Applied Psycholinguistics **36** (2015), 115–128
doi:10.1017/S0142716414000459

DISCUSSION ARTICLE

Representation of second language phonology

ANNE CUTLER
University of Western Sydney, Max Planck Institute for Psycholinguistics, Nijmegen, and Radboud University Nijmegen

ADDRESS FOR CORRESPONDENCE
Anne Cutler, MARCS Institute, University of Western Sydney, Locked Bag 1797, Penrith South, NSW 2751, Australia. E-mail: a.cutler@uws.edu.au

ABSTRACT
Orthographies encode phonological information only at the level of words (chiefly, the information encoded concerns phonetic segments; in some cases, tonal information or default stress may be encoded). Of primary interest to second language (L2) learners is whether orthography can assist in clarifying L2 phonological distinctions that are particularly difficult to perceive (e.g., where one native-language phonemic category captures two L2 categories). A review of spoken-word recognition evidence suggests that orthographic information can install knowledge of such a distinction in lexical representations but that this does not affect learners' ability to perceive the phonemic distinction in speech. Words containing the difficult phonemes become even harder for L2 listeners to recognize, because perception maps less accurately to lexical content.

The phonology of the native first language (L1) is largely acquired during the first year of life (Kuhl et al., 2008), and that alone tells us that the source of information that supports its learning is spoken input only. In particular, this applies to the set of phonetic contrasts relevant for distinguishing words of the L1 vocabulary: the full set is to a significant extent in place as the earliest word forms are being stored in the learner's personal lexicon (Saffran, Werker, & Werner, 2006).

In contrast, most second language (L2) learners, certainly those who are the concern of the present Special Issue, have an extensive native vocabulary already in place. Moreover, they have acquired additional relevant skills, in particular reading (also social skills and explicit learning skills). All this makes it clear that L2 learners, in principle, have multiple options for finding information about the newly experienced phonology. They can ask teachers or other informants outright (*a hammer* and *an 'ammer*—is that the same thing?), they can elicit information indirectly through conversation, they can look it up in textbooks and grammars, and they can acquire it from experience with the language, not only listening experience, but also reading, whatever the topic or material. Written language contains a wealth of information about segmental phonological regularities (and in English, this information is certainly of use to L1 users who have been deaf for their entire life; Hanson, 1982; Hanson, Shankweiler, & Fischer, 1983).

Experience with the L1 leads to expectations about which concepts should map to words (does *we* mean both "you and I" and "my partner and I but not you"?), about what words should be like and what internal structure they may have, and of course about which phonetic contrasts are relevant for distinguishing words. Each such expectation may or may not be met in an L2. Where L1-based expectations mismatch with the L2, fully or partially erroneous representations can arise. The research questions for applied psycholinguists are then how such representations are formed, what effect they have on L2 proficiency, and how the representations used by L2 learners may best be brought into line with the equivalent representations of native speakers.

TESTING THE NATURE OF L2 PHONOLOGICAL REPRESENTATIONS

In psycholinguistic studies of spoken-word recognition (by L1 or L2 listeners), methods combining spoken and visual input enable examination of what words listeners think they are hearing in spoken input. One such widely used method is eye tracking, in which listeners' involuntary looks to elements of a display reveal moment-by-moment availability of alternative interpretations of the speech that they are processing. Cutler, Weber, and Otake (2006) and Weber and Cutler (2004) used this method to examine the nature of the lexically stored phonological representations for certain words that L2 listeners appear not to be able to accurately perceive. These two reports were part of a larger series of studies targeted at mapping the range of processing consequences of phonemic confusions for L2 listeners. There is an extremely large literature on the issue of phonemic confusions, especially the single category case (Best & Tyler, 2007), in which a contrasted pair of sounds in the L2 maps to a single sound of the native language. Confusion in such cases is both part of the public consciousness about certain foreign accents (to confirm this for the English-speaking consciousness, one need only conduct a web search on *flied lice*) and a cause of much agonizing about accent among L2 learners (consider the many web sites devoted to r/l remedial attempts). However, is such agonizing justified? Does it really matter if listeners misperceive an occasional speech sound? Should it usually not be possible to recover the intended meaning because one of the two word options would be a far better fit than the other to the utterance topic?

The outcome of the research program was a clear answer: it does matter. However, that is not because of anything to do with accent in speech production or with minimal word pairs such as *fly* and *fry* or *rice* and *lice*. Languages abound in homophones (words that sound the same but have different meanings), and adding a few more homophones to the lexicon because of an inability to distinguish between minimal pairs will really make very little difference to the listener's task. We all have to refer to the utterance context to decide whether a speaker has said *mail* or *male*, *paste* or *paced*, and *rain*, *rein*, or *reign*, or to decide which meaning of *mouse*, *spot*, or *condition* is the appropriate one in a particular case.

Languages do not avoid homophones; in contrast, they seem to prefer to label new concepts with old names rather than new ones. My computer has a mouse, and it can access the web, and it seems to bother no one that *mouse* and *web* still also retain the old meanings they had before computers came along. The result is that word forms with multiple meanings are simply the normal case for vocabularies. If all the minimal pairs of English words differing only in /r/ versus /l/ were to be interpreted as containing /r/, then 287 homophones would be added to the vocabulary; if they were all interpreted as containing /l/, 311 new homophones would result (Cutler, 2005). (The numbers are not quite symmetrical because some pairs contain both sounds; *rightly* and *lightly* are a minimal pair, but turning all /l/ to /r/ produces *rightry*, a nonword, while turning /r/ to /l/ makes them both the real word *lightly*.) These few hundred extra homophones are a trivial matter in the context of vocabularies of tens of thousands of words, a very large proportion of which have multiple meanings.

However, a much more serious problem is caused by temporary overlap among words. Spoken-word recognition is a process that is both rapid and efficient, mainly because our brain's processing system does not wait till a word has been fully spoken, but generates possible word candidates to be compared against incoming input. The more ambiguous the input, the more competing interpretations there are, at least for a while, and it is known that word recognition is slowed by an increase in the number of such competitors (Norris, McQueen, & Cutler, 1995). In L2 listening, extra competition is caused by words that would not cause such competition for L1 listeners (Broersma, 2012; Broersma & Cutler, 2008, 2011), and the scale of possible occurrences of such temporary overlap far outstrips the impact of the potential homophones (Cutler, 2005). Thus, for /r/–/l/ confusions, once again, *leg* will compete when the input is actually *regular*, *crow* will compete when the input is *clothing*. Both of these are short words that happen to be fully embedded in longer words if the phonemic minimal pair is overlooked; but temporary overlap also includes cases such as *televise/terrify*, *breathing/bleeding*, *risen/lizard*, and literally thousands more. This is where phonemic confusion causes real problems for L2 spoken-word recognition. In all listening, mismatching input has an almost instant effect of reducing the set of potential competitors. However, *leg* mismatches *regular*, and *terr-* mismatches *tel-*, *only* if the phonemic difference is registered by the listener. Thus, these inputs mismatch for L1 listeners; but L2 listeners, who overlook the difference, have no benefit from this mismatch and must wait until a later mismatch arrives that can be detected. As these cases indicate, that may be one, two, or several phonemes later.

The empirical question posed in the eye-tracking studies concerned the nature of the lexical representations accessed for words containing such confusable sounds; if a listener cannot tell the difference between two L2 sounds, are the listener's lexical representations of L2 words containing those sounds phonologically identical in the relevant respect? Again, minimal pairs are the minority of this: are the stored representations of *write* and *light* phonologically identical, that is, effectively homophones, just as *rain*, *rein*, and *reign*? However, the question addresses all the overlap cases as well: do the representations of *leg* and *regular* contain identical initial sounds? If so, does the sound map more closely to one or the other of the two L2 phonemes, or does it embrace both sounds equally well?

The results of Weber and Cutler's (2004) and Cutler et al.'s (2006) experiments, conducted, respectively, with Dutch learners of English (who confuse the vowel sounds in *had* vs. *head*) and with Japanese learners of English (who confuse the initial consonants of *write* vs. *light*), were quite surprising. They clearly showed that the lexical representations could not be identical; they must be distinct. In the experiments, listeners saw a display that included pictures with names that contained the confusable sounds, for example, a <u>pan</u>da and a <u>pen</u>cil for the Dutch listeners, or a <u>rock</u>et and a <u>lock</u>er for the Japanese participants. They were instructed to click on one of the objects, and their eye movements were monitored as they did so, so as to provide information on the stored representations that were being activated by the speech input they heard.

If the lexical representations of the two sounds were identical, then the incoming confusable speech portions (*pan-*/*pen-* or *rock-*/*lock-*) should contact both of the lexical entries, just as homophones such as *mail*/*male* activate both (semantically distinct) lexical entries that they match (Grainger, Van Kang, & Seguí, 2001). However, they did not. Each listener group showed a tendency to look at only one of the phonological interpretations. The Dutch listeners preferred the options with the /ɛ/ sound, and looked more at the pencil whether they heard *pen-* or *pan-*. The Japanese listeners preferred the options with the /l/ sound, and looked more at the locker whether they heard *lock-* or *rock-*. The direction of the preference can be easily explained as choice of the L2 sound that has better goodness of fit to the single L1 category. However, what cannot be explained so easily is why this preference appears not to have carried through to the lexical representations. If *pan-* is perceived as *pen-*, why is the lexical representation of the animal's name not stored as an equivalent to *penda*? If *rock-* is perceived as *lock-*, why is a rocket not assigned a phonological representation with /l/? In neither language is this the case, because the experiments showed that looks went to the objects the names of which really did have the preferred sound: the pencil and the locker. How do these listeners know which word is supposed to contain the sound they are choosing to interpret the input as?

The distinctions in the lexicon have clearly not come from these listeners' speech perception, because perception delivers essentially the same result for each phoneme in the respective minimal pair. In speech, each sound is being preferentially interpreted as the option closest to the L1 category, and this goes for whichever L2 category the speaker intended. The learners' lexicons must therefore have been affected by direct information in each case that a distinction between the two categories exists. Such information about a practically imperceptible but in principle necessary L2 distinction probably came from teachers or from textbooks. It is abstract knowledge, drawn not from listening experience but from metalinguistic sources. That is, because *panda* is spelled *panda* (not *penda)*, and *rabbit* is spelled *rabbit* (not *labbit*), the critical sound in the phonological content of each lexical entry becomes labeled as <u>not</u> the one closest to the L1 category.

Recall that neither of these two studies had any element of orthographic input, and the experimental question in no way concerned the issue of orthographic knowledge. However, with respect to the storage in the lexicon of abstract knowledge about a distinction that cannot be reliably perceived, it is surely likely that listeners are helped in this way because each distinction is coded orthographically. Escudero, Hayes-Harb, and Mitterer (2008) subsequently showed that provision of orthographic information about nonwords such as *tenzer* and *tandik* resulted in distinct representations of the initial syllables for Dutch listeners, whereas without the orthographic information the initial syllables seemed to be treated as homophonous. This is further evidence that auditory perception of the spoken items is not the source of the lexical distinction and that orthographic information alone, without accompanying teacher exposition, can cause the distinction to figure in the phonological representations that are lexically stored.

In the results from the eye-tracking studies, the role of orthographic knowledge can only have been indirect: it may have been the source of information that allowed an imperceptible distinction to be lexically coded. Orthography cannot have played a role in the listeners' responses (e.g., by causing listeners to look at the item with the name that would be written with the letter used for the single L1 sound), because the two language groups' results patterned inversely. In Dutch, the single L1 sound in the æ/ε vowel space is written with the letter *e*, and this corresponds to the letter used in English in the names of the items looked at by preferences (such as *pencil*); but in Japanese, the single L1 sound in the r/l space is transliterated in English with *r* (consider *tempura* and *Kirin*), but the items looked at by preference were those spelled with *l* (such as *locker*; note that the subjects in the Japanese study were students and hence must have been familiar with Romanji transliterations of their native words). Goodness of fit of the two L2 phonemes to the L1 sound certainly differs in the way reflected in the results (e.g., Japanese listeners rate English /l/ as closer than English /r/ to their native rhotic; Iverson et al., 2003); goodness of fit thus seems to be the better fitting and sufficient explanation of why there is a preference. Orthography's role is played out in influence on the initial construction of the abstract phonological representation in the lexical entry.

Phonological representations in a language learner's lexicon therefore are not compiled solely from an accumulation of speech perception episodes. They are influenced also by nonspeech/metalinguistic information such as instruction that two sounds are supposed to be distinct, not identical. This abstract distinction in the resulting representations then controls whether each representation is accessed during phonemic interpretation of speech input. When the interpretation delivers the preferred or dominant phoneme (which is almost always), the representations coded as containing that preferred phoneme are contacted. Because the interpretation virtually never delivers the nonpreferred phoneme, the representations coded as containing that nonpreferred phoneme are virtually never contacted.

LEXICAL REPRESENTATION WITHOUT PERCEPTION: HELP OR HINDRANCE FOR THE LEARNER?

In L1, the lexicon and input match: the contrasts that are phonologically represented in the lexicon are those that are perceptually distinguished in the input. However, what we see here is a mismatch between input and lexicon for listeners attempting to acquire these difficult contrasts. What implications follow from this for the learner of an L2?

As described by Escudero (2015 [this issue]), the inclusion in the lexicon of a contrast that is indistiguishable in perception has been viewed in some literature as evidence that orthographic information (the implied source of the distinction) can assist the L2 learner. On the face of it, this seems a plausible interpretation: if the phonemic distinction is lexically coded, then the phonological representations in the learner's lexicon are more in line with those of native listeners, which is certainly desirable. In addition, distinct lexical representations may prompt listeners to attempt a distinction in speech production, making the L2 speaker (if the attempt is successful) more intelligible to native listeners; and addition of unwanted homophones to the learner's stored vocabulary can be avoided. Of course, as pointed out above, homophones are not really a problem for any listener in any language, first or second: homophony is the normal currency of vocabularies. Thus, this "help" with homophony is at best of minor usefulness.

However, attaining mastery of a confusable L2 distinction would deliver a huge gain for learners perceptually. The statistics computed over the English lexicon by Cutler (2005) for the æ/ε and r/l confusions show that each can lead to enormous increases in potential competition during listening to English. The number of spuriously embedded words (such as *leg* in *regular*) is increased, per contrast, by between 10,000 and 70,000 extra embeddings for every million words of speech, and the number of temporarily activated competitors (such as *televise* given the beginning of *terrify*) also hugely increases. Thus, it would really be helpful if establishing a lexical representation of a distinction were to assist listeners in correctly perceiving the distinction in speech input. However, there is no evidence that this happens. The evidence clearly indicates that the lexical coding of such contrasts is not helpful at all: avoiding collapse of the contrast in the lexicon does not lead to the contrast collapse being avoided in perception. In the eye-tracking studies described above, Dutch listeners proved to have distinct lexical representations of /æ/ and /ε/ but to treat both equivalently (as more likely to be /ε/) when hearing them. Japanese listeners likewise had distinctly stored representations of /r/ and /l/ and nevertheless treated either when heard as more likely to be /l/. Asian users of the many websites devoted to the English /r/–/l/ contrast presumably know that there is a distinction that they are trying to acquire, and therefore are likely to have incorporated it in their lexical representations. If knowing that the distinction exists were enough to perceive it correctly in the input, much anguish and effort would be instantly avoided!

Coding of the distinction in lexical entries is not only unable to help where it should help (i.e., in reducing competition), but it also exercises exactly the opposite effect. Incorporating a distinction at the lexical level without being able to perceive it in the lexicon works substantially against the learner's interest, in that it increases competition. Broersma and Cutler (2011; see also Broersma, 2012; Broersma & Cutler, 2008) examined L2 listeners' sensitivity to embedded-word competitors, using a cross-modal priming paradigm, which provides an appropriate index of word activation. In this task, listeners hear spoken input, but their nominal task is to perform lexical decision on letter strings that appear on a computer screen as they listen. Repetition of a lexical item, even in a different modality, always speeds decision time for the second presentation. Thus, if listeners hear *deaf*, then see the printed word DEAF on the screen, their response time to accept it as a real word will be faster than if they had just heard some unrelated (control) word instead of *deaf*. This is true for both L1 and L2 listeners. *Deaf* can also occur as a fully embedded form, for instance, in *definite*, and the phonemically identical *def-* extracted from *definite* also led to priming for DEAF, again for both L1 and L2 listeners. (Note that *def-* in *definite* will be shorter in duration than *deaf* in isolation, but this did not remove the priming effect when the extracted token was itself presented in isolation.)

Given the vowel confusions that Dutch listeners to English make, Dutch listeners' responses to DEAF were also faster after hearing the syllable *daff-*, extracted from *daffodil* and presented in isolation (native English listeners showed no priming at all in this case; *daff-* was as ineffective as an unrelated control prime in influencing responses to DEAF).

However, the most interesting case for our present purposes is what happens when mismatching input is available. This can be tested by presenting in full a spoken *definite*, containing within it the syllable identical to *deaf* (which, extracted from *definite* and presented in isolation, will prime DEAF, as just described). In this case, there is no facilitation of DEAF at all, either for L1 or for L2 listeners, because the rapidly arriving subsequent context *-inite* mismatches *deaf*, as well as providing increased matching support to *definite*. This is the way mismatch and competition are supposed to work: longer words will triumph over the shorter words embedded within them, even though the lexical representations of these shorter words will be briefly activated. The mechanisms function in the same way in L1 and L2. When the L2 listener has an accurate phonological representation in the lexicon that matches correctly to the perceived input (as is the case with *deaf/definite*), everything works just as it should.

However, what if the lexical representation does not match the perceived input (as is the case with *leg/regular* or *deaf/daffodil*)? We have already seen that the truncated longer word (in the experiment, for instance, *daff-*) will activate the pseudomatched word (*deaf*), so we know that the representation of this shorter word will be activated. However, will the subsequently arriving context rapidly trounce this activation and favor the longer word, as it should? No, it will not, because the input is being perceived as beginning with *def-*, and this does not match *daffodil*'s lexical representation, which is correctly stored as not containing the same vowel as in *deaf* and *definite*, even though its first vowel when heard is never actually interpreted as different from that of *deaf* or *definite*. Thus, the rapid effect of a mismatch that was seen for *deaf/definite* does not happen for L2 listeners with *deaf/daffodil*. Broersma and Cutler (2011) observed that there is significant facilitation of DEAF when the spoken input is *daffodil*. (L1 listeners, who never activated *deaf* given *daff-* anyway, show no such priming from *daffodil* either, of course. Table 1 summarizes the pattern of results across listener groups and input conditions.)

Thus, a very serious hindrance arises when a distinction that cannot be perceived in the input is incorporated in the lexicon. Presumably, an L2 listener who perceives the input as equivalent to *defodil* will eventually realize that this heard form refers to the flower, given the evidence from the second and third syllables. The situation will be exactly like recognizing an actually mispronounced word; the listener will have to ignore the perceptual evidence for that one sound and achieve recognition

Table 1. *Priming for DEAF according to Broersma & Cutler (2011)*

	Listeners	
Input	L1	L2
deaf	YES	YES
def-	YES	YES
definite	NO	NO
daff	NO	YES
daffodil	NO	YES

on the basis of all the correctly perceived sounds. However, this will not happen quickly enough to get rid of the competition from *deaf* (in the way that, when *definite* is heard, competition from *deaf* is so rapidly discarded). In other words, competition from spuriously embedded forms such as *leg* in *regular* or *deaf* in *daffodil* is not only present in L2 listening but also particularly persistent: it is in effect *stronger* competition than the expected competition such as *deaf* in *definite* or *leg* in *legacy*. By implication, the L2 listeners would have been better off if the pronunciation of *daffodil* had been incorrectly stored as *defodil*, because then the situation would have been comparable to the *definite* case. (This cannot be tested in the cross-modal priming task with real words, because all Dutch users of English are aware that the æ/ε distinction exists, and they know how to spell the English words that they can use. However, it might be possible to test such a prediction with a variant of the Escudero et al., 2008, nonword technique.)

In summary, the distinction that has been accurately incorporated in the lexicon, but crucially on the basis not of phonetic perception but of nonspeech information, has created a situation in which mismatch cannot work, and in doing so, it has exacerbated the competition problem in spoken-word recognition rather than alleviating it. L2 learners are doing themselves no perceptual favor at all by incorporating into their lexicon (whether by using information from orthography or from any other source) a distinction that they cannot reliably perceive in the input.

ORTHOGRAPHIC INFORMATION ABOUT PHONOLOGY

This discussion so far and the surveyed research has concerned the intensively researched topic of L1/L2 comparisons of phoneme repertoire. However, phonemic repertoire is not the only phonological information that L2 learners must master. Phonotactic constraints are language-specific also (and L1 phonotactics can interfere with L2 listening even at very high levels of L2 proficiency; Weber & Cutler, 2006). Suprasegmental information such as pitch can be lexically distinctive in one of a speaker's languages but quite irrelevant for distinguishing between words in another. The phonological shape of words differs across languages, as does lexical prosody, and prosodic structure above the word potentially further complicates phonological learning.

While alphabetic writing systems provide reasonably good information about phonemic repertoires (especially when grapheme-to-phoneme mappings are highly consistent), the rest of the phonological information that L2 learners need to acquire is hardly available from reading at all. Give or take a few indirect implications and an occasional language that makes certain phonological structures explicit, the natural availability of nonsegmental phonological information in orthography is unimpressive:

1. *Sentence prosody:* Intonation, focus, and contrastive emphasis can be in part expressed syntactically, but they are principally expressed in suprasegmental dimensions; because they are very largely determined by contextual factors, they vary widely and are essentially never coded orthographically. Even though aspects of sentence prosody vary across languages, L2 learners must derive knowledge about this from listening alone. Some prosodic effects, such as final lengthening or phrase-boundary marking, arise from syntactic structure that is partly coded in written texts by punctuation marks; question marks may likewise be said to represent information that can also be expressed prosodically. However, punctuation gives no direct information about how or where to realize such prosodic effects (consider that *wh*-questions and yes–no questions in English differ intonationally but are written with the same question mark).
2. *Lexical prosody:* Stress languages do not, in general, incorporate stress in writing, although exceptions may be orthographically marked where general rules account for the majority of stress placements, as in Spanish (note that even then, the number of cases counting as exceptional, and hence receiving such explicit realization, varies across varieties of the language, for example, between European and South American Spanish). L2 learners of English and other Germanic languages, or of Russian, or Arabic, or Indonesian, however, receive not even such partial indications of stress placement. In pitch accent languages such as Japanese, likewise, there is no orthographic coding of accent patterns.
3. *Phonotactics:* Some phonotactic constraints can be inferred from the available orthographic symbols in explicitly syllabic orthographies such as Korean, or the kana orthographies of Japanese, but in alphabetic orthographies, especially those with higher degrees of opacity, again there is no orthographic support for learning.
4. *Casual speech processes:* In informal speech, speakers reduce, delete, or assimilate sounds that may be explicitly articulated in more formal registers, and again, none of this is ever encoded in formal writing (though novelists occasionally use "eye dialect" to represent some aspects of casual speech, such as syllable deletion, or assimilation as in *doncha* for *don't you*; and some processes also turn up in text messages, the driving force there, however, being space saving, not verisimilitude). Such processes are sometimes language specific, although cross-language comparisons have suggested that perceptual compensation owes more to auditory than to linguistic processes, in that listeners without relevant language experience show perceptual patterns similar to those of native listeners (Gow & Im, 2004; Mitterer, Csépe, Honbolygo, & Blomert, 2006).
5. *Lexical tone:* Tone languages with alphabetic orthographies often encode syllabic tones (both contour tones, as in Vietnamese, and register tones, as in many

> African languages). This explicit marking underlines the lexical status of tone as being functionally more similar to segmental structure than to prosodic suprasegmentals. There is some debate about the value of tone marking to native users of African languages, largely on the grounds that the tone diacritics make text too visually complex (see Bird, 1999a; Roberts, 2009, 2011), and in one experiment prompted by this debate (Bird, 1999b), the presence of tone marks was found to decrease L1 reading fluency. For L2 learners, however, written tone information may well be useful if it can be correctly stored and correctly matched to perceptual input. Showalter and Hayes-Harb (2013) found that tone marks on a *pinyin* representation of Mandarin assisted English speakers (without prior knowledge of Mandarin) to store and recall tones assigned to nonwords.

In summary, very little information about phonology is available in written text. What alphabetical orthographies can tell readers about the citation pronunciation of words (segmental sequence, plus in certain cases tonal realization) is the best phonological information that is to be found in this way.

Moreover, on the basis of their L1 experience, L2 learners will, in general, expect orthography to offer very little phonological information beyond the segmental level. It is therefore reasonable to ask what they might expect to get out of such information if it were to be presented. The results of the phonemic confusion studies discussed in detail above strongly suggest that any available information will only be useful if lexical encoding of a contrast can be accompanied by perceptual discrimination of the same contrast. Otherwise, using the information is likely only to result in worse perceptual performance, as proved to be the case when single category phoneme contrasts were represented in the lexicon.

There is at least one nonsegmental type of phonological information for which evidence exists that some L2 listeners are likely to have great difficulty mastering the perceptual contrast, and that is lexical stress. Listeners whose L1 does not have lexical stress are known to experience great difficulty encoding and recalling stress patterns (Dupoux, Pallier, Sebastián-Gallés, & Mehler, 1997, comparing French listeners' processing of Spanish stress; see Peperkamp, Vendelin, & Dupoux, 2010, for similar evidence from Finnish and Hungarian listeners). L2 learners from such backgrounds also perform poorly on tests of stress mastery in their L2 (Tremblay, 2008, for French learners of English; Schmidt-Kassow, Rothermich, Schwartze, & Kotz, 2011, for French learners of German; Archibald, 1997, for Chinese and Japanese learners of English), even when they can accurately perceive an equivalent nonlinguistic contrast (Schmidt-Kassow et al., 2011). Conversely, listeners whose L1 is another stress language are able to perceive and store the stress patterns of their L2, even if the L1 and L2 stress placement rules differ (Cooper, Cutler, & Wales, 2002, for Dutch learners of English; Guion, Harada, & Clark, 2004, for Spanish learners of English; Suárez & Goh, 2013, for English learners of Spanish). This points to the likelihood that explicitly providing stress information would be useful only to learners whose L1 was also a stress language (such as English learners of Russian; Hayes-Harb & Hacking, 2013).

Note that in English stressed and unstressed syllables usually differ in containing full versus reduced vowels, making additional attention to suprasegmental cues redundant, and in consequence L1 listeners do not actually attend to suprasegmental cues to stress in recognizing words (Cooper et al., 2002; Fear, Cutler, & Butterfield, 1995). Nonetheless, the suprasegmental cues are present in English speech, and they can be put to good use by L2 listeners whose L1 does require attention to such stress cues (Cooper et al., 2002; Cutler, 2009). Furthermore, that English listeners to Spanish (which has no vowel reduction, and hence no stressed–unstressed vowel difference for English listeners to exploit) can perform well with Spanish stress (Suárez & Goh, 2013) suggests that the English-speaking listeners' native patterns have laid the groundwork for perception of stress distinctions in a non-English manner (i.e., using the Spanish-like suprasegmental cues) as well. However, it may still be the case that in an L2 that does have vowel reduction but does not use it as systematically as English in unstressed syllables (e.g., Dutch or Russian), English listeners may experience difficulty in storing the precise phonological form (Hayes-Harb & Hacking, 2013), just as learners from such languages often fail to store and produce English stress distinctions in a nativelike manner (causing consequent comprehension difficulties for native English listeners: Braun, Lemhöfer, & Mani, 2011; Cutler, in press).

CONCLUSION

In practice, learning an L2 involves learning from multiple types of input. Further, L2 learners literate in their L1 usually want to be able to read and write their new language as well as speak and understand it. Then they will also want to represent in memory everything they learn about the words they have encountered.

Learning an L2 is subject to interference from the L1 at all levels, so that any stored aspect of lexical information may be difficult to encode if the L1 has induced prior expectations that do not match the L2 requirements. Certainly this includes orthography. Learning of novel words is negatively affected when the forms are spelled in a way unusual for their pronunciation in the participants' L1, even when the L1 in question (English) has an opaque orthography itself (Hayes-Harb, Nicol, & Barker, 2010; Rastle, McCormick, Bayliss, & Davis, 2011). The nature of the orthography used by the L1 then also plays an important role with novel L2 words; difficulty can arise even if both L1 and L2 use the same Roman alphabet, when the degree of grapheme–phoneme transparency differs (for instance, L1-induced expectations that words may be opaquely coded in writing are helpful when a new opaque orthography is encountered, while expectations of transparency are a hindrance in the same case; Erdener & Burnham, 2005). However, the L1 orthography also shapes coding of L2 words when the L2 grapheme–phoneme correspondences are inadequately mastered, leading to well-known effects of mispronunciation (for relevant references, see Bassetti & Atkinson, 2015 [this issue]). In perceptual tasks, such influence can also arise (Escudero & Wanrooij, 2010), presumably due to the stored representations being phonemically categorized in a way that would be appropriate for the matching orthographic representation in the L1. The greater the learners' proficiency in their L2, the greater such effects of congruence of spelling between L1 and L2 can be (Veivo & Järvikivi, 2013).

None of this suggests that L2 learners can expect much true and lasting help from recourse to orthography in their quest to master an L2 phonology. Obviously mastering the orthography of an L2 is necessary for full use of the language, and it may be helpful in many ways beyond reading and writing, such as understanding relationships between words, appreciating morphological structure, and more. However, in the realm of segmental phonology helpful effects are strongest where the L1 has already paved the way, that is to say, with L2 contrasts that are congruent with contrasts in the L1 (Simon & Escudero, 2013). Precisely the segmental contrasts where the L2 learner needs most help turn out, on the evidence described in the central sections of this paper, to be the situation where the effects of a lexically coded distinction are least helpful.

In both L1 and L2, lexical representations are abstracted from input evidence, and every aspect of such a representation that has as its source other than spoken input must be solely based on abstract knowledge. It is worth pointing out that the limitations of experiential episodes in the creation of L2 phonological representations not only highlight the disconnection between the way a phonemic contrast is represented in stored word forms and the perceptual processing of the cues to the same contrast but also emphasize the crucial role of abstraction in the lexicon in general. In conclusion, using orthography or other abstract speech-external information to incorporate a distinction into lexical representations certainly works. However, the side effects of gaining the distinction solely in such a way, without perceptually attainable support, are added competition, and hence processing delay in word recognition, which can be quite inimical to learners' progress in their L2.

REFERENCES

Archibald, J. (1997). The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics, 35*, 167–181.

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics, 36*, 67–91.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro & O.-S. Bohn (Eds.), *Second language speech learning* (pp. 13–34). Amsterdam: John Benjamins.

Bird, S. (1999a). Strategies for representing tone in African writing: A critical review. *Written Language and Literacy, 2*, 1–44.

Bird, S. (1999b). When marking tone reduces fluency: An orthography experiment in Cameroon. *Language and Speech, 42*, 83–115.

Braun, B., Lemhöfer, K., & Mani, N. (2011). Perceiving unstressed vowels in foreign-accented English. *Journal of the Acoustical Society of America, 129*, 376–387.

Broersma, M. (2012). Increased lexical activation and reduced competition in second-language listening. *Language and Cognitive Processes, 27*, 1205–1224.

Broersma, M., & Cutler, A. (2008). Phantom word recognition in L2. *System, 36*, 22–34.

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology, 64*, 74–95.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and nonnative listeners. *Language and Speech, 45*, 207–228.

Cutler, A. (2005). *The lexical statistics of word recognition problems caused by L2 phonetic confusion*. Paper presented at Interspeech 2005, Lisbon, Portugal. Retrieved from http://www.isca-speech.org/archive/interspeech_2005/i05_0413.html

Cutler, A. (2009). Greater sensitivity to prosodic goodness in non-native than in native listeners. *Journal of the Acoustical Society of America, 125*, 3522–3525.

Cutler, A. (in press). Lexical stress in English pronunciation. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation*. New York: Wiley–Blackwell.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*, 269–284.

Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36*, 406–421.

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning, 55*, 191–228.

Escudero, P. (2015). Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words. *Applied Psycholinguistics, 36*, 7–22.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel L2 words and asymmetric lexical access. *Journal of Phonetics, 36*, 345–360.

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech, 53*, 343–365.

Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America, 97*, 1893–1904.

Gow, D. W., Jr., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language, 51*, 279–296.

Grainger, J., Van Kang, M. N., & Seguí, J. (2001). Cross-modal repetition priming of heterographic homophones. *Memory & Cognition, 29*, 53–61.

Guion, S. G., Harada, T., & Clark, J. J. (2004). Early and late Spanish–English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition, 7*, 207–226.

Hanson, V. L. (1982). Use of orthographic structure by deaf adults: Recognition of fingerspelled words. *Applied Psycholinguistics, 3*, 343–356.

Hanson, V. L., Shankweiler, D., & Fischer, F. W. (1983). Determinants of spelling ability in deaf and hearing adults: Access to linguistic structure. *Cognition, 14*, 323–344.

Hayes-Harb, R., & Hacking, J. (2013). T*he influence of written stress marks on native English speakers' acquisition of Russian lexical stress contrasts.* Paper presented to 23rd EuroSLA, Amsterdam, August.

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech, 53*, 367–381.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Ketterman, A., et al. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47–B57.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language. *Philosophical Transactions of the Royal Society B, 363*, 979–1000.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science, 30*, 451–479.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1209–1228.

Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics, 38*, 422–430.

Rastle, K., McCormick, S. F., Bayless, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1588–1594.

Roberts, D. (2009). Visual crowding and the tone orthography of African languages. *Written Language and Literacy, 12*, 140–155.

Roberts, D. (2011). A tone orthography typology. *Written Language and Literacy, 14*, 82–108.

Saffran, J. R., Werker, J., & Werner, L. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In R. Siegler & D. Kuhn (Eds.), *Handbook of child development* (pp. 58–108). Hoboken, NJ: Wiley.

Schmidt-Kassow, M., Rothermich, K., Schwartze, M., & Kotz, S. A. (2011). Did you get the beat? Late proficient French–German learners extract strong–weak patterns in tonal but not in linguistic sequences. *NeuroImage, 54*, 568–576.

Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research, 29*, 185–200.

Simon, E., & Escudero, P. (2013). *Incongruent grapheme-to-phoneme correspondences hinder L2 word learning.* Paper presented to EUROSLA 23, Amsterdam, August.

Suárez, L., & Goh, W. D. (2013). Recognition memory of foreign language lexical stress. *Memory & Cognition, 41*, 872–885.

Tremblay, A. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics, 29*, 553–584.

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

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1–25.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America, 119*, 597–607.