It has been generally assumed the representation of native phonemes will lack the fuzziness of novel L2 phonemes’ representations, and that native speakers should therefore not exhibit the asymmetries that have been associated with L2 learners’ lexical decision performance (e.g., Darcy et al., 2013).

This assumption appears to be implied in studies lacking control groups of native speakers (e.g., Hayes-Harb and Barrios, 2019; Llompart, 2020; Llompart and Reinisch, 2019; Llompart and Reinisch, 2020). However, inspection of reported response patterns reveals that native speakers quite often show asymmetric lexical decision performance (Sebastián-Gallés et al., 2005; Sebastián-Gallés et al., 2006; Díaz et al., 2012; Darcy et al., 2013; Melnik and Peperkamp, 2019). Recognizing the prevalence of such lexical decision asymmetries among native speakers, Melnik and Peperkamp (2019) note that it is “important to always compare the learners’ performance to that of native speakers, such as to clearly identify asymmetries that are specific to L2 processing” (p. EL17).

Research employing auditory LD tasks to investigate L2 phonological processing has varied in the ways in which the data and patterns are presented, in addition to the ways in which various asymmetries are interpreted with respect to the unique influence of perceptual and/or phonological representations on L2 learners’ performance. The first goal of the present research is thus to report LD data for multiple L2 contrasts across multiple L1 groups (representing previously unstudied L1-L2 combinations) and to evaluate the findings (taking into account patterns of performance on both word and nonword stimuli) with respect to the predictions of the scenarios described above. Research in this area has also varied in whether or not learner performance is compared to performance by native speakers. To the extent that asymmetries in LD performance by learners are interpreted as evidence of nonnative-like perceptual and/or phonological representations, they must be considered in relation to performance by native speakers (see also Melnik and Peperkamp, 2019). The second goal of this study is to document potential native speaker LD asymmetries to allow for comparison of performance by native speakers and L2 learners, and the inclusion of multiple contrasts and multiple groups of L2 learners allows more opportunities for this comparison than does a more targeted and less exploratory study. The third goal of this work is to add to the representation of individuals who have emigrated to L2-dominant settings (in this case, late learners of English in the United States) in the literature (see Extra and Verhoeven, 2011; Paradis et al., 2020). With a few exceptions (e.g., Darcy and Thomas, 2019), studies on this particular topic have focused on instructed learners in “foreign language” settings (e.g., Darcy et al., 2013; Cook and Gor, 2015) or early bilinguals in a bilingual speech community (e.g., Sebastián-Gallés et al., 2005; Sebastián-Gallés et al., 2006; Amengual, 2016a; Amengual, 2016b).

We selected two segmental contrasts known to be difficult for L2 learners of English from a variety of native language backgrounds: the /æ/-/ɛ/ vowel contrast and the /l/-/ɹ/ liquid contrast (e.g., Flege et al., 1997; Aoyama et al., 2004). Moreover, Cutler (2005) notes that failing to maintain these particular contrasts in English can lead to a substantial increase in lexical competition. The phonological processing of English /æ/ and /ɛ/ has been extensively studied in the context of Dutch and German (Weber and Cutler, 2004; Escudero et al., 2008; Díaz et al., 2012); however, this contrast is known to pose a challenge for learners of English from a wide range of language backgrounds, including native speakers of Mandarin and Korean. Native speakers of Mandarin have been shown to experience difficulty identifying, discriminating, and producing English /æ/ and /ɛ/ (Wang, 1997;

Chen et al., 2001; Jia et al., 2006), neither of which is nominally present in the Mandarin vowel inventory. Native speakers of Korean also experience difficulty perceiving and producing the English /æ/-/ɛ/ contrast (Tsukada et al., 2005; Kim, 2010; Hong, 2012). Unlike Mandarin, however, the Korean vowel inventory nominally contains the vowel /ɛ/, and some evidence suggests that /ɛ/ might be expected to behave like the dominant vowel of the two. For example, Yang (1996) found that Korean /ɛ/ is acoustically more similar to English /ɛ/ than to English /æ/, and Flege et al. (1997) reported that native Korean speakers produce English /ɛ/ more accurately than English /æ/. However, other studies have found higher perception and production accuracy for English /æ/ than /ɛ/ (Tsukada et al., 2005; Cho and Jeong, 2013). With respect to predictions regarding the dominance status of the two vowels, we know of no data that points to the dominance of /ɛ/ or /æ/ for the L1 Mandarin speakers, and for L1 Korean speakers, evidence regarding dominance is contradictory.

Mandarin is characterized as having a lateral approximant phoneme /l/, but not /ɹ/ (e.g., Brown, 2000). Nonetheless, Brown (1998) demonstrated that native Mandarin speakers who were late learners of English and living in North America exhibited near-ceiling accuracy on both AX discrimination and forced-choice picture selection (e.g., hear “rake”, choose between pictures of a “lake” or a “rake”) tasks, with highly accurate performance maintained across onset, cluster, and coda positions. Brown (1998) notes that Mandarin /l/ does not vary allophonically between [l] and [ɹ], which may reduce the likelihood that native Mandarin speakers neutralize English /l/ and /ɹ/ to a single category. Korean has a singleton liquid phoneme--sometimes characterized as /l/ (e.g., Brown, 2000)--that is realized as [r] (initially) and [l] (elsewhere). On the one hand, the status of these two phones in Korean as conditioned variants of the same phoneme might make the English /ɹ/ - /l/ contrast difficult for native speakers of Korean. Consistent with this prediction, Brown (2000) demonstrated that native Korean speakers less accurately perceived and lexically encoded the English /l/- /ɹ/ contrast than did native Mandarin speakers (who do not have the allophonic experience with [l] and [ɹ] that would encourage neutralization). However, the geminate liquid in Korean is realized as [ll] intervocalically, resulting in a [r] - [ll] (singleton-geminate) contrast intervocalically, which Kim (2007) characterizes as a latent /ɹ/ - /l/ contrast. And indeed, native speakers of Korean have exhibited fairly accurate identification of English /l/ and /ɹ/, in intervocalic position (Ingram and Park, 1998; Hazan et al., 2006). Using a cross-language identification and category goodness task, Schmidt (1996) showed that native Korean speakers identify both English /l/ and /ɹ/ as Korean /l/ in initial position, but that English /l/ is more similar to Korean /l/ than is English /ɹ/ (see also Park, 2013 for discussion of the “new” versus “similar” status of English /l/ and /ɹ/ for native speakers of Korean). Concerning the dominance status of English /l/ and /ɹ/ for these two groups of learners, the limited available evidence of similarities between English and both Mandarin and Korean /l/ suggest that English /l/ may be dominant for native Mandarin and Korean speakers.

TABLE 3 | Summary of participant characteristics.

	Mean age in years (SD; range)	Sex	Mean LOR in months (SD); range	Self-rated English Listening	(Additional) L2s	Counter- balance List
Native English (<i>n</i> = 56) 18 P, 38 SP	26.7 (10.7; 18–60)	20 male 35 female 1 non-binary	n/a	n/a	23 reported one or more L2s Arabic(1), ASL(1), Cantonese(1), Danish(1), Farsi(1), French(3), German (2), Italian(1), Japanese(4), Laotian(1), Mandarin (3), Portuguese(1), Spanish (14), Swedish (1), Thai (1)	28 List 1 28 List 2
Native Mandarin (<i>n</i> = 18) 6 P, 12 WoM	31.8 (8.3; 19–51)	8 male 10 female	97.4 (57.0; 36–267); missing = 2	3 fair 15 good	8 reported additional L2s Cantonese(1), French(2), German(2), Japanese(4), Korean (2), Thai (1)	9 List 1 9 List 2
Native Korean (<i>n</i> = 24) 4 P, 20 WoM	44.0 (9.1; 28–62)	4 male 20 female	175.3 (125.6; 6–480)	13 fair 11 good	5 reported additional L2s Japanese(4), Mandarin (1)	11 List 1 13 List 2

P, Prolific; *WoM*, Word of mouth; *SP*, Linguistics study pool; *LOR*, length of residence in the United States.

TABLE 4 | Lexical characteristics of the /æ/-/ɛ/ and /l/-/ɹ/ stimulus sets.

	Words			Nonwords	
	Mean (SD) Subtlex frequency	Mean (SD) # neighbors	Mean (SD) neighbor frequency	Mean (SD) # neighbors	Mean (SD) neighbor frequency
Underlying /æ/	25.33 (11.77)	4.75 (3.19)	5.25 (1.91)	3.67 (3.77)	21.39 (20.20)
Underlying /ɛ/	25.33 (20.89)	4.00 (4.112)	7.08 (9.45)	4.00 (4.00)	23.72 (15.97)
/æ/-/ɛ/ comparison	$t(22) = 0.000, p = 1.000$	$t(22) = 0.499, p = 0.623$	$t(22) = -0.658, p = 0.517$	$t(22) = -0.210, p = 0.836$	$t(22) = -0.313, p = 0.757$
Underlying /l/	25.83 (23.87)	3.42 (2.61)	7.08 (8.28)	2.00 (1.28)	21.64 (24.43)
Underlying /ɹ/	17.83 (8.32)	3.58 (3.12)	9.17 (10.18)	4.08 (3.99)	19.13 (17.36)
/l/-/ɹ/ comparison	$t(22) = 1.096, p = 0.285$	$t(22) = -0.142, p = 0.888$	$t(22) = -0.550, p = 0.588$	$t(22) = -1.723, p = 0.099$	$t(22) = 0.291, p = 0.774$
Filler	22.25 (8.58)	5.08 (2.84)	9.26 (8.69)	3.75 (2.53)	13.79 (6.01)

METHODS

Participants

There are three groups of participants in this approximately 15 min online study: 18 native Mandarin and 24 native Korean L2 learners of English, and a control group of 56 native speakers of English, with participants randomly assigned to two counterbalancing conditions. We pursued three avenues for participant recruitment: our department's participant pool, Prolific (www.prolific.co), and word-of-mouth. The participant pool connects students enrolled in linguistics courses to studies for course credit. Prolific connects participants to paid research studies. Due to limitations associated with Prolific's participant screening and recruitment policies, data from several participants recruited via Prolific was discarded because their responses to the post-task questionnaire indicated that they did not meet the study's inclusionary criteria. Participants recruited via Prolific were paid between \$4.00 and \$5.00 USD, with variation in compensation resulting from author experimentation with the system. Word-of-mouth recruitment involved emailing participants from previous studies who had opted into our recruitment list, in addition to asking colleagues to distribute a recruitment message to potential participants on our behalf. These participants were compensated with a \$5 Amazon gift card.

The control group of 56 native speakers of English was recruited through the Linguistics study pool and Prolific and self-identified as native/first language speakers of English only.

Data from an additional three native English speakers recruited via Prolific was excluded because they indicated that they were not familiar with some of the English words used in the study. Given the potential for systematic differences between participants recruited via the multiple avenues, we attempted to balance where participants were recruited from and counterbalanced list assignment. The native Mandarin and Korean speakers were born in China and Korea and considered Mandarin and Korean, respectively, to be their only first and native languages. They identified English a second language, and in order to ensure substantial exposure to North American dialects of English, were living in the United States at the time of the study, and had arrived in the United States no earlier than 12 years of age. Additionally, participants selected for inclusion in this study self-rated their English listening ability as “fair” or “good,” on the four-point scale labeled “poor - fair - good - near-native.” They exhibited a range of ages of English language acquisition, age of arrival in the United States, and length of residence in the United States (see **Table 3** for participant characteristics).

Materials

Native Mandarin and native Korean participants who were recruited by word of mouth (rather than a participant management platform) completed a pre-task questionnaire. They were asked a series of questions to confirm that they met the inclusionary criteria (native language, countries of birth and current residence, and status of English as a second

language), and also took a brief Mandarin or Korean vocabulary test serving as a screening to reduce participation by individuals who are not speakers of these languages.

For the auditory lexical decision task, three sets of English word-nonword pairs were created by first querying the MRC Psycholinguistics Database (Coltheart, 1981) using the following criteria: two-syllable English words beginning with CV, a stress-unstress pattern, a sublex frequency between 10 and 100. We then removed all proper nouns. From the resulting set, we selected words with initial /l/ or /ɹ/ (the /l/-/ɹ/ set) or with /æ/ or /ɛ/ in the first syllable (the /æ/-/ɛ/ set), and removed from these sets all words whose /l/-/ɹ/ or /æ/-/ɛ/ counterparts were real words (e.g., “/l/iver” and “/ɹ/iver”). In addition, we selected a separate set of candidate filler words that met all of the same criteria and created nonword counterparts for these words (e.g., “/k/otton”-“/p/otton”). We then assessed the neighborhood density, neighborhood frequency of all resulting words and nonwords, in addition to sublex per million frequency of the words, and crafted the three stimulus sets, attempting to balance these lexical characteristics across the sets and with the aim of including only nouns (due to a limited number of word options, some monomorphemic verbs and -y adjectives are included in the sets). Because our stimuli met very specific criteria and we wanted to ensure there was a sufficient number of items for each category, there was overlap between some of the test and filler words (e.g., “rabbit” was used both as an [æ] and an [ɹ] word). Care was taken to ensure that multiple tokens of the same word were spoken by different talkers on a particular list. The lexical characteristics of the resulting set are presented in **Table 4**. Comparison of the lexical characteristics of /ɹ/ vs. /l/ and /æ/ vs. /ɛ/ stimuli reveal no significant differences. The complete list of stimuli provided in **Supplementary Table 1**, and the auditory lexical decision task materials are available at <https://osf.io/9mnav/>.

The word and nonword stimuli were recorded by two female native speakers of American English. Three recordings of each were made, and Praat was used to identify and extract each auditory stimulus. The first production was chosen for presentation in the study unless it contained artifacts. Stimuli were scaled such that their average intensity was 65 dB. A Praat script provided formant values at the segment (vowel or liquid) midpoints (Lennes, 2003). Analysis of the acoustics of the stimuli revealed greater variability in F1/F2 for /æ/-/ɛ/ and in F3 for /l/-/ɹ/ in nonwords than words, leading to potentially confounding differences in stimulus acoustics. We thus eliminated responses to all tokens (words and nonwords) that were more than two standard deviations greater or less than the F1 or F2 means (/æ/-/ɛ/) or the F3 mean (/l/-/ɹ/) of the word tokens, separately for each talker. This resulted in the exclusion of participants' responses to six word tokens and 14 nonword tokens from the analyses (see **Supplementary Table 2**) and amounted to the exclusion of 8.3% of the data from each group (560 of 6,780 observations, 180 of 2,160 observations, and 238 of 2,880 observations, for the native English, Mandarin, and Korean speakers, respectively).

In a post-task questionnaire, participants were asked to provide basic demographic information (e.g., age, sex) in

addition to information about their native and second language experience, location of birth and current residence. The native Mandarin and Korean speakers were additionally asked to detail their English language experience. They were asked to self-assess their own speaking, listening, reading, and writing ability in English on a four-point scale (poor to near-native). In addition, they were asked to indicate which of the stimulus words were unfamiliar to them by checking a box next to the word.

Procedures

All parts of the study were conducted online via computer and headphones, with participants using their own equipment in locations of their choice. Participants first were presented with the informed consent text, and registered their consent to participate by pressing a button labeled “Agree.” This part of the study was conducted in Qualtrics. Participants were next directed to Pavlovia, where the listening task was hosted. The listening task involved an auditory lexical decision. Participants were told that they would hear many words, some of which were real English words and some not. Their task was to decide as quickly and accurately as possible whether the word they heard was a real English word. Participants were instructed to press the “y” key on the keyboard to indicate a YES response, and a “n” key to indicate a NO response. The task was self-paced, and a button press was required to advance. The next trial was presented 1s after the participant's response was made. The listening task consisted of a total of 120 test trials. Test trials were preceded by 4 practice trials (disk, d[u]sk, simple, s[u]mple) without feedback. The entire listening task took approximately 10 min. Upon completion of the listening task, participants were directed back to the Qualtrics platform to complete the post-task questionnaire, and then they were either directed back to the appropriate participant management platform for credit (departmental participant pool) or payment (Prolific), or asked to provide their name and email for the purpose of sending a gift card.

RESULTS

The data and analysis code are available at <https://osf.io/psxw9/>. We took several steps to ensure the quality of the data. In addition to excluding responses to some stimuli based on their acoustic properties (see above), we also considered participants' self-reported familiarity with the stimulus words. Because participants' familiarity with the words is crucial to our ability to interpret their lexical decision responses, we excluded responses from each participant for all words and associated nonwords that were unfamiliar for that individual. Words identified as unfamiliar by one or more Mandarin native speakers included lousy(3), medal(1), pattern(1), radar(2), ransom(7), reckon(7), rhythm(1), ruin(1), tunnel(1), and warrant(2). This amounted to the exclusion of an additional 66 observations (3% of the data). One or more Korean native speakers indicated the following words were unfamiliar: battle(1), lecture(1), lobby(1), legend(1), rabbit(1), ransom(3), reckon(10), rotten(1), rubber(1), supper(2), and

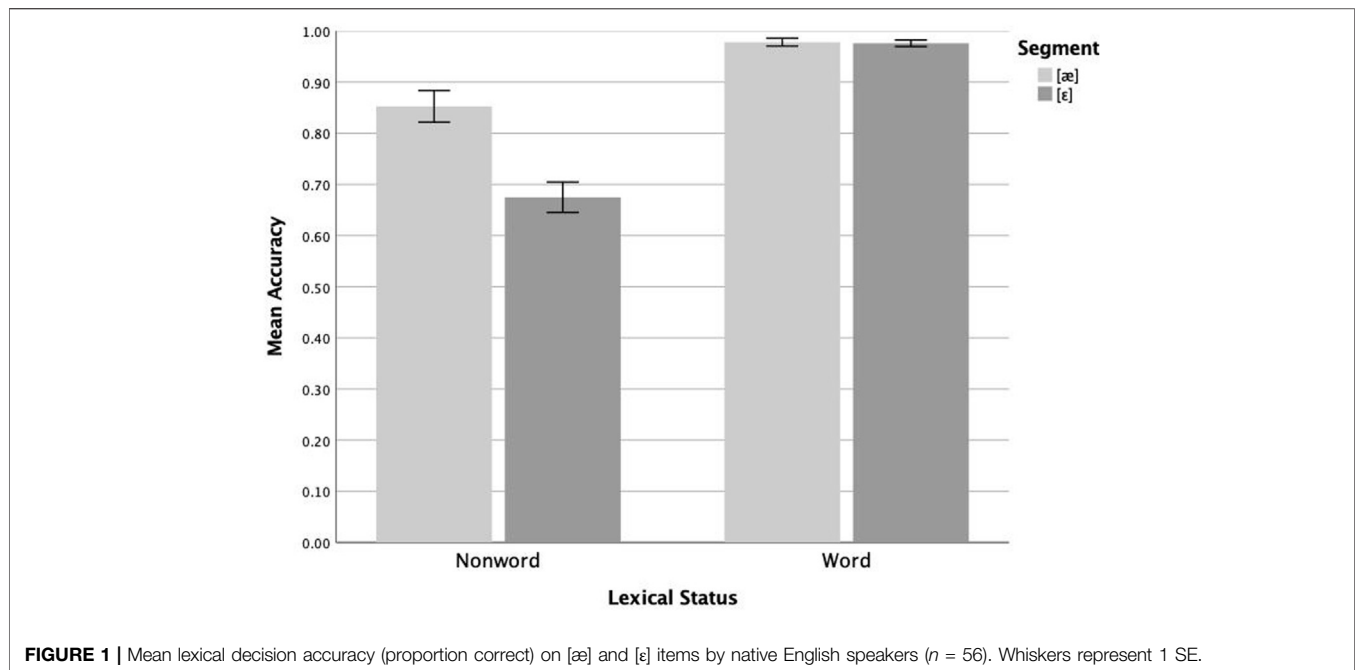


FIGURE 1 | Mean lexical decision accuracy (proportion correct) on [æ] and [ɛ] items by native English speakers ($n = 56$). Whiskers represent 1 SE.

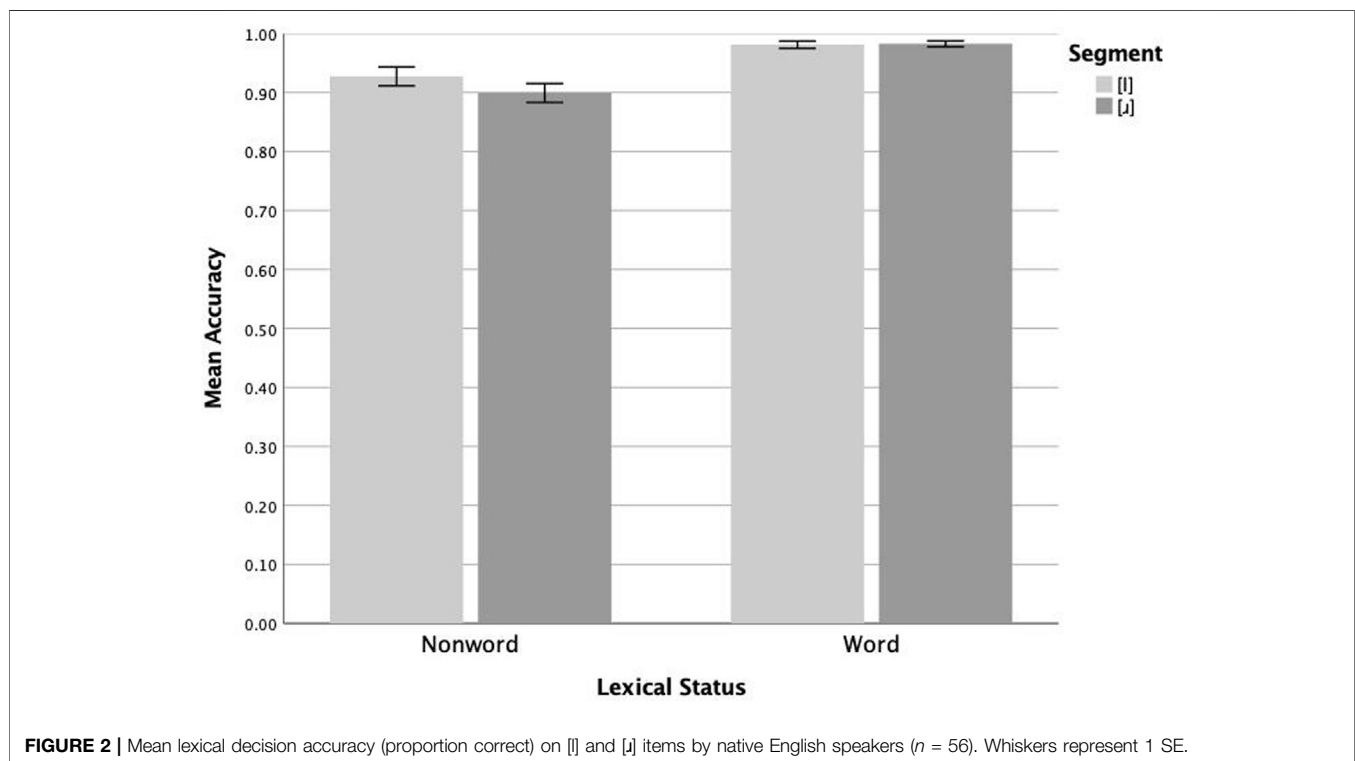
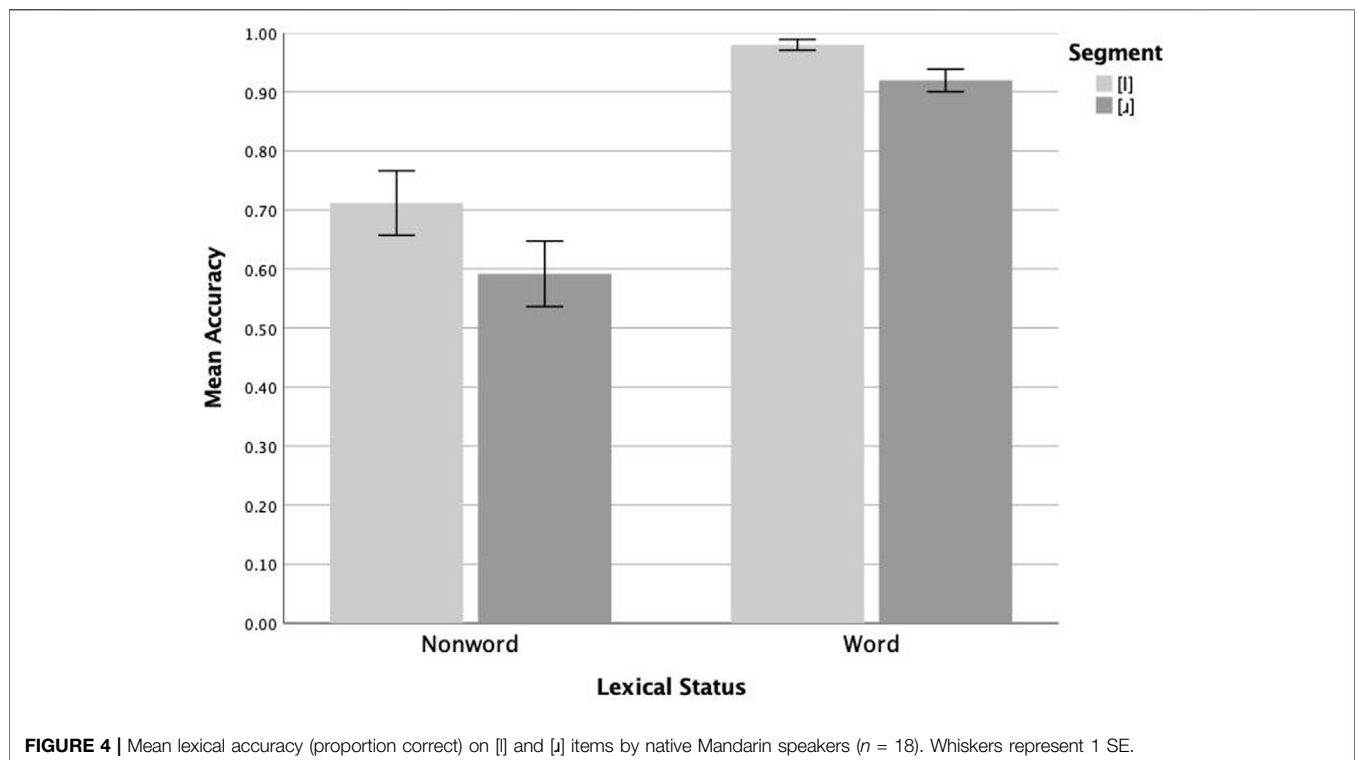
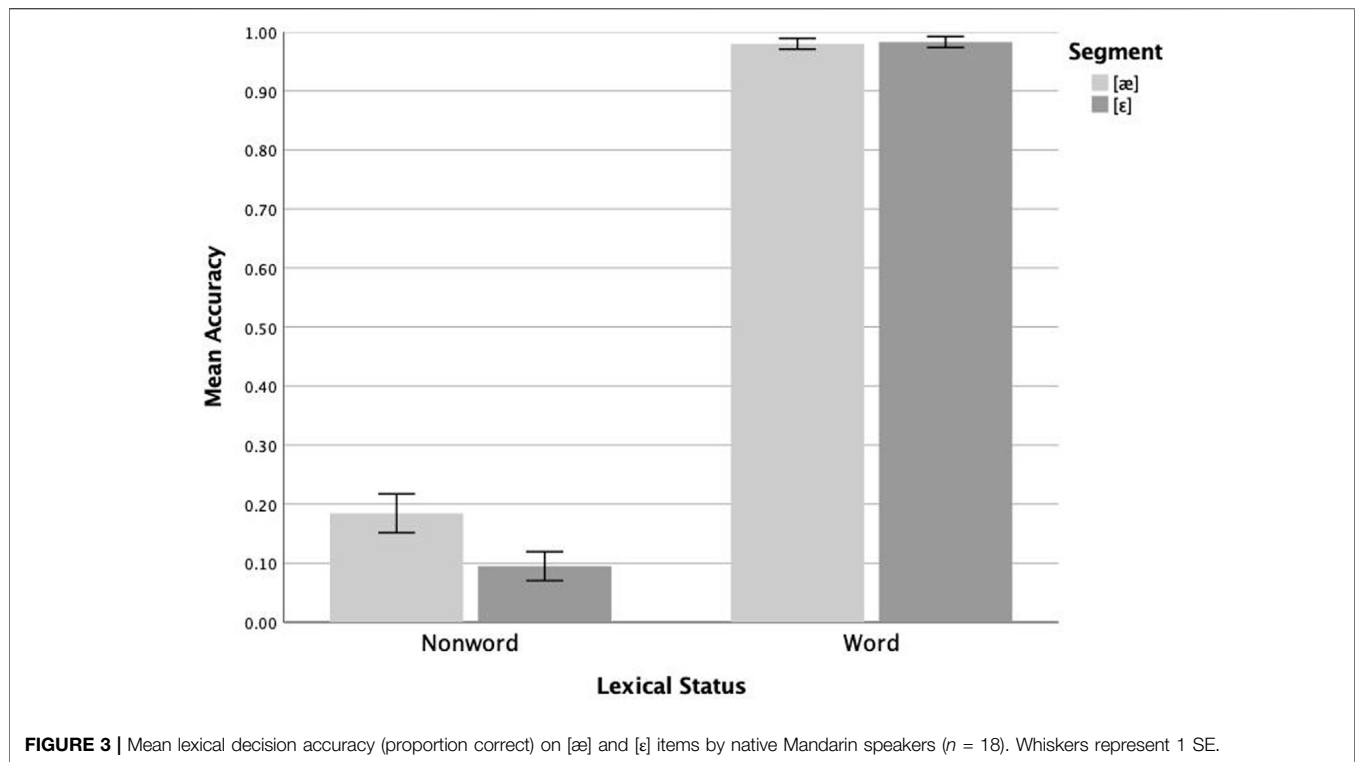


FIGURE 2 | Mean lexical decision accuracy (proportion correct) on [l] and [ɹ] items by native English speakers ($n = 56$). Whiskers represent 1 SE.

warrant(2). Exclusion of these words and their corresponding nonwords resulted in the exclusions of 78 of 2,880 observations (2.7% of the data) for the Korean native speakers.

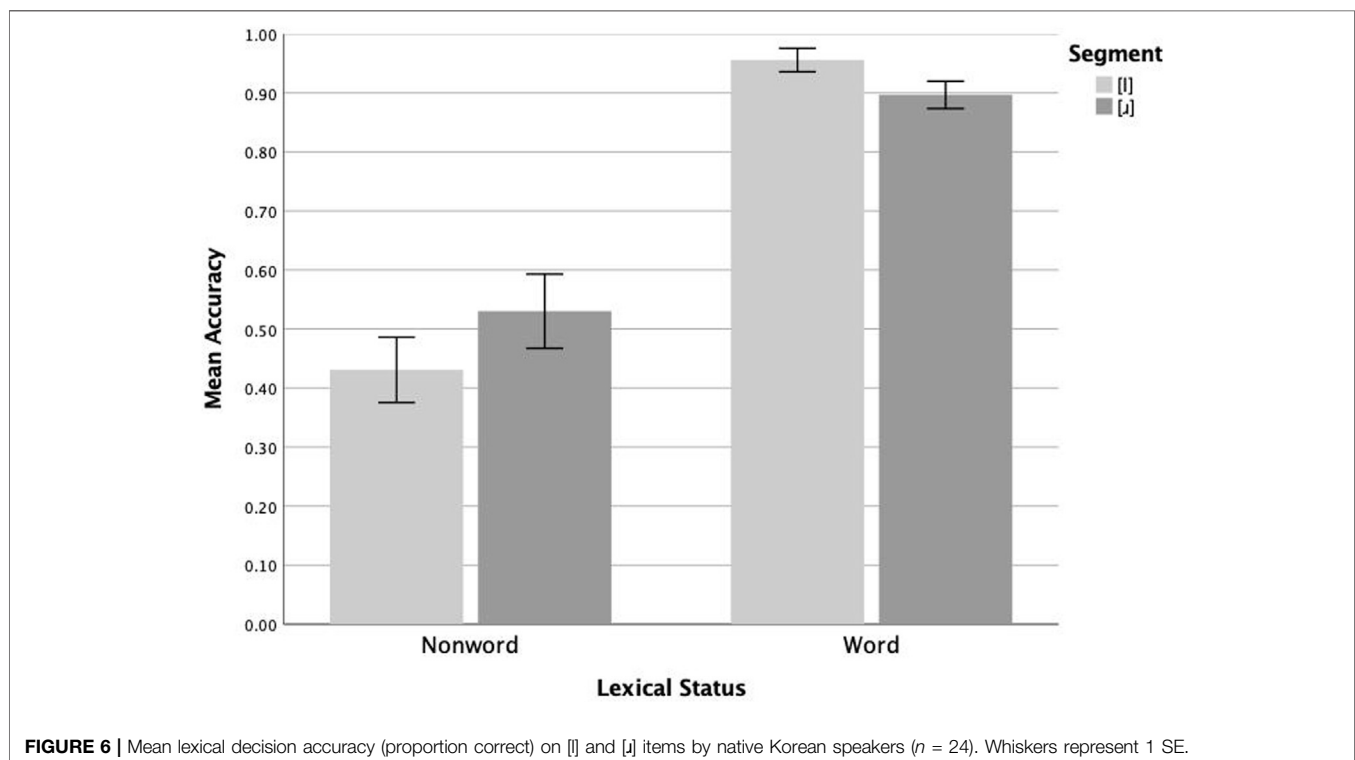
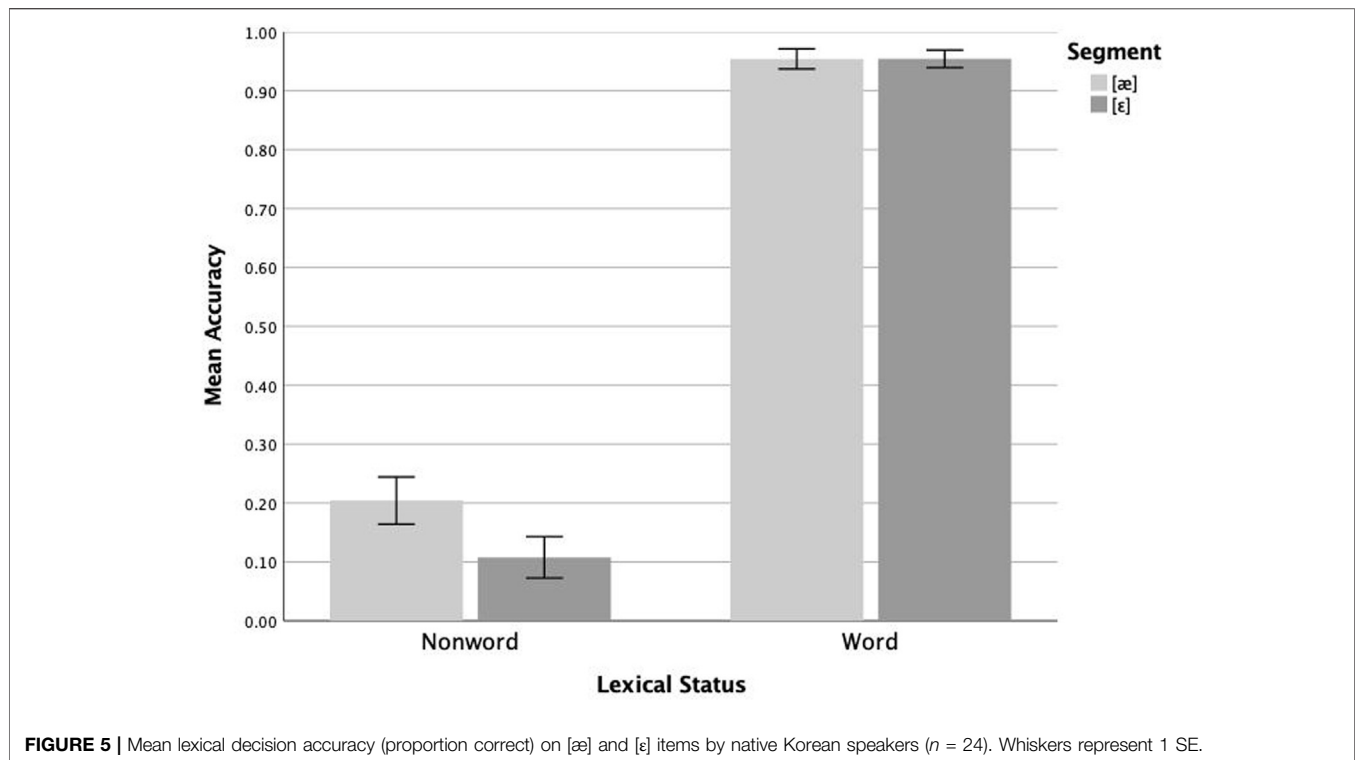
To further control for potentially problematic stimuli, following the exclusion of tokens with outlier acoustics and unfamiliar words and corresponding nonwords just described,

we also excluded tokens where the mean proportion correct performance was greater or less than 2 SD of their group mean for a given target segment ([æ], [ɛ], [l], [ɹ], or filler) and lexical status (word, nonword). This resulted in the exclusion of twelve tokens for native English speakers (336 observations), 10 tokens for native Mandarin speakers (99 observations), and 12



tokens for native Korean speakers (140 observations; the excluded tokens are listed in **Supplementary Table 2**). Following all exclusions, data analysis was conducted on the remaining 5,824

observations provided by 56 native speakers of English (86.7%), 1827 observations from 18 native speakers of Mandarin (84.5%), and 2,435 observations from 24 native speakers of Korean (84.5%).



Native English Speakers

Proportion correct responses (YES to words and NO to nonwords) of the native English speakers were submitted to

two separate repeated measures ANOVAs. Because filler items were designed only to distract, involved a variety of segmental contrasts, and were not associated with any predictions regarding

relative difficulty, they are not considered in these analyses (mean proportion correct responses by native English speakers to filler nonwords was 0.913 and to filler words was 0.986). In the first repeated measures ANOVA, surface segment (two levels: [æ], [ɛ]) and lexical status (two levels: nonword, word) were within-subjects independent variables and proportion correct was the dependent variable. The main effect of surface segment was significant ($F(1,55) = 35.979, p < 0.005; \eta_p^2 = 0.395$), as were the main effect of lexical status ($F(1,55) = 53.466, p < 0.005; \eta_p^2 = 0.493$) and the interaction of the two ($F(1,55) = 34.399, p < 0.005, \eta_p^2 = 0.395$). Follow-up pairwise analyses revealed that native English-speaking participants responded correctly more often to [æ] (mean = 0.853) than to [ɛ] (mean = 0.675) nonwords ($F(1,55) = 39.179, p < 0.005, \eta_p^2 = 0.416$), but no significant difference in responses to [æ] (mean = 0.978) and [ɛ] (mean = 0.976) words ($F(1,55) = 0.047, p = 0.828, \eta_p^2 = 0.001$). These results are plotted in **Figure 1**. The pattern observed here does not readily match any of the predicted scenarios, perhaps due to the near-ceiling performance on word stimuli. However, we do observe an asymmetry in performance on nonwords, with higher accuracy for [æ] nonwords (those derived from words containing /ɛ/) than for [ɛ] nonwords (derived from words containing /æ/).

For the /l/-/ɹ/ data, a second repeated measures ANOVA with surface segment and lexical status as within-subjects variables and proportion correct as the dependent variable revealed a non-significant main effect of surface segment ($F(1,55) = 2.192, p = 0.144, \eta_p^2 = 0.038$), a significant main effect of lexical status ($F(1,55) = 24.973, p < 0.005, \eta_p^2 = 0.312$), and a non-significant interaction of the two ($F(1,55) = 2.512, p = 0.119, \eta_p^2 = 0.044$). Follow-up pairwise analyses revealed that native English-speaking participants did not differ in their accuracy for [l] (mean = 0.928) and [ɹ] (mean = 0.899) nonwords ($F(1,55) = 2.680, p = 0.107, \eta_p^2 = 0.046$), or for [l] (mean = 0.981) and [ɹ] (mean = 0.983) words ($F(1,55) = 0.051, p = 0.822, \eta_p^2 = 0.001$). These results are plotted in **Figure 2**. This response pattern, with near-ceiling performance for words and slightly lower but symmetric accuracy for nonwords, is consistent with the predictions of scenario 8, where participants benefit from perceptual and phonolexical representations that are differentiated and precise (or the functionally equivalent “not X”; scenario 7).

Native Mandarin Speakers

Mean proportion correct for native Mandarin speakers overall was 0.698. Mean proportion correct for filler words was 0.926 and filler nonwords 0.612. Analysis of the /æ/-/ɛ/ data from the native Mandarin speakers revealed significant main effects of surface segment ($F(1,17) = 6.171, p = 0.024, \eta_p^2 = 0.266$) and lexical status ($F(1,17) = 1,206.335, p < 0.005, \eta_p^2 = 0.986$), and a significant interaction of the two ($F(1,17) = 8.035, p = 0.011, \eta_p^2 = 0.321$). Follow-up pairwise analyses revealed that these participants responded correctly more often to [æ] (mean = 0.184) than to [ɛ] (mean = 0.095) nonwords ($F(1,17) = 7.820, p = 0.012, \eta_p^2 = 0.315$), with no significant difference in responses to [æ] (mean = 0.980) and [ɛ] (mean = 0.983) words ($F(1,17) = 0.080, p = 0.781, \eta_p^2 = 0.005$). These results are plotted in **Figure 3**. Like the native English speakers, the native Mandarin speakers exhibited highly

accurate performance on [æ] and [ɛ] words. However, their relatively low accuracy on the nonword stimuli points to a YES bias in their responses. In addition, like the native English speakers, they show an asymmetry in responses to these nonwords with more accurate performance on [æ] than [ɛ] nonwords.

Analysis of the /l/-/ɹ/ data from the native Mandarin speakers revealed significant main effects of surface segment ($F(1,17) = 13.586, p = 0.002, \eta_p^2 = 0.444$) and lexical status ($F(1,17) = 37.385, p < 0.005, \eta_p^2 = 0.687$), and a non-significant interaction of the two ($F(1,17) = 1.221, p = 0.285, \eta_p^2 = 0.067$). Follow-up pairwise analyses revealed that these participants responded correctly significantly more often to [l] (mean = 0.712) than to [ɹ] (mean = 0.592) nonwords ($F(1,17) = 6.603, p = 0.020, \eta_p^2 = 0.280$), and significantly more often to [l] (mean = 0.980) than to [ɹ] (mean = 0.920) words ($F(1,17) = 7.743, p = 0.013, \eta_p^2 = 0.313$). These results are plotted in **Figure 4**. In contrast to the native English speakers, the native Mandarin speakers exhibited asymmetries for both words and nonwords with higher accuracy on [l] than [ɹ] stimuli. This pattern of responses is consistent with scenarios 2, 3 and 4, where a neutralized perceptual representation of [ɹ] as [l] contacts phonolexical representations that are ambiguous (scenario 2), ‘not X’ (scenario 3) or precise (scenario 4).

Native Korean Speakers

Mean proportion correct for native Korean speakers overall was 0.669. Mean proportion correct for filler words was 0.955 and filler nonwords 0.705. Analysis of the /æ/-/ɛ/ data from the native Korean speakers revealed significant main effects of surface segment ($F(1,23) = 7.182, p = 0.013, \eta_p^2 = 0.238$) and lexical status ($F(1,23) = 464.043, p < 0.005, \eta_p^2 = 0.953$), and a marginal interaction of the two ($F(1,23) = 3.949, p = 0.059, \eta_p^2 = 0.147$). Follow-up pairwise analyses reveal that these participants responded correctly more often to [æ] (mean = 0.204) than to [ɛ] (mean = 0.108) nonwords ($F(1,23) = 6.651, p = 0.017, \eta_p^2 = 0.224$), with no significant difference in responses to [æ] (mean = 0.954) and [ɛ] (mean = 0.954) words ($F(1,23) < 0.005, p = 0.994, \eta_p^2 < 0.005$). These results are plotted in **Figure 5**. Like the native English and native Mandarin speakers, the native Korean speakers exhibited highly accurate performance on the [æ] and [ɛ] word stimuli, as well as an asymmetry in responses to these nonwords with more accurate performance on [æ] than [ɛ] nonwords. Like the Mandarin speakers, the Korean speakers exhibited a bias towards YES responses.

Analysis of the /l/-/ɹ/ data from the native Korean revealed a non-significant main effect of surface segment ($F(1,23) = 0.679, p = 0.418, \eta_p^2 = 0.029$) and a significant main effect of lexical status ($F(1,23) = 58.740, p < 0.005, \eta_p^2 = 0.719$), and a significant interaction of the two ($F(1,23) = 11.961, p = 0.002, \eta_p^2 = 0.342$). Follow-up pairwise analyses reveal that these participants responded correctly significantly more often to [ɹ] (mean = 0.530) than to [l] (mean = 0.431) nonwords ($F(1,21) = 5.512, p = 0.028, \eta_p^2 = 0.193$), and significantly more often to [l] (mean = 0.956) than [ɹ] (mean = 0.897) words ($F(1,23) = 7.734, p = 0.011, \eta_p^2 = 0.252$). These results are plotted in **Figure 6**. In contrast to the native English and the native Mandarin speakers, the native

Korean speakers exhibited asymmetries for both words and nonwords, but in opposite directions, with higher accuracy on [ɹ] than [l] nonwords but higher accuracy on [l] than [ɹ] words. This pattern of responses is consistent with scenarios 5 and 6, where distinctive perceptual representations of [ɹ] and [l] contact phonological representations containing familiar /l/ and neutralized (scenario 5) or ambiguous (scenario 6) phonological representations of /l/.

DISCUSSION

Our first goal was to report LD data for the same two L2 contrasts across three L1 groups, taking into account patterns of performance on both word and nonword stimuli so as to evaluate the findings with respect to the predictions of the scenarios described above. The LD task revealed that native English, Mandarin, and Korean speakers all exhibit near-ceiling acceptance of English words containing [æ] and [ɛ], and are less accurate when it comes to rejecting [æ] and [ɛ] nonwords. While native English speakers' performance on nonwords was more accurate than that of the two learner groups, all three groups exhibited the same nonword asymmetry with more accurate responses to [æ] than to [ɛ] nonwords. Participants' near-ceiling performance on these word stimuli is reminiscent of the findings of a number of other studies (Sebastián-Gallés and Baus, 2005; Sebastián-Gallés et al., 2005; Sebastián-Gallés et al., 2006; Díaz et al., 2012; Llompart, 2020), and is often cited as the reason for reporting analyses of A' or d' scores (collapsing across words and nonwords; Sebastián-Gallés and Baus, 2005; Díaz et al., 2012; Llompart and Reinisch, 2019) or reporting analyses of nonword data only (Llompart, 2020). However, as noted earlier, in the absence of certainty regarding the dominance of the phonemes, an asymmetry for nonwords only is ambiguous with respect to whether it provides evidence for difficulties at the phonetic or phonological levels.

A very different picture emerged regarding the /l/-/ɹ/ contrast. Native speakers' performance was consistent with the predictions of scenario 8 (and 7), where both phonetic and phonological representations unambiguously encode the contrast. The native Mandarin speakers' performance matched the ordinal accuracy predictions of the perceptual coding scenarios (scenarios 2, 3, and 4), where the learner is understood to experience difficulty at the level of speech perception. In contrast, the native Korean speakers' performance was consistent with the ordinal accuracy predictions of the lexical coding scenarios (scenarios 5 and 6). We thus have evidence that the same materials can elicit three distinct patterns of performance depending on L1 background. The native speakers of English performed as if presumably expected of native speakers (symmetric and highly accurate performance in all conditions); the native Mandarin participants' performance suggested difficulty at the perceptual level with /l/ behaving as the dominant category; and the native Korean participants' performance was consistent with difficulty at the level of lexical coding, also with /l/ behaving as the dominant category. That /l/ appears to be the dominant

category for the native Mandarin and Korean speakers is unsurprising given the status of /l/ in Mandarin. In the case of Korean, the combination of word and nonword asymmetries found here suggests both that /l/ is dominant for this group and that the locus of difficulty is at the lexical encoding level. A question that remains is whether the data presented here allows us to distinguish between the predictions of scenarios 2 vs. 3 and 4 or 5 vs. 6 (which produce the same ordinal accuracy predictions but differ in the magnitude of the asymmetries). The simple answer is "no", though the availability of these more nuanced predictions among the set of scenarios proposed here begs the question of how such distinctions might arise. It is possible that developmental data involving learners at different stages of L2 acquisition will provide evidence that distinguishes these scenarios (see, e.g., Darcy et al., 2013; see also Broersma and Cutler, 2011 for discussion concerning the assessment of the amount of spurious activation experienced by learners). Future work is needed to determine whether these predictions are indeed evidenced in learners.

Our second goal was to allow for the comparison of performance by native and nonnative speakers by documenting native speaker LD performance (in addition to the performance of two different nonnative speaker groups). As already noted, doing so has revealed both expected and unexpected patterns in native English speakers. With respect to the /l/-/ɹ/ contrast, native English speakers performed as expected with high accuracy for both word and nonword stimuli, suggesting precise and detailed perceptual and lexical coding of the contrast. Unexpectedly, near-ceiling effects on word stimuli coupled with high false alarm rates and asymmetric performance on nonwords were observed for the /æ/-/ɛ/ contrast in all three groups. While asymmetries in LD performance by L2 learners have been attributed to difficulties with perceptual and/or phonological processing, asymmetric LD performance has also been observed in native speakers (Díaz et al., 2012; Darcy et al., 2013; Melnik and Peperkamp, 2019), as well as in learners processing of words containing so-called "easy" contrasts (e.g., native German speakers exhibited higher d' scores for /I/-items (winter/*w[i]nter) than /i/-items (needle/*n[I]ddle) in Llompart and Reinisch, 2019). Specifically for the /æ/-/ɛ/ contrast and native English speakers, high false alarm rates for [æ] and [ɛ] nonwords have been attested (Broersma and Cutler, 2011; Díaz et al., 2012). Curiously, native English speakers' /æ/-/ɛ/ response patterns here are similar to those reported by Díaz et al. (2012), except that the American English speakers in the present study exhibit an asymmetry in favor of [æ] nonwords, and the British English speakers in their study perform more accurately on [ɛ] nonwords. Díaz et al. (2012) speculated that the differential performance for word and nonwords containing [æ] and [ɛ] in their native English speakers may be due to 1) higher frequency of /ɛ/ than /æ/ in English, and/or 2) possible idiosyncrasies in speaker's pronunciation of /æ/ and /ɛ/ (noting, however, that there were no such asymmetries for same speaker in perception-only task with different speech tokens). The explanation based on frequency is unlikely to account for the pattern we observed here, since the nonword asymmetry goes in the opposite direction; however, the question of stimulus phonetics is an important one

that is expected to interact with speech perception behavior. It is worth noting that these vowel segments [æ] and [ɛ] are heavily overlapped in the acoustic space as produced by native speakers of American English (Hillenbrand et al., 1995), and the acoustic properties of the segments involved in the contrast may play a role.

It is also possible that native (and nonnative) speaker asymmetries in lexical activation can be understood in the context of asymmetries in speech perception. Peripheral vowels are known to behave as perceptual anchors and to be more readily detected than less peripheral vowels (Polka and Bohn, 2003; Polka and Bohn, 2011), and that formant proximity and stimulus prototypicality influence perceptual asymmetries (Liu et al., 2021). Having observed an asymmetry in German learners of English in the lexical processing of the /æ/-/ɛ/ contrast with more accurate performance for [æ] nonwords over [ɛ] nonwords (like the one we report here for our three listener groups), Llompart and Reinisch (2019) and Llompart (2020) suggest that the relatively more peripheral position of /æ/ than /ɛ/ make it a better perceptual anchor. Llompart and Reinisch (2019) also attribute the unexpected asymmetry observed for the “easy” /i/-/ɪ/ contrast noted above to the more peripheral nature of /i/. At this point it is unclear why the native speakers exhibited the particular asymmetry found here, though stimulus properties, including the statistical distributions of the phonemes involved, lexical properties of words containing those phonemes, and the acoustic-phonetic properties of the stimuli themselves, may all influence LD performance. Ultimately, however, the utility of LD tasks in the investigation of L2 phonolexical processing depends on an understanding of why asymmetries are sometimes also observed in native speakers, and future work must explore the reasons behind native speaker asymmetries in order to determine the appropriate interpretation of L2 learners’ asymmetric patterns.

Our third goal was to increase the representation of late learners of English in the US in the body of research on L2 phonolexical processing. By recruiting via Prolific and word-of-mouth, we were able to include participants outside of the undergraduate and graduate student population that is typically represented in studies of both native speakers and language learners. While we did not systematically collect information about education level or reasons for emigrating to the United States, interaction with participants as well as their responses to open-ended questions about their language backgrounds revealed that many of the native Mandarin and Korean participants emigrated to the United States for reasons other than post-secondary education, and therefore represent a variety of English language learners in the United States. Indeed, a consequence of this recruitment strategy is that our participants’ linguistic backgrounds—in particular, the circumstances under which they acquired English—varied widely, with potential impacts on their performance in the present study at individual and/or group levels, emphasizing the need for further investigation of both individual differences and learner backgrounds in the study of L2 phonolexical processing.

The LD task has been widely employed in the study of the lexical processing of language learners. However, a strong

response bias often leads to high hit rates (and ceiling effects for words) and high false alarm rates for nonwords; for this reason, researchers have often chosen to report signal detection measures such as A’ or d’ scores (e.g., Sebastián-Gallés and Baus, 2005; Díaz et al., 2012; Llompart and Reinisch, 2019). These signal detection measures factor out response bias by simultaneously taking into account the proportion of accurate YES responses to words (hit rate) and the proportion of inaccurate YES responses to nonwords (false alarm rate) and provide a single measure of sensitivity. However, the practice of reporting only these measures is problematic, as it obscures away from the raw data, and more importantly for our purposes because it conceals possible asymmetries in lexical processing which we have argued may be helpful for understanding the locus of difficulty for learners with respect to novel phonological contrasts.

Ceiling effects for words are also problematic for understanding LD data with respect to the scenarios we have fleshed out, since asymmetries in words and their direction relative to nonword asymmetries are required to distinguish between the perceptual coding and lexical coding scenarios, in the absence of strong a priori reasons to believe that a particular category will be dominant for a learner population. Take, for example, the case in which robust asymmetries are observed for nonwords only. On the one hand, data of this sort will almost certainly result in an interaction between lexical status and surface segment, which may be indicative of lexical coding difficulty (see Darcy et al., 2013). However, nonword differences might alternatively reflect difficulty at the level of phonetic coding if ceiling effect for word stimuli (due either to overall accurate performance on the task or a strong word bias) obscure differences in performance. As a result, word data that is not at ceiling would seem to be crucial for understanding these asymmetries. Moreover, the fact that robust asymmetries have been observed for word stimuli when nonword asymmetries are absent (Melnik and Peperkamp, 2019), also provides empirical grounds for not ignoring word data. It is possible that presenting the auditory stimuli in noise,⁴ or manipulating characteristics of control or filler items, would make the task more challenging or help to reduce response bias, moderating ceiling effects for words.

As highlighted above, a firm understanding of which member of a new contrast functions as the dominant category vs. the non-dominant category is crucial for being able to distinguish the perceptual coding from the lexical coding scenarios, particularly in the face of ceiling effects for words. Despite its importance, there has been little discussion of the best diagnostics for determining dominance. Some practices include inventory comparisons, consultation of existing perception data, or production acoustics. The problem is not unlike the challenge of

⁴Cutler et al. (2004) report data L1 English and L2 Dutch listeners suggesting that English phoneme identification is impacted by noise to an equal extent, suggesting that embedding words in noise may be a feasible option.

predicting phonological similarity (see, e.g., Barrios et al., 2016), and requires further exploration in this literature.

An additional limitation of this study is that we did not systematically evaluate the prosodic positions of the studies segments. The effect of position on the difficulty posed by the English /l/-/ɹ/ contrast has been documented for native speakers of both Mandarin and Korean (Ingram and Park, 1998; Hazan et al., 2006). More generally, phonological context and its effects on L2 speech perception is likely to affect the lexical encoding of both consonants and vowels, and should be investigated in future work in this area.

In conclusion, the auditory lexical decision task is attractive as a time- and cost-effective method that holds the promise of simultaneously providing information about perceptual and lexical coding. It is further appealing because it can be readily carried out online, thus reaching more than the typical convenience sample of participants. The proliferation of studies reporting LD data in recent years has resulted in inevitable variability in implementation and interpretation, and has revealed some potential pitfalls of the method with respect to clarifying the locus of difficulty in L2 lexical processing. Here we spell out the predictions of several logically possible scenarios, and provide new data for the /æ/-/ɛ/ contrast that illustrates the difficulty of interpreting asymmetries for nonwords only, in addition to the analytical challenge that arises when native speakers also exhibit LD asymmetries. The /l/-/ɹ/ contrast materials elicited evidence of three distinct scenarios, providing new data demonstrating the effects of L1 background on the perceptual and/or lexical coding difficulties experienced by language learners. We believe that these findings together highlight the need for further research that explores and addresses the promise and drawbacks of the LD task in the study of L2 phonolexical processing.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://osf.io/5z6y7/>.

REFERENCES

- Amengual, M. (2016a). Cross-Linguistic Influence in the Bilingual Mental Lexicon: Evidence of Cognate Effects in the Phonetic Production and Processing of a Vowel Contrast. *Front. Psychol.* 7, 617. doi:10.3389/fpsyg.2016.00617
- Amengual, M. (2016b). The Perception of Language-Specific Phonetic Categories Does Not Guarantee Accurate Phonological Representations in the Lexicon of Early Bilinguals. *Appl. Psycholinguist.* 37, 1221–1251. doi:10.1017/S0142716415000557
- Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., and Yamada, T. (2004). Perceived Phonetic Dissimilarity and L2 Speech Learning: the Case of Japanese /r/ and English /l/ and /r/. *J. Phonet.* 32, 233–250. doi:10.1016/s0095-4470(03)00036-6
- Barrios, S., Jiang, N., and Idsardi, W. J. (2016). Similarity in L2 Phonology: Evidence from L1 Spanish Late-Learners' Perception and Lexical Representation of English Vowel Contrasts. *Second Lang. Res.* 32, 367–395. doi:10.1177/02676583166630784
- Best, C. T. (1995). "A Direct Realist View of Cross-Language Speech Perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Editor W. Strange (Timonium, MD: York Press), 171–204.
- Broersma, M., and Cutler, A. (2011). Competition Dynamics of Second-Language Listening. *Q. J. Exp. Psychol.* 64, 74–95. doi:10.1080/17470218.2010.499174
- Broersma, M. (2012). Increased Lexical Activation and Reduced Competition in Second-Language Listening. *Lang. Cogn. Process.* 27, 1205–1224. doi:10.1080/01690965.2012.660170
- Brown, C. A. (1998). The Role of the L1 Grammar in the L2 Acquisition of Segmental Structure. *Second Lang. Res.* 14 (2), 136–193. doi:10.1191/026765898669508401
- Brown, C. A. (2000). "The Interrelation Between Speech Perception and Phonological Acquisition from Infant to Adult," in *Second Language Acquisition and Linguistic Theory*. Editor J. Archibald (Malden, Massachusetts: Wiley-Blackwell), 4–64.
- Chen, Y., Robb, M., and Gilbert, H. (2001). Vowel Production by Mandarin Speakers of English. *Clin. Linguist. Phonet.* 15, 427–440.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Utah's Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The project was initiated by RH-H, and joined by SB following a pilot phase. For the data reported here, both authors contributed to materials development. SB set up online data collection, and RH-H conducted the analyses. Both contributed to all parts of manuscript preparation.

FUNDING

This research was funded in part by a Kickstart Grant from the College of Humanities at the University of Utah awarded to RH-H and SB.

ACKNOWLEDGMENTS

We are grateful to our University of Utah colleagues in the Speech Acquisition Lab and the Cognitive Language Aficionados Research Group, as well as the editor and reviewers, for their feedback on this project. We are also grateful to the very many colleagues who supported our participant recruitment efforts, and to the study participants, for their time and attention.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2021.689470/full#supplementary-material>

- Cho, M.-H., and Jeong, S. (2013). Perception and Production of English Vowels by Korean Learners: A Case Study. *Stud. Phonet. Phonol. Morphol.* 19, 155–177. doi:10.17959/sppm.2013.19.1.155
- Chrabaszcz, A., and Gor, K. (2014). Context Effects in the Processing of Phonolexical Ambiguity in L2. *Lang. Learn.* 64, 415–455. doi:10.1111/lang.12063
- Chrabaszcz, A., and Gor, K. (2017). Quantifying Contextual Effects in Second Language Processing of Phonolexically Ambiguous and Unambiguous Words. *Appl. Psycholinguist.* 38, 909–942. doi:10.1017/s0142716416000497
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Q. J. Exp. Psychol. Section A* 33, 497–505. doi:10.1080/14640748108400805
- Cook, S. V., and Gor, K. (2015). Lexical Access in L2. *ML* 10, 247–270. doi:10.1075/ml.10.2.04coo
- Cook, S. V., Pandža, N. B., Lancaster, A. K., and Gor, K. (2016). Fuzzy Nonnative Phonolexical Representations lead to Fuzzy Form-to-Meaning Mappings. *Front. Psychol.* 7, 1345. doi:10.3389/fpsyg.2016.01345
- Cutler, A., Weber, A., and Otake, T. (2006). Asymmetric Mapping from Phonetic to Lexical Representations in Second-Language Listening. *J. Phonet.* 34, 269–284. doi:10.1016/j.wocn.2005.06.002
- Cutler, A., Weber, A., Smits, R., and Cooper, N. (2004). Patterns of English Phoneme Confusions by Native and Non-Native Listeners. *J. Acoust. Soc. Am.* 116, 3668–3678. doi:10.1121/1.1810292
- Cutler, A. (2005). “The Lexical Statistics of Word Recognition Problems Caused by L2 Phonetic Confusion,” in *INTERSPEECH-2005*, 413–416.
- Darcy, I., Daidone, D., and Kojima, C. (2013). Asymmetric Lexical Access and Fuzzy Lexical Representations in Second Language Learners. *Mental Lexicon* 8, 372–420. doi:10.1075/ml.8.3.06dar
- Darcy, I., and Thomas, T. (2019). When Blue Is a Disyllabic Word: Perceptual Epenthesis in the Mental Lexicon of Second Language Learners. *Bilingualism* 22, 1141–1159. doi:10.1017/S1366728918001050
- Díaz, B., Mitterer, H., Broersma, M., and Sebastián-Gallés, N. (2012). Individual Differences in Late Bilinguals’ L2 Phonological Processes: From Acoustic-Phonetic Analysis to Lexical Access. *Learn. Individ. Differ.* 22, 680–689. doi:10.1016/j.lindif.2012.05.005
- Díaz, B. (2021). *Personal Communication with R. Hayes-Harb on Mar 23, 2021*.
- Escudero, P., Hayes-Harb, R., and Mitterer, H. (2008). Novel Second-Language Words and Asymmetric Lexical Access. *J. Phonet.* 36, 345–360. doi:10.1016/j.wocn.2007.11.002
- Extra, G., and Verhoeven, L. (2011). *Bilingualism and Migration*. Berlin, NY: Walter de Gruyter. Available at: <https://play.google.com/store/books/details?id=ot6vWLDxMl0C>.
- Flege, J. E., Bohn, O.-S., and Jang, S. (1997). Effects of Experience on Non-Native Speakers’ Production and Perception of English Vowels. *J. Phonet.* 25, 437–470. doi:10.1006/jpho.1997.0052
- Hayes-Harb, R., and Barrios, S. (2019). “Investigating the Phonological Content of Learners’ “Fuzzy” Lexical Representation for New L2 Words,” in Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference. Editors J. M. Levis, C. L. Nagle, and E. Todey, Ames, IA: Iowa State University, 55–69. Available at: <https://apling.engliastate.edu/wp-content/uploads/sites/221/2019/07/10th-Annual-PSLLT-Proceedings-FINAL.pdf>.
- Hayes-Harb, R., and Masuda, K. (2008). Development of the Ability to Lexically Encode Novel Second Language Phonemic Contrasts. *Second Lang. Res.* 24, 5–33. doi:10.1177/0267658307082980
- Hazan, V., Sennema, A., Faulkner, A., Ortega-Llebaria, M., Iba, M., and Chung, H. (2006). The Use of Visual Cues in the Perception of Non-Native Consonant Contrasts. *J. Acoust. Soc. Am.* 119, 1740–1751. doi:10.1121/1.2166611
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). Acoustic Characteristics of American English Vowels. *J. Acoust. Soc. America* 97 (5), 3099–3111. doi:10.1121/1.411872
- Hong, S. (2012). The Relative Perceptual Easiness Between Perceptually Assimilated Vowels for University-Level Korean Learners of American English and Measurement Bias in an Identification Test. *Stud. Phonet. Phonol. Morphol.* 18, 491–511. doi:10.17959/sppm.2012.18.3.491
- Ingram, J. C. L., and Park, S.-G. (1998). Language, Context, and Speaker Effects in the Identification and Discrimination of English /r/ and /l/ by Japanese and Korean Listeners. *J. Acoust. Soc. Am.* 103, 1161–1174. doi:10.1121/1.421225
- Jia, G., Strange, W., Wu, Y., Collado, J., and Guan, Q. (2006). Perception and Production of English Vowels by Mandarin Speakers: Age-Related Differences Vary with Amount of L2 Exposure. *J. Acoust. Soc. Am.* 119, 1118–1130. doi:10.1121/1.2151806
- Kim, J.-E. (2010). Perception and Production of English Front Vowels by Korean Speakers. *Phonetics Speech Sci.* 2, 51–58. Available at: <http://www.koreascience.or.kr/article/JAKO201019455946033.page>.
- Kim, J. G. (2007). “Evidence of /l/-r/ Contrast in Korean,” in ICPHS 2007 Proceedings. Editors J. Trouvain and W. J. Barry (Saarbrücken, Germany: University des Saarlandes), 649–652.
- Lennes, M. (2003). Analyze Formant Values from Labeled Segments in Files [Praat Script]. Available at: http://phonetics.linguistics.ucla.edu/facilities/acoustic/collect_formant_data_from_files.txt (Accessed February 20, 2020).
- Liu, Y. Y., Polka, L., Masapollo, M., and Ménard, L. (2021). Disentangling the Roles of Formant Proximity and Stimulus Prototypicality in Adult Vowel Perception. *JASA Express Lett.* 1 (1), 015201. doi:10.1121/10.0003041
- Llompарт, M. (2021). *Personal communication with R. Hayes-Harb and S. Barrios on Mar 26, 2021*.
- Llompарт, M. (2020). Phonetic Categorization Ability and Vocabulary Size Contribute to the Encoding of Difficult Second-Language Phonological Contrasts into the Lexicon. *Bilingualism* 24, 481–496. doi:10.1017/S1366728920000656
- Llompарт, M., and Reinisch, E. (2017). Articulatory Information Helps Encode Lexical Contrasts in a Second Language. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1040–1056. doi:10.1037/xhp0000383
- Llompарт, M., and Reinisch, E. (2019). Robustness of Phonolexical Representations Relates to Phonetic Flexibility for Difficult Second Language Sound Contrasts. *Bilingualism* 22, 1085–1100. doi:10.1017/S1366728918000925
- Llompарт, M., and Reinisch, E. (2020). The Phonological Form of Lexical Items Modulates the Encoding of Challenging Second-Language Sound Contrasts. *J. Exp. Psychol. Learn. Mem. Cogn.* 46, 1590–1610. doi:10.1037/xlm0000832
- Melnik, G. A., and Peperkamp, S. (2019). Perceptual Deletion and Asymmetric Lexical Access in Second Language Learners. *J. Acoust. Soc. America* 145, EL13–EL18. doi:10.1121/1.5085648
- Pallier, C., Colomé, A., and Sebastián-Gallés, N. (2001). The Influence of Native-Language Phonology on Lexical Access: Exemplar-Based versus Abstract Lexical Entries. *Psychol. Sci.* 12, 445–449. doi:10.1111/1467-9280.00383
- Park, H. (2013). Detecting Foreign Accent in Monosyllables: The Role of L1 Phonotactics. *J. Phonet.* 41 (2), 78–87. doi:10.1016/j.wocn.2012.11.001
- J. Paradis, X. Chen, H. Ramos, and M. Crago (Editors) (2020). The Language, Literacy and Social Integration of Refugee Children and Youth [Special Issue]. *Appl. Psycholinguist* 41 (6), 1251–1254. doi:10.1017/S0142716420000788
- Polka, L., and Bohn, O.-S. (2003). Asymmetries in Vowel Perception. *Speech Commun.* 41, 221–231. doi:10.1016/S0167-6393(02)00105-X
- Polka, L., and Bohn, O.-S. (2011). Natural Referent Vowel (NRV) Framework: An Emerging View of Early Phonetic Development. *J. Phonet.* 39, 467–478. doi:10.1016/j.wocn.2010.08.007
- Schmidt, A. M. (1996). Cross-Language Identification of Consonants. Part 1. Korean Perception of English. *J. Acoust. Soc. Am.* 99 (5), 3201–3211. doi:10.1121/1.414804
- Sebastián-Gallés, N., and Baus, C. (2005). “On the Relationship between Perception and Production in L2 Categories,” in *Twenty-first Century Psycholinguistics: Four Cornerstones*. Editor A. Cutler (Mahwah, New Jersey: Erlbaum), 279–292.
- Sebastián-Gallés, N., Echeverría, S., and Bosch, L. (2005). The Influence of Initial Exposure on Lexical Representation: Comparing Early and Simultaneous Bilinguals. *J. Mem. Lang.* 52, 240–255. doi:10.1016/j.jml.2004.11.001
- Sebastián-Gallés, N., Rodríguez-Fornells, A., De Diego-Balaguer, R., and Díaz, B. (2006). First- and Second-Language Phonological Representations in the Mental Lexicon. *J. Cogn. Neurosci.* 18, 1277–1291. doi:10.1162/jocn.2006.18.8.1277
- Simon, E., Sjerps, M. J., and Fikkert, P. (2014). Phonological Representations in Children’s Native and Non-Native Lexicon. *Bilingualism* 17, 3–21. doi:10.1017/S1366728912000764
- Simonchyk, A., and Darcy, I. (2017). “Lexical Encoding and Perception of Palatalized Consonants in L2 Russian,” in Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference. Editors M. G. O’Brien and J. M. Levis (Ames, IA: Iowa State University), 121–132.
- Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., and Flege, J. (2005). A Developmental Study of English Vowel Production and Perception by Native

- Korean Adults and Children. *J. Phonet.* 33, 263–290. doi:10.1016/j.wocn.2004.10.002
- Wang, X. (1997). The Acquisition of English Vowels by Mandarin ESL Learners: A Study of Production and Perception. Available at: <http://summit.sfu.ca/system/files/iritems1/7404/b18794348.pdf> (Accessed March 19, 2021).
- Weber, A., and Cutler, A. (2004). Lexical Competition in Non-Native Spoken-word Recognition. *J. Mem. Lang.* 50, 1–25. doi:10.1016/S0749-596X(03)00105-0
- White, E. J., Titone, D., Genesee, F., and Steinhauer, K. (2017). Phonological Processing in Late Second Language Learners: The Effects of Proficiency and Task. *Bilingualism* 20, 162–183. doi:10.1017/S1366728915000620
- Yang, B. (1996). A Comparative Study of American English and Korean Vowels Produced by Male and Female Speakers. *J. Phonet.* 24, 245–261. doi:10.1006/jpho.1996.0013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Barrios and Hayes-Harb. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.