Darcy, I. & Holliday, J. J. (2019). Teaching an old work new tricks: Phonological updates in the L2 mental lexicon. In J. Levis, C. Nagle, & E. Todey (Eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Ames, IA, September 2018 (pp. 10-26). Ames, IA: Iowa State University.

INVITED TALK

TEACHING AN OLD WORD NEW TRICKS: PHONOLOGICAL UPDATES IN THE L2 MENTAL LEXICON

Isabelle Darcy, Indiana University, Bloomington (USA)
Jeffrey J. Holliday, Korea University, Seoul (South Korea)

We examine the dynamic relationship between perception of challenging phonological dimensions and their lexical representation in second language learners. We ask whether learners update the phonological form of words in their mental lexicon for all words simultaneously as a result of perception improvements, or whether updates are word-specific. Taking into account the trajectory of perceptual development and word learning over time, we examine this question using a lexical decision task targeting two vowel contrasts in Korean: /o/-/ʌ/ (test) and /o/-/a/ (control), the test contrast being especially challenging for L1 Mandarin learners of Korean. Participants also completed a vowel identification task, a background questionnaire, and a word familiarity questionnaire. The results confirmed that several learners had imprecise phonolexical representations, especially for words containing the test contrast. While most learners were very accurate at identifying this contrast, those with the most confusions were also least accurate in lexical decision. We also observed a trend towards word-specific phonolexical updates: words that were learned more recently were encoded more accurately than words learned earlier. The data raise the question of which lexical representations bilinguals create for words they learn, and how pronunciation instruction can help addressing phonological issues in the bilingual mental lexicon.

INTRODUCTION

This study examines how second language (L2) learners mentally store words. Lexical encoding of difficult L2 phonemic contrasts has been shown to be challenging (e.g. Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008; Ota, Hartsuiker, & Haywood, 2009), but little is known about how learners update initially inaccurate lexical representations. Over time, sustained input helps the L2 phonological system develop, and processing of phonological dimensions becomes more accurate (Flege, Bohn, & Jang, 1997). Similarly, learners' lexical representations become more accurate over time (Darcy, Dekydtspotter, Sprouse et al., 2012). Given the assumed link between perceptual ability and word form learning (Pallier, Bosch, & Sebastian-Gallés, 1997; Pallier, Colomé, & Sebastian-Gallés, 2001), it is possible that improvements in lexical encoding depend on improvements in perceptual accuracy – yet this is still an open question. In this paper, we ask whether lexical representations are updated as people learn more about the phonology of their L2.

Updates to lexical representations

As adult speakers of a language, we know tens of thousands of words. We learn words our whole life, in our first language, but also in languages learned after the first. In a second language, just as in the first, some words are acquired very early, others much later. For example, for the first author who has been learning English as a second language for about 30 years, the words “hello” and “soffit” are both part of her lexicon, but one was learned in the very first days of learning English, whereas the second one was learned 30 years later, during the summer of 2018.

For each word we know (e.g. “mouse”), we also know many things about it. We know what it means, and we know what it sounds like – that is, we have a PHONOLEXICAL REPRESENTATION of the word stored in memory. We also know how it is written, that it is a noun, that its plural is *mice*, etc.

But language users don’t just learn new words, they sometimes need to change – or update – the words that they already know. For example, many of us needed to add a new meaning to the lexical representation of “mouse” in order to refer to a computer device used to position the cursor and click anywhere on the monitor. In fact, for some of us, this meaning may be the dominant meaning, while the other meaning referring to a small animal may be more weakly or not at all activated.

As this example shows, lexical representations can be updated by adding to or expanding the meanings attached to an item. The same is true for the phonological representations that we have stored in the mental lexicon – and in particular, for the words we learn in an L2. One would expect that lexical representations can evolve to reflect improved phonological knowledge and possibly show fewer instances of L1 influence in the phonolexical representation, for example. However, we don’t know whether or how this process happens, and what factors lead to updates in the phonological representations stored in the mental lexicon.

Phonolexical representations for L2 words change over time

Research evidence showing that L2 learners improve in lexical tasks suggests that it is possible to update the phonolexical representations they created for the L2 words they learned. For example, more advanced learners outperform intermediate learners in lexical decision tasks where they have to tell words from nonwords (Darcy, Daidone, & Kojima, 2013; Darcy, Park, & Yang, 2015; Darcy & Thomas, 2019). Higher proficiency also helps them detect mispronunciations in words (Simonchyk & Darcy, 2018), or helps them experience less competition from similar words (Cook, Pandža, Lancaster, & Gor, 2016; Veivo & Järvikivi, 2013). These findings suggest that learners’ lexical representations become more precise over time, and possibly become less influenced by the L1 phonology.

At the same time, a number of studies have shown that with accruing experience of the L2, learners can develop a more accurate knowledge of its phonological system. While most learners initially often misperceive phonetic and phonological dimensions, improvements can be seen in both perception and production – at least for some phonetic/phonological dimensions (e.g. Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Levy & Strange, 2008; Wayland & Guion, 2003).

Interestingly, these improvements in perception or production do not seem to parallel those in lexical tasks, which suggests that having acquired a contrast in perception does not guarantee accurate lexical representations (e.g. Darcy et al., 2012; Simonchyk & Darcy, 2017). In Simonchyk and Darcy, a group of intermediate and very advanced learners of Russian took part in a perception task in which they had to distinguish palatalized from non-palatalized consonants, and in a lexical task with words and nonwords involving the palatalization contrast. Knowledge of all words in the lexical task was verified with a word familiarity questionnaire. Despite high familiarity with all words, error rates on the test condition of the lexical task were high – the minimum was 30 % error, and some participants made 80–90% errors. When comparing error rates on the perception task to those in the lexical task, no clear link emerged: even one subgroup of participants with a homogeneous error rate around 15% in the perception task displayed wide-ranging error scores on the lexical task (30-90% across both groups). This means that even the listeners who were quite accurate at distinguishing the palatal and non-palatal sounds in Russian did not always know which of the two should be stored in the lexical representation of words they were very familiar with. This kind of data brings home the point that the mental lexicon in learners can be quite abstract or separate from phonetic categorization performance. We can therefore conclude that lexical updates are not always automatic, and that even at high levels of proficiency learners don't necessarily have accurate phonolexical representations.

An additional issue regarding phonolexical representations is that few studies consider item variance on lexical tasks. In a recent study (Darcy & Thomas, 2019), an item analysis revealed substantial variability in error rates between items, indicating that the lexical representations for some words seem to get updated, while other words appear to be much harder to update. Interestingly, the variability was not easily explained by item characteristics such as familiarity, frequency, loanword status, cluster type, among others.

Even though at first glance perception and word familiarity appear to have limited explanatory power in terms of the mechanisms behind lexical updates, it is possibly because both factors have mostly been considered in a "static" manner. As a large number of studies show, perception evolves over time, it is not static. Considering perception in a dynamic way might reveal a clearer relationship with the form of lexical representations (see also Nagle, 2018, on the link between perception and production). Similarly, familiarity considered alone may have limited explanatory power. But in interaction with changes in perception, it might offer an interesting picture regarding the form of lexical representations. In a nutshell, familiar words learned when perception was at a "beginner-level" may be encoded with a lack of phonological precision. If learners do not update these initial representations, these would then be reinforced over the years. By contrast, more recent familiar words, even though they may have been encountered in fewer instances, may be encoded more precisely as a result of developments in perception.

In this study, we examine the interaction of these factors in a dynamic way, considering the trajectory of perceptual development and word learning over time in learners. Simply put, this interaction could be thought of as the timing of ***when*** each word was learned. This could explain at least some of the variability previously encountered in lexical tasks even for familiar words: it is possible to imagine that words learned early (and which have been part of the mental lexicon for a long time) may be represented in less target-like ways compared to words learned more recently, if the phonolexical representations of words indeed reflect the acquisition "stage" of the L2 phonological system. We set out to investigate this idea that updates may be influenced by the timing of when a given word was learned.

We call this possibility the AGE OF WORDS HYPOTHESIS. In a nutshell, if words that contain a difficult contrast are updated according to the timing of learning, those learned early (that is, a long time ago) might be harder to update than words learned recently. In this scenario, updates are word-specific and depend on when a word was learned; updates are therefore also learner-specific, and selectively apply first to words learned *after* a new perceptual dimension has been acquired. In that case, a word's phonolexical form would also partly reflect the learner's perceptual ability regarding the difficult contrast at the time of learning. Phonolexical representations of words learned at early stages of phonological development would then reflect this earlier version of the perceptual system, while words learned more recently would reflect the perceptual progress made since. This dynamic process, by which updates first enter the lexicon through new words, and then gradually permeate the system retroactively to update older forms, is expected to take time until all words are updated, and would effectively result in old words lagging behind recent words in terms of the accuracy of lexical representations.

A different hypothesis is that word age does not play a role for the updates. Instead, learners may update their lexicon wholesale: as they acquire a new perceptual dimension (e.g. a specific vowel contrast), all lexical representations containing this dimension are updated simultaneously (PHONOLOGICAL UPDATE HYPOTHESIS). If this hypothesis is correct, there would be no effect of age.

Research questions and predictions

The following question guided the current investigation: Are updates to phonolexical representations influenced by the timing of when a given word was learned?

We hypothesize that the timing of when a word is learned indeed matters, and that there is a difference in how accurately the phonological form of words is represented in a learner's mental lexicon based on when the words were learned.

To test this hypothesis, we first establish to what extent learners have developed perceptual mastery for a pair of difficult vowels (indexed by their ability to perceptually identify each vowel); we then use a lexical decision task to probe the form of learners' lexical representations, comparing responses for words that were learned early ("old") vs. more recently ("young"). Specifically, we selected a number of test words (both old and young words) which contain the difficult vowel pair. We then manipulated these items to create nonwords by switching one vowel by its counterpart. For instance, for the difficult vowel pair /o/ and /ʌ/, an example word containing /o/ [sogɛ] 'introduce' would become the nonword [sʌgɛ] by switching out the /o/ with the /ʌ/ vowel. In a lexical decision task, the learner would ideally respond "yes" (it is a real word of Korean) to [sogɛ], but respond "no" (is it not a real word in Korean) to [sʌgɛ]. This latter response is what we call "nonword rejection", when participants successfully reject nonwords in the task. Since learners are likely to have built initially imprecise phonolexical representations for these words due to the difficult vowels they contain (Darcy, Daidone & Kojima, 2013), we interpret the failure to reject these nonwords (that is, incorrectly accepting [sʌgɛ] as a real word – see methods) to mean that their phonolexical representations are still imprecise with respect to that vowel. Likewise, successfully rejecting these nonwords indexes the accuracy with which the corresponding words are lexically represented. We compare the nonword rejection rate for nonwords based on "young" vs. "old" words. If indeed the age of words matters for phonological updating in lexical representations, the test vowels in *young* words should be encoded more accurately than in *old* words, for learners who have acquired this vowel contrast. Thus, learners should fail to reject nonwords that are based on *old* words more often than those based on *young* words, and nonword rejection rate will be higher for young words. Notably, this effect is not expected to hold for control words and nonwords (that is, those that do not contain a difficult vowel). It may even be reversed, simply because listeners are more familiar with old words and less so with newly learned words (which therefore may be harder to either accept or reject). If phonological updates happen wholesale, no such advantage for young words is expected on the test items.

METHODS

In this study, we worked with Chinese L2 learners of Korean. To examine whether these learners have acquired the difficult vowel contrast in their L2, we used a perceptual identification task for Korean vowels. The vowel contrast of interest is the /o/-/ʌ/ contrast which is reportedly difficult for Chinese learners of Korean. To probe the form of learners' lexical representations, comparing old and young words, we used an auditory lexical decision task. Participants also completed a background questionnaire, and a word familiarity and word learning history questionnaire.

Participants

Twenty-seven native speakers of Mandarin Chinese participated (mean age = 22.5 years; SD = 1.99). They were enrolled students at a major South Korean university and were living in Korea at the time of testing (mean length of residence = 1.93 years, SD = 1.52). We did not administer a separate Korean proficiency test but all participants had spent more than 1 year studying at a Korean university, where most classes are held in Korean, and they were able to fill out an extensive background questionnaire in Korean (mean age of arrival in Korea = 20.6 years, SD = 1.65).

Materials and procedure

Vowel identification task. In the vowel identification task, participants listened to 150 CV syllables excised from running speech productions. The stimuli were presented via headphones one by one, while a screen (see Figure 1) was displayed. The task was a forced choice with the 5 vowel categories (as Korean letter symbols) given as choices. They chose the letter for the vowel they heard (the IPA symbols were not displayed; they are provided here for convenience). In Korean, grapheme-phoneme correspondences for vowels are very transparent, and all learners were familiar with the letter symbols. The CV syllables contained 30 tokens each of five different vowels, including 30 /o/ and 30 /ʌ/. We look specifically at the /o/ and /ʌ/ contrast because it is the test contrast for the lexical decision task. We counted the confusions for these vowels (out of 60, expressed in %), as well as for the whole set of vowels.

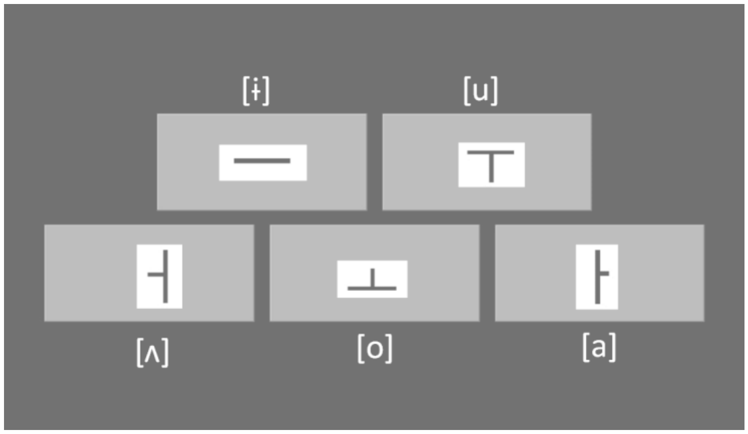

Figure 1. Screen display for the forced choice vowel identification task.

Auditory lexical decision task. This task was designed to compare the phonolexical representations for older and younger words in the same participants. Two vowel contrasts of Korean were used: /o/-/ʌ/ (test) and /o/-/a/ (control). For each contrast, stimuli were 16 words designated as "old" (i.e. likely learned a long time ago), and 16 words designated as "young" (likely learned more recently) based on Korean textbooks. Each word (8 for each vowel within the contrast) was modified to create a paired nonword by switching the vowel (e.g. /o/ for /ʌ/ and vice-versa). This resulted in 128 experimental items (64 items per contrast: 32 test, 32 control), to which 160 distractors were added. Stimuli were split into two lists, such that a word and its paired nonword never appeared in the same list. Lists were assigned randomly to participants. Each participant only heard one of the two lists, to avoid priming a response by presenting the paired item in the other list as well. The list of stimuli is presented in the Appendix A1. Since old and young items necessarily were different words, efforts were made to keep them phonologically as comparable as possible with respect to length and segments: for each old word, a young word was chosen that was as similar as possible. Table 1 displays examples of these old and young word pairs in both conditions (test and control). They were not presented as pairs in the task.

Table 1

Example stimuli for old and young words in each condition

Condition	**Old word**	(gloss)	(nonword)	**Young word**	(gloss)	(nonword)
Test	tɕ**ʌ**njʌk̚	*dinner*	tɕ**o**njʌk̚	tɕ**ʌ**nsɛ	*security deposit*	tɕ**o**nsɛ
Test	tɕ**ʌ**nhwa	*telephone*	tɕ**o**nhwa	tɕ**ʌ**npʰa	*propagation*	tɕ**o**npʰa
Test	s**o**gɛ	*introduce*	s**ʌ**gɛ	s**o**dɨk̚	*income*	s**ʌ**dɨk̚
Test	m**o**du	*all*	m**ʌ**du	m**o**sun	*contradiction*	m**ʌ**sun
Control	k**a**k*ɨm	*sometimes*	k**o**k*ɨm	k**a**nɨm	*estimate*	k**o**nɨm
Control	tɕʰ**o**rok̚	*green*	tɕʰ**a**rok̚	tɕʰ**o**biŋ	*invitation*	tɕʰ**a**biŋ

All stimuli were recorded in a sound-isolated recording booth by a phonetically trained female native speaker of Korean from Seoul. Each individual item was saved into a separate sound file for presentation by the stimulus presentation software; all files were normalized for amplitude. Participants were asked to listen to each item and decide if they heard a real word or not. They indicated their response by pressing buttons on a computer keyboard. An accurate response in this task is to say “yes” to words, and “no” to nonwords. Saying “no” to a nonword, which is always a possible word, is only possible if the phonolexical representation of the word is precise, and if the phonological difference between the word and the nonword is perceived. If – for example – a learner does not know exactly which vowel is supposed to be in the word /s**o**gɛ/ “introduce”, then, s/he might think that /s**ʌ**gɛ/ (the nonword) is actually the real word. Therefore, this task allows us to probe the form of lexical representations for these words.

All items in a given list were presented auditorily through high quality headphones in a random order, only once. As soon as participants made their answer, the next item was presented after a brief delay. The task was not speeded. Prior to the test phase, participants were given 10 practice trials with feedback. Stimuli presentation was controlled by the software OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). Accuracy and RT were measured, but only accuracy is used as a dependent variable in this paper.

Procedure. All procedures were approved by the Indiana University Review board. First, participants took part in the lexical decision task, followed by the vowel identification task. At the end of the experiment, each participant filled out a background questionnaire as well as a word-familiarity and learning history questionnaire, asking learners whether they knew each word, and when they thought they had learned it (Appendix A2). The entire testing session lasted about 45 mins. Participants were tested in a quiet computer room on a South Korean University campus, and were paid for participating.

RESULTS

All trials containing nonwords based on a word the listener reported not knowing (16.2% of the trials) were excluded. We also re-coded word age if, for example, a word we designated as “old” was reported by a listener to have been learned more recently. Thus, we obtained a listener-specific coding of word age, which we used in the analysis. In order to get interpretable datasets, it was necessary to exclude participants who made too many errors on the lexical decision distractors. In addition, because the coding of word-age was learner specific, some learners reported knowing too few words in certain conditions, for instance, only having two “young” items. A minimum of 5 trials in any given condition was the criterion we used for inclusion. After this exclusion procedure, the final participant sample were 13 native speakers of Mandarin who were learning Korean as L2.

Vowel identification task

Table 2 presents the rates of correct identification in the vowel identification task for each learner. Overall, all vowels were identified with relatively high accuracy, and nine out of 13 participants obtained overall average scores higher than 80% correct. When looking at the critical test contrast, we derive the accuracy in Table 2 from the number of trials on which /o/ was misperceived as /ʌ/

and vice versa. Thus, a score of 85% correct means that in 15% of the /o/ and /ʌ/ trials, the vowel was confused with the other. Participants were overall accurate in identifying these two vowels (average of 87.3% correct). Ten out of 13 participants scored 80% correct or above for this test contrast.

Table 2

Accuracy (%) in the vowel identification task (test vs. all contrasts) for each participant

Learner	27f1	25f1	18f2	22f2	13f1	05f1	24f2	09m1	11f1	10f2	07f1	14f2	21f1
/o/-/ʌ/	65	68.3	75	80	85	86.7	88.3	93.3	95	98.3	98.3	100	100
All vowels (incl. /o/-/ʌ/)	61.7	60	70	78.3	80	85	85	88.3	91.7	93.3	91.7	88.3	93.3

Auditory lexical decision task

In the current analysis, we defined "accuracy" as the nonword rejection rates in each trial type (test vs. control). Generally, we expected an effect of trial type, that is, higher nonword rejection rates for control items compared to test items. This first prediction was confirmed. A mixed effects logistic regression model with fixed effects of *word age* and *trial type* revealed a significant effect of trial type, where test items were less accurate than control items; $\beta = 1.17$, $p < .001$).

Second, we examine any potential effects of word age by comparing responses for items based on old words vs. young words (declaring the variable of word age in the regression model). For the test contrast, accuracy for nonwords based on old words was on average 13.7% lower than for those based on young words, whereas for the control contrast, this difference was much smaller: accuracy for items based on old words was 6% lower than for items based on young words. Figure 2 shows the mean accuracy in each trial type and word age: control vs. test, and old vs. young.

Descriptively, the hypothesis that word age matters for updates is supported. Statistically however, it was not. The tendency observed in the test trials is in the predicted direction: Nonwords based on young words are rejected more successfully than nonwords based on old words. This might indicate that young words are easier to update (to correct), and therefore nonwords based on these are easier to reject. However, this trend did not produce a significant interaction between age and condition, possibly due to the small sample size and the fact that variability between participants is large. The mixed effects logistic regression model revealed a significant effect of trial type on accuracy scores (test < control), but no significant effect of word age, and no significant interaction between age and trial type. Together with the correlation presented above, the absence of a statistically significant effect of word age tentatively supports the phonological update hypothesis, in which individual lexical forms appear to be updated wholesale as learners learn more about the phonological system of their L2.

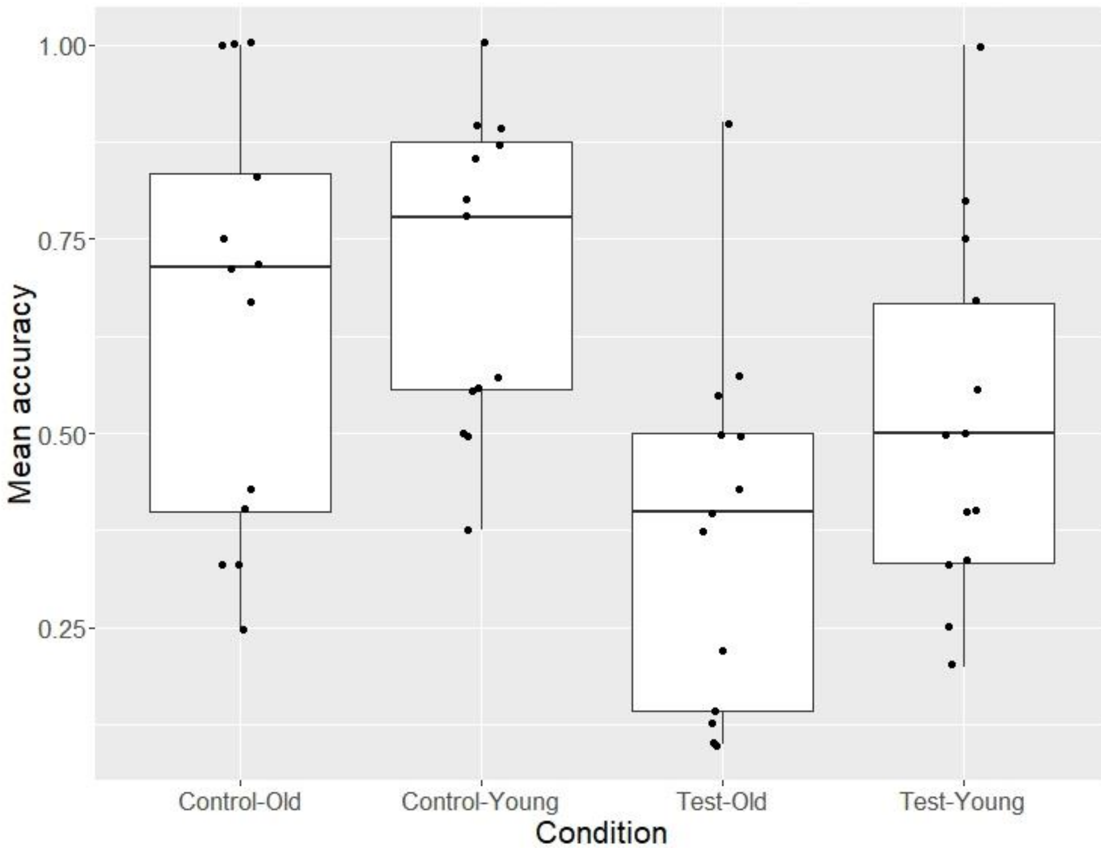

Figure 2. Boxplot/strip chart for mean accuracy (nonword rejection rate) in the four experimental conditions (trial type x word age). Each dot represents one participant (n = 13).

Finally, we examined the link between vowel identification and lexical decision accuracy on test items (across both old and young items). A two-tailed non-parametric Spearman correlation showed that vowel identification accuracy was positively correlated with accuracy in lexical decision ($r_s = .71$, n = 13, $p = .007$). The scatterplot in Figure 3 shows the relationship between the two measures. The red square across the top highlights that even at very high levels of identification accuracy for this contrast, performance on the lexical decision task is very variable (ranging from about 40% to 95% correct), thus again indicating that even near perfect identification does not guarantee lexical accuracy.

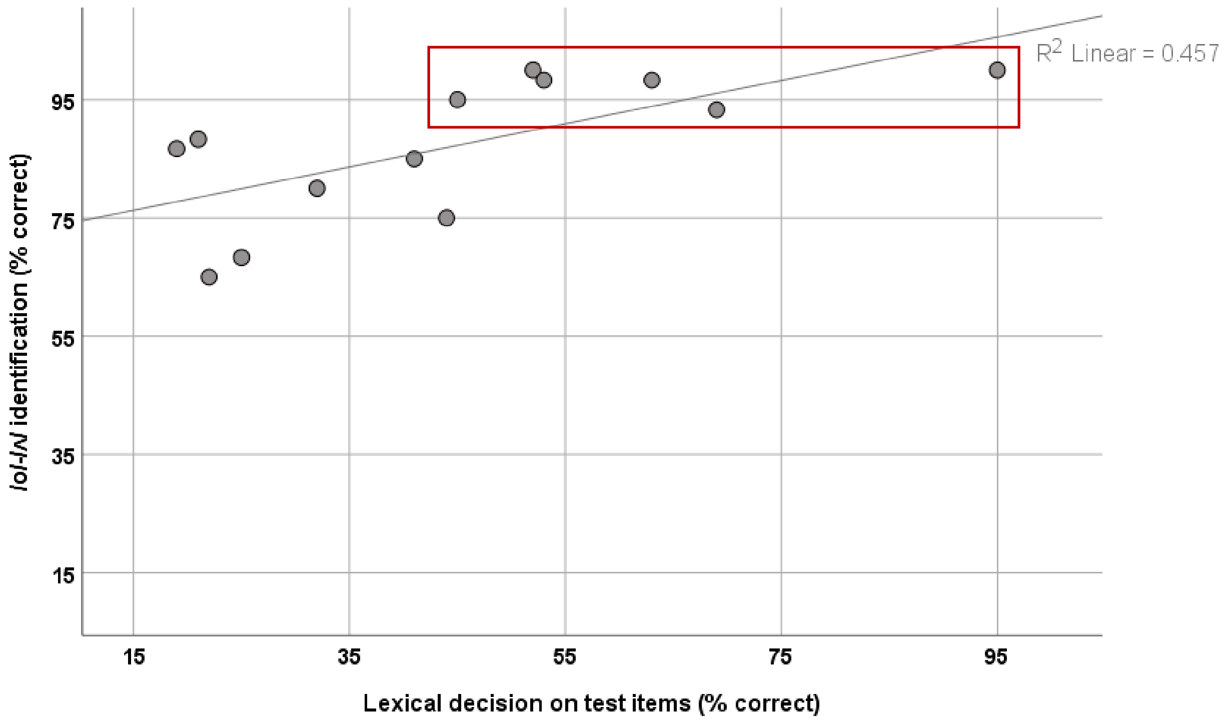

Figure 3. Scatterplot showing the relationship between identification and lexical decision scores. Each dot is one participant.

In the global regression analysis above, all participants were included. Yet, word age effects, if any, are expected to be emerging mainly for those participants who have acquired the contrast successfully, because in the word age hypothesis, any effects *depend on* perceptual mastery of a contrast. Therefore, it is important to examine this pattern in participants with an excellent perceptual mastery of the vowel contrast. Figure 4 shows the relationship between perceptual identification accuracy for the test vowels and the lexical decision scores on the test items, split by word age (young: empty vs. old items: filled). Participants are ranked on the x-axis from least to most accurate in the perceptual identification task for this contrast (%correct). We found that listeners who confused the contrast more often were also significantly more likely to incorrectly accept nonwords based on *both* old and young words (both $R^2 = 0.352$, $p < .02$), perhaps suggesting that in their mental lexicon, words do not encode the contrast reliably, regardless of word age. However, a difference between old and young test items emerges more reliably for those participants with the highest mastery of the contrast: for all participants with a perceptual identification score above 90%, young items are responded to more accurately than old ones. This relationship was not visible in the control contrast.

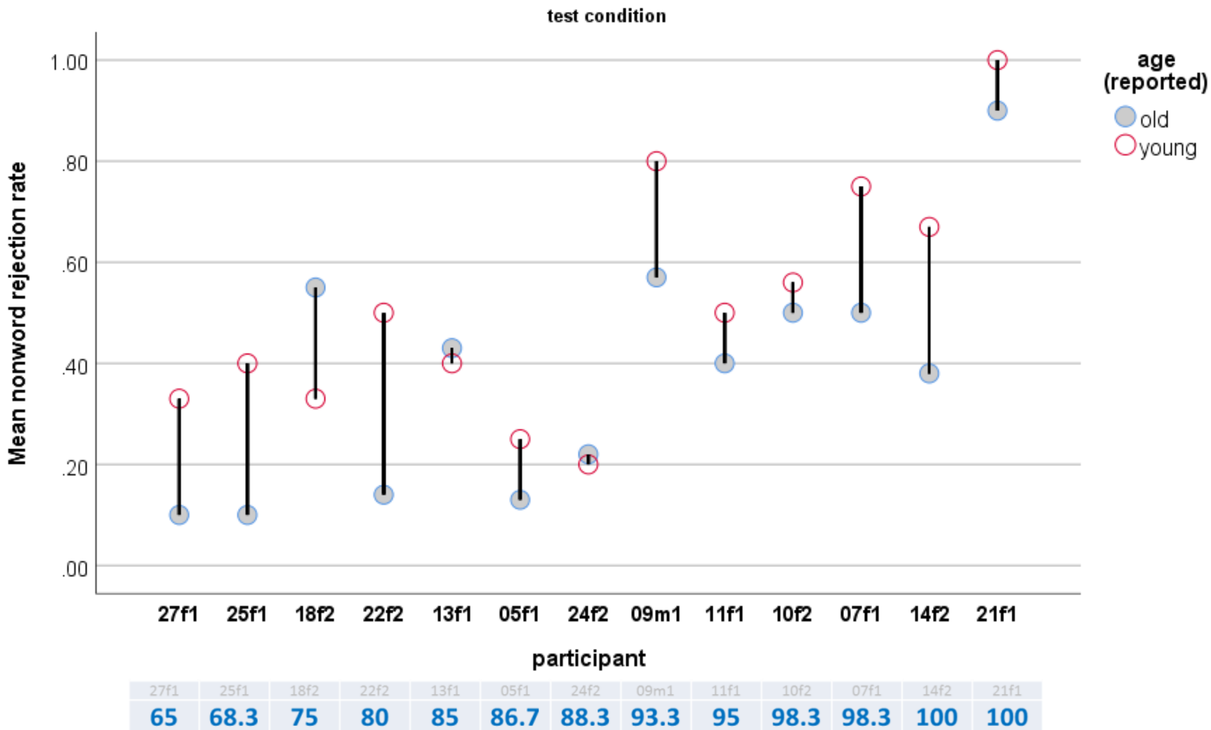

Figure 4. Relationship between perceptual identification accuracy for the test vowel and the lexical decision scores for test items, split by word age. The bottom numbers indicate individual identification accuracy (%) for the test contrast, ranked from lowest to highest.

DISCUSSION

The absence of a statistically significant effect of word age tentatively supports the phonological update hypothesis. Phono-lexical representations appear to be updated wholesale as learners learn more about the phonological system of their L2, given the fact that we observed a correlation between perception accuracy and nonword rejection rate for both young and old items. Yet, a difference by word age becomes most clearly visible in participants with the highest mastery of the vowel contrast. Thus, we cannot fully reject the possibility that an effect of age might emerge. However, at this stage this effect is small and must remain tentative.

Despite the preliminary nature of these findings, they suggest that potentially, once a contrast is acquired and part of a learner's phonological knowledge, it is first represented in the phonolexical representation of more recent words, and little by little, may reach more entrenched representations. The limited sample size of this investigation limits the conclusions we can draw, but our exploratory study raises interesting questions.

The most important question raised by our findings concerns the inventory of lexical representations in the bilingual mental lexicon. It is indeed possible that phonolexical representations are not actually updated (corrected), but rather that new ones are added to previously existing ones, and co-exist for a time. Accordingly, the earlier representations remain accessible during word recognition. One benefit of this kind of scenario is that word recognition would be facilitated for both target-like input as well as for non-target-like input (which follows the L1 phonological specifications). Concretely, this means that an L2 learner who builds a new accurate representation next to an earlier one may be as fast to recognize the word when spoken by a native speaker as when spoken by another L2 learner who confuses the same contrast in their production. In the long term however, L2 learners ideally need to inhibit the activation of – or even fully erase – the earlier phonolexical representations. Whether learners are able to do this at all remains a question for future research.

Our findings further indicate that older lexical representations may be more resistant to updates than recent lexical representations. If this finding is confirmed, it will need to be reconciled with other data showing that more entrenched lexical representations also lead to higher accuracy and less variability in specific lexical tasks (e.g. Cook et al., 2016). Our results suggest the opposite. However, the two possibilities need not be incompatible: they may derive from different tasks targeting different phonological contrasts and proficiency levels. In particular, findings such as ours are likely to apply mostly to difficult phonological dimensions, whereas Cook and colleagues obtained data across a wider range of phonological dimensions, not all of them being perceptually challenging.

Finally, our findings relate to L2 pronunciation in two main ways. The first is that globally, lexical updates appear to rely on improvements in phonological knowledge; yet, even among the learners who have mastered the nonnative vowel contrast in perception, accurate or updated lexical representations were not guaranteed. If these findings are indeed solid (similar findings were obtained in Simonchyk & Darcy, 2017), they lead us to assume a kind of 'hierarchy' in the order of acquisition such that phonological knowledge of sounds may precede phonolexical updates, and not the other way around (that is, in L2, phonological knowledge may not emerge from inferences over the mental lexicon, a scenario that has been proposed for L1 acquisition by, among others, Munson, Edwards and Beckman [2005, p. 198]: "Higher level phonological knowledge emerges as a consequence of word learning and serves to facilitate future word learning"). While this assumption clearly needs to be fully tested for L2 learners, given its relevance for instruction, it also begs the question of why updating lexical representations lags behind improvements in phonological knowledge. One possible answer is that the kind of perceptual improvements underlying accurate identification in a task such as ours are not linked to actual word representations as they occur, thus hindering a direct influence on these word representations. Another possible answer is that updating lexical entries takes more time or more directed effort. Thus, teachers should be aware of this potential discrepancy in learners' productions, which are likely to reflect fuzzy lexical representations for a long time. This would not necessarily mean that learners have not yet acquired the corresponding phonological knowledge, it may just indicate that some words have not yet been updated. Clearly, the precise mechanisms by which perceptual and phonolexical improvements take place as well as their time course are yet to elucidate. But there is potentially a role for pronunciation instruction to help bridge the gap between perception and lexical representations, and this is the other way in which our findings relate to pronunciation instruction.

Our findings highlight the need for integrating phonological knowledge into lexical knowledge. Importantly, we are not suggesting that teachers should first ensure learners have acquired the phonological system of the L2 before they start learning words. This approach would not only be unrealistic, it may also be counterproductive, since in both L1 and L2, learning words might still be a driving force in developing the phonological system in the first place (Munson et al., 2005). A much more effective approach would be to incorporate pronunciation instruction about words at the earliest possible time by integrating new vocabulary teaching with pronunciation, and by revisiting known words often while focusing on their phonological form. This approach might prove successful in preventing too many fuzzy lexical representations from getting set up in the first place.

ACKNOWLEDGMENTS

The authors wish to thank Zoie Hancock and Joshua Lee for their help with the stimuli creation. We also thank audience members at the 10th PSLLT conference in Ames, IA and at LabPhon 16 in Lisbon, Portugal, for stimulating discussion.

ABOUT THE AUTHORS

Isabelle Darcy is associate professor of second language psycholinguistics in the Department of Second Language Studies at Indiana University. She obtained a PhD in Linguistics and Cognitive Science from the Ecole des Hautes Etudes en Sciences Sociales in Paris (France) and from the Johannes Gutenberg University in Mainz (Germany). Her research focuses on the acquisition of second language phonology, pronunciation instruction, and on L1/L2 speech perception and word recognition.

Indiana University
Department of Second Language Studies
Morrison Hall 231
1165 E. 3rd St.
Bloomington, IN 47405, USA
Phone: +1 812 855 0033
E-mail: idarcy@indiana.edu

Jeffrey J. Holliday is assistant professor in the Department of Korean Language and Literature at Korea University in Seoul, South Korea. He obtained a PhD in Linguistics from The Ohio State University. His research interests include laboratory phonology, perception and production of Korean as a second language, socio-phonetics, and individual differences.

Department of Korean Language & Literature
Korea University
145 Anam-ro Seongbuk-gu
Seoul 02841
South Korea
E-mail: holliday@korea.ac.kr

REFERENCES

Bradlow, A., Akahane-Yamada, R., Pisoni, D., & Tohkura, Y. i. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, 61, 977-985. doi:10.3758/BF03206911

Cook, S. V., Pandža, N. B., Lancaster, A. K., & Gor, K. (2016). Fuzzy Nonnative Phonolexical Representations Lead to Fuzzy Form-to-Meaning Mappings. *Frontiers in Psychology, 7*(1345). doi:10.3389/fpsyg.2016.01345

Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon, 8*, 372-420.

Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., & Scott, J. H. G. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English-L2 French acquisition. *Second Language Research, 28*, 5-40. doi:10.1177/0267658311423455

Darcy, I., Park, H., & Yang, C.-L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences, 40*, 63-72. doi:http://dx.doi.org/10.1016/j.lindif.2015.04.005

Darcy, I., & Thomas, T. (2019). When *blue* is a disyllabic word: Perceptual epenthesis in the mental lexicon of second language learners. *Bilingualism: Language and Cognition*, 1-19. doi: 10.1017/S1366728918001050 (first view).

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*, 682-706.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.

Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics, 36*, 141-157.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. doi:10.3758/s13428-011-0168-7

Munson, B., Edwards, J., & Beckman, M. E. (2005). Phonological knowledge in typical and atypical speech-sound development. *Topics in Language Disorders, 25*, 190-206.

Nagle, C. L. (2018). Examining the temporal structure of the perception–production link in second language acquisition: A longitudinal study. *Language Learning, 68*, 234-270. doi:10.1111/lang.12275

Ota, M., Hartsuiker, R. J., & Haywood, S. L. (2009). The KEY to the ROCK: Near-homophony in nonnative visual word recognition. *Cognition, 111*, 263-269.

Pallier, C., Bosch, L., & Sebastian-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64,* B9-B17.

Pallier, C., Colomé, A., & Sebastian-Gallés, N. (2001). The influence of native-language phonology on lexical access: exemplar-based versus abstract lexical entries. *Psychological Science, 12*, 445-449.

Simonchyk, A., & Darcy, I. (2017) Lexical encoding and perception of palatalized consonants in L2 Russian. In M. O'Brien & J. Levis (Eds). *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Calgary, AB, August 2016 (pp. 121-132). Ames, IA: Iowa State University.

Simonchyk, A., & Darcy, I. (2018) The Effect of Orthography on the Lexical Encoding of Palatalized Consonants in L2 Russian. *Language and Speech, First Published March 18, 2018*. doi: 10.1177/0023830918761490

Veivo, O., & Järvikivi, J. (2013). Proficiency modulates early orthographic and phonological processing in L2 spoken word recognition. *Bilingualism: Language and Cognition, 16*(04), 864-883. doi:10.1017/S1366728912000600

Wayland, R., & Guion, S. (2003). Perceptual discrimination of Thai tones by naive and experienced learners of Thai. *Applied Psycholinguistics, 24*, 113-129.

APPENDIX

Table A1

List of stimuli used in the lexical decision task

	Test /o/ words					
	Old words			**Young words**		
	IPA	**(gloss)**	**(nonword)**	**IPA**	**(gloss)**	**(nonword)**
1	norɛ	*song*	nʌrɛ	nodoŋ	*labor*	nʌdoŋ
2	modu	*all*	mʌdu	mosun	*contradiction*	mʌsun
3	koŋbu	*study*	kʌŋbu	koŋdʑi	*notification*	kʌŋdʑi
4	koŋtɕ*a	*free*	kʌŋtɕ*a	koŋhak̚	*engineering*	kʌŋhak̚
5	totɕʰak̚	*arrive*	tʌtɕʰak̚	tosʌ	*book*	tʌsʌ
6	toŋsɛŋ	*younger sibling*	tʌŋsɛŋ	toŋtɕʰaŋ	*alumnus*	tʌŋtɕʰaŋ
7	sogɛ	*introduce*	sʌgɛ	sodɨk̚	*income*	sʌdɨk̚
8	onɨl	*today*	ʌnɨl	onsu	*hot water*	ʌnsu
	Test /ʌ/ words					
1	tɕʌnjʌk̚	*dinner*	tɕonjʌk̚	tɕʌnsɛ	*security deposit*	tɕonsɛ
2	mʌndʑʌ	*first*	mondʑʌ	mʌŋgɛ	*sea squirt*	moŋgɛ
3	pʌls*ʌ	*already*	pols*ʌ	pʌltɕʰik̚	*penalty*	poltɕʰik̚
4	ʌlgul	*face*	olgul	ʌlluk̚	*stain*	olluk̚
5	tɕʌnhwa	*telephone*	tɕonhwa	tɕʌnpʰa	*propagation*	tɕonpʰa
6	tɕʰʌŋso	*clean*	tɕʰoŋso	tɕʰʌŋtɕʰun	*adolescence*	tɕʰoŋtɕʰun
7	ʌnni	*older sister*	onni	ʌndʌk̚	*hill*	ondʌk̚
8	tɕʌmɕim	*lunch*	tɕomɕim	tɕʌmtɕʰa	*gradually*	tɕomtɕʰa
	Control /o/ words					
	Old words			**Young words**		
	IPA	**(gloss)**	**(nonword)**	**IPA**	**(gloss)**	**(nonword)**
1	komin	*trouble*	kamin	kotɕʰɨŋ	*high floor*	katɕʰɨŋ
2	koguŋ	*palace*	kakuŋ	kogal	*depletion*	kagal
3	mok̚tɕʌk̚	*goal*	mak̚tɕʌk̚	mok̚s*um	*life*	mak̚s*um
4	motɕip̚	*recruit*	matɕip̚	mobʌm	*model*	mabʌm
5	momsal	*flu*	mamsal	momtɕit̚	*gesture*	mamtɕit̚
6	tɕʰorok̚	*green*	tɕʰarok̚	tɕʰobiŋ	*invitation*	tɕʰabiŋ
7	onmom	*whole body*	anmom	onɕil	*heated room*	anɕil
8	otɕ*aŋ	*closet*	atɕ*aŋ	ogok̚	*five grains*	agok̚
	Control /a/ words					
1	kak*ɨm	*sometimes*	kok*ɨm	kanɨm	*estimate*	konɨm
2	kasɨm	*chest*	kosɨm	kamum	*drought*	komum
3	nampʰjʌn	*husband*	nompʰjʌn	namgɨk̚	*south pole*	nomgɨk̚

4	tanpʰuŋ	*maple*	tonpʰuŋ	tansʌ	*cue*	tonsʌ
5	tap̚tɕ*aŋ	*reply*	top̚tɕ*aŋ	tap̚ɕ*in	*reply (formal)*	top̚ɕ*in
6	annɛ	*information*	onnɛ	angɛ	*fog*	ongɛ
7	tɕʰaŋmun	*window*	tɕʰoŋmun	tɕʰaŋdʑo	*creation*	tɕʰoŋdʑo
8	hanɨl	*sky*	honɨl	hatɕʰɛ	*lower body*	hotɕʰɛ

Note. Darker shading are items assigned to list 1, those with no shading to list 2.

Date____________________ IRB Study #__________

Participant _________________
N° Group _________________

When did you learn these words?

Please estimate when, during your Korean learning experience, you have learned these words:
* If the word is among those you learned first → please check box #1: "In my first year"
* If you have learned this word some time ago, but not in your first year → check #2 "in-between"
* If you have learned the word not very long ago → check #3: "quite recently"

Item	In my first year	in-between	quite recently		Item	In my first year	in-between	quite recently
song					labor			
all					notification			
study					engineering			
free					book			
arrive					alumnus			
younger sibling					income			
introduce					hot water			
today					security deposit			
dinner					sea squirt			
first					penalty			

Figure A2. Sample questionnaire items used to estimate listener-specific word age.