

# Perceptual representations of phonotactically illegal syllables

Mara Breen · John Kingston · Lisa D. Sanders

Published online: 18 November 2012  
© Psychonomic Society, Inc. 2012

**Abstract** Listeners often categorize phonotactically illegal sequences (e.g., /dla/ in English) as phonemically similar legal ones (e.g., /gla/). In an earlier investigation of such an effect in Japanese, Dehaene-Lambertz, Dupoux, and Gout (2000) did not observe a mismatch negativity in response to deviant, illegal sequences, and therefore argued that phonotactics constrain early perceptual processing. In the present study, using a priming paradigm, we compared the event-related potentials elicited by Legal targets (e.g., /gla/) preceded by (1) phonemically distinct Control primes (e.g., /kla/), (2) different tokens of Identity primes (e.g., /gla/), and (3) phonotactically Illegal Test primes (e.g., /dla/). Targets elicited a larger positivity 200–350 ms after onset when preceded by Illegal Test primes or phonemically distinct Control primes, as compared to Identity primes. Later portions of the waveforms (350–600 ms) did not differ for targets preceded by Identity and Illegal Test primes, and the similarity ratings also did not differ in these conditions. These data support a model of speech perception in which veridical representations of phoneme sequences are not only generated during processing, but also are maintained in a manner that affects perceptual processing of subsequent speech sounds.

**Keywords** Speech perception · Phonotactics · Event-related potentials · Priming

---

M. Breen (✉)  
Department of Psychology and Education,  
Mount Holyoke College,  
South Hadley, MA 01075, USA  
e-mail: mbreen@mtholyoke.edu

J. Kingston  
Department of Linguistics,  
University of Massachusetts,  
Amherst, MA, USA

L. D. Sanders  
Department of Psychology;  
Neuroscience and Behavior Program,  
University of Massachusetts,  
Amherst, MA, USA

Categorization of nonnative speech sounds and patterns is strongly influenced by native-language experience. In many situations, listeners assimilate nonnative sounds to native ones. As a result, patterns that do not exist in their language are reported to be ones that do (Brown & Hildum, 1956; Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Dupoux, Pallier, Kakehi, & Mehler, 2001; Hallé & Best, 2007; Hallé, Best, & Bachrach, 2003; Hallé, Segui, Frauenfelder, & Meunier, 1998; Massaro & Cohen, 1983; Moreton, 2002). The goal of the present study was to determine whether, in addition to assimilated native-language categories, veridical representations of nonnative sound patterns are also stored.

## Assimilation

The assimilation of nonnative to native patterns has been reported for a variety of languages. One example comes from Japanese, which bans adjacent consonants with different places of articulation. Illegal consonant clusters in loan words are repaired by inserting a vowel between the consonants. For example, *MacDonald's* becomes *makudonarudo*, where the underlined vowels are inserted to make the word phonotactically legal in Japanese. In addition to influencing production, this phonotactic constraint influences what Japanese listeners report hearing. In response to an /ebuzo/ to /ebzo/ continuum constructed by incrementally removing the /u/ between the /b/ and /z/, the percentage of trials on which Japanese listeners reported that the vowel was present only dropped from 100 % to 70 % (Dupoux et al., 1999). In contrast, French listeners' vowel-present responses across the continuum were accurate, dropping from 100 % to 5 %; French differs from Japanese in permitting consonants like /b/ and /z/ to occur next to one another. These results suggest that the constraints on Japanese syllable structure can cause native speakers to perceive an absent vowel between consonants that cannot occur next to one another. Further evidence has indicated that the source of the inserted

vowel is Japanese phonotactics rather than its lexicon (Dupoux et al., 2001).

Japanese speakers' assimilation of nonnative clusters in production and perception clearly shows that the phonotactics of their native language influence their representations of strings of segments. However, this evidence does not indicate whether *all* of the representations of these segments are shaped by native-language phonotactics. Does phonotactic knowledge completely alter listeners' representations of the segment strings, or only how they categorize their responses? That is, do the only representations that Japanese listeners have of illegal consonant clusters include intervening vowels? In *interactive* models such as TRACE (Elman & McClelland, 1988; McClelland & Elman, 1986; McClelland, Mirman, & Holt, 2006a, 2006b), listeners' linguistic knowledge is applied as soon as it is activated by the incoming signal, and this activation of linguistic knowledge alters the strength of activation of candidate sounds as well as words. In *autonomous* models such as Merge (McQueen, Norris, & Cutler, 2006a; Norris, McQueen, & Cutler, 2000), linguistic knowledge does not feed back onto auditory representations of speech sounds unless and until the experimental task demands it.

### Interactive and autonomous models

Most of the debate between proponents of interactive and autonomous models has focused on the relationship between lexically-driven categorization of ambiguous sounds (Ganong, 1980; Pitt & Samuel, 1993) and compensation for coarticulation. Compensation for coarticulation resembles the effects of phonotactics on perception, in that the listener's percept is adjusted depending on the context in which a sound occurs. Some have argued that the observed compensation effects are directly caused by lexical categorization (Elman & McClelland, 1988; Magnuson, McMurray, Tanenhaus, & Aslin, 2003a, 2003b; Samuel & Pitt, 2003). Others have attributed these effects to statistical generalizations across the lexicon (McQueen, 2003; McQueen, Jesse, & Norris, 2009; Pitt & McQueen, 1998) or to the task that listeners were asked to perform (Norris et al., 2000). Evidence that linguistic knowledge always influences perception of speech sounds supports interactive models; evidence that the influence of linguistic knowledge on speech perception can be disrupted supports autonomous models.

Listeners' lexical knowledge has also been shown to influence their perceptual adjustment to idiosyncrasies in pronunciation (Eisner & McQueen, 2005; Kraljic & Samuel, 2005, 2006, 2007; Maye, Aslin, & Tanenhaus, 2008; McQueen, Cutler, & Norris, 2006; McQueen et al., 2009; McQueen, Norris, & Cutler, 2006b; Norris, McQueen, & Cutler, 2003).

For example, after listening to a fricative that is ambiguous between /s/ and /f/ in multiple contexts in which only /s/ results in a word (e.g., following the context *pea*, /s/ produces the word *peace*, but /f/ produces the nonword *peaf*), listeners are more likely to categorize the ambiguous sound as /s/, even in a context that is not lexically biased. The finding that listeners can learn to associate specific acoustic characteristics with different phonological categories over time indicates that some stored representations of sounds must be acoustically veridical; having stored representations that are not modulated by linguistic knowledge is consistent with autonomous models.

In the present study, we tested the predictions of interactive and autonomous models using a phonotactic constraint found in both English and French. The clusters /dl/, /tl/, and /sr/ can never occur within a syllable,<sup>1</sup> while the onset clusters /dr/, /tr/, and /sl/ are legal and frequent (e.g., *drip*, *trip*, and *slip*). Massaro and Cohen (1983) presented an /r/-/l/ continuum in the contexts /t\_i/ and /s\_i/ to native English speakers. The listeners identified the ambiguous tokens more often as /r/ after /t/ and more often as /l/ after /s/. Similarly, Moreton (2002) presented native English speakers with a stop from a /d/-/g/ continuum, followed by a sonorant from an /l/-/w/ continuum; /dw/, /gw/, and /gl/ occur as onset clusters in English words, but /dl/ does not. When listeners reported hearing /d/, they were significantly less likely to report that the following sound was /l/. Furthermore, Hallé and Best (2007) explored the effects of native-language phonotactics in French, English, and Hebrew listeners. As described above, the clusters /dl/ and /tl/ can never occur within a syllable in French or English. In contrast, these clusters do occur as onsets in Hebrew. The legality of these clusters in Hebrew is important, because the stimuli employed in the present study were produced by native Hebrew speakers. All three groups were successful at discriminating /d/ from /g/ and /t/ from /k/ before /r/ in a three-interval oddity task (accuracy > 95 %). However, only Hebrew listeners reached this level of accuracy discriminating these stops before /l/. Furthermore, the Hebrew listeners responded an average of 600 ms before a triplet ended in this condition. In contrast, for French and English speakers listening to the stops followed by /l/, accuracy dropped below 80 %, and responses slowed to around 500 ms after the end of a triplet. Finally, the more likely that an individual French or English listener was to categorize /d/ as /g/ or /t/ as /k/ before /l/ in an independent categorization task, the poorer this individual was at discriminating between /d/ and /g/ or /t/ and /k/ in this context.

The findings from these studies of phonotactic constraints appear to support an interactive model of speech perception, in that they show that listeners' responses to speech sounds are influenced by their knowledge of which clusters are

<sup>1</sup> /d/ or /t/ cannot occur together with /l/ in an onset, but they can readily precede /l/ in a different syllable, as in *bed.lam*, *todd.ler*, *at.las*, and *tatt.ler*, where the periods mark the syllable boundaries, or in *saddle*, and *bottle*, where the /l/ is syllabic (see also Moreton, 2002).

phonotactically legal. Specifically, the phonotactic constraint against /dl/ and /tl/ onset clusters in French and English forced a systematic misperception of /dl/ as /gl/ and of /tl/ as /kl/. Importantly, this pattern was observed using both identification and discrimination tasks. Therefore, it appears that French and English listeners were unable to access veridical representations of /d/ and /t/ before /l/. Hebrew listeners, who do experience these clusters as onsets in their native language, were able either to assign these stops to their correct categories or, as suggested by the speed of their responses, to access veridical auditory representations.

### Event-related potentials

Using behavioral data to adjudicate between the autonomous and interactive views risks circularity; if effects of language experience are found, the task must reflect categorical representations, and tasks that reflect categorical representations are defined by showing the effects of language experience. The temporal resolution of event-related brain potentials (ERPs) makes it possible to distinguish between levels of representation that precede the listener's overt behavioral responses. Although there is clearly no single time boundary that neatly differentiates between perceptual representations and higher-level representations, earlier portions of the waveform (i.e., less than 350 ms after stimulus onset) are typically affected by stimulus characteristics and are more promising for indexing the veridical representations posited by autonomous models (Luck, 2005; Picton et al., 2000). Therefore, evidence that early portions of ERP waveforms are immune to the influence of language experience would support the autonomous view. In contrast, later portions of the waveform are more likely to be affected by task demands and are more likely to index the categorical representations posited by both the interactive and autonomous models.

Previous studies have employed ERPs to index the effects of language experience on speech processing using an oddball paradigm (Cheour et al., 1998; Dehaene-Lambertz, Dupoux, & Gout, 2000; Näätänen et al., 1997; Winkler et al., 1999). This procedure requires presenting many repetitions of a standard sound, interrupted infrequently and randomly by a deviant sound or sounds. The deviant sounds typically elicit a mismatch negativity (MMN) that peaks between 150 and 350 ms after onset. The MMN is commonly assumed to reflect the automatic updating of working memory when listeners are confronted with new input (Näätänen, Paavilainen, Rinne, & Alho, 2007). Although the short latency of the MMN (i.e., less than 350 ms after the point at which a deviant differs from the standard) is used to argue that it indexes perceptual

processing, perceptual representations and early ERPs can be modulated by top-down influences, including selective attention (Hansen, Dickstein, Berka, & Hillyard, 1983; Hillyard, 1981; Hink & Hillyard, 1976; Näätänen, 1990; Schröger & Eimer, 1997). As such, the effect of language experience on early ERPs could reflect either the categorical nature of all language representations posited by interactive models, or the top-down modulation of perceptual processing by attention.

Näätänen et al. (1997) recorded ERPs while presenting Estonian and Finnish speakers with a series of speech sounds in which the vowel /e/ was the standard and the vowels /o/, /ö/, and /õ/ were deviants. The deviants /o/ and /ö/ contrast with the standard /e/ in both languages, but /õ/ only contrasts with /e/ in Estonian. An MMN was observed in response to all three deviants in Estonian listeners, but only for /o/ and /ö/ in Finnish listeners. This result was interpreted as evidence that the MMN is sensitive to whether two sounds contrast in the listeners' native language. A similar difference was reported for a Finnish vowel contrast not found in Hungarian; Hungarian speakers who were fluent in Finnish showed an MMN, but other Hungarian speakers who had no prior exposure to Finnish did not show an MMN (Winkler et al., 1999).

To assess the immediacy of the application of phonotactic knowledge during speech processing, Dehaene-Lambertz et al. (2000) used a modified oddball design in investigating the representations of adjacent consonants by Japanese and French listeners. As stated earlier, Japanese bans adjacent consonants with different places of articulation, while French has no such prohibition; sequences such as /gm/ are illegal in Japanese but legal in French. Dehaene-Lambertz et al. presented listeners with sequences of five stimuli, consisting of four tokens of a VC.CV or V.Cu.CV string and a fifth token, either of the same string or of the other string. For example, if the first four stimuli were /igmo/, the deviant at the end of the sequence was /igumo/, and if the first four were /igumo/, the deviant was /igmo/. In this trial structure, the fifth sound was a standard if it had the same sequence of segments as the previous four stimuli and was a deviant if a vowel had been inserted or removed between the two consonants. Behaviorally, French listeners distinguished the deviants from the standards, but Japanese listeners did not. In addition, the deviant stimulus elicited an MMN for the French listeners but not for the Japanese listeners. Dehaene-Lambertz et al. interpreted this result as evidence that a filter repairs the nonnative consonant clusters to conform to native-language phonotactics at the earliest stages of processing at which the brain's responses were measured. These results, as well as those reported by Näätänen et al. (1997) and Winkler et al. (1999), can be interpreted as supporting the interactive hypothesis.

## Present study

The present study was also designed to investigate the levels of representation that are repaired for phonotactic violations. However, it differs methodologically in two important ways from Dehaene-Lambertz et al.'s (2000) study. First, their conclusion that listeners immediately repair illegal phonotactic patterns to legal ones rested on a between-subjects comparison; specifically, Japanese listeners demonstrated no early ERP difference, but French listeners did. However, this result could be due to Japanese listeners explicitly reporting a deviant sound on far fewer trials than the French listeners: Japanese listeners reported that the fifth stimulus was different on only 8.6 % of deviant trials, as compared to French listeners' 95.1 %. This difference in behavioral responses could induce a bias toward "same" responses and potentially reduce Japanese listeners' attention and sensitivity to stimulus differences. Therefore, in the present study, we tested the levels of representation that are affected by assimilation in the same individuals listening to both legal and illegal patterns. A second major difference between the present study and that of Dehaene-Lambertz et al. is that we did not adopt an oddball design of the type that elicits an MMN. It has been widely argued that the amplitude and latency of the MMN are closely linked to the ability of listeners to discriminate between the standard and deviant sounds (Horváth et al., 2008; Näätänen, 2001, 2008; Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007). As such, the MMN may be influenced by the same type of representation that listeners use to make overt discrimination decisions, and therefore may not provide additional information beyond behavioral measures. Of course, many other differences exist between the present study and Dehaene-Lambertz et al.'s, including the structure of the languages involved and the position of illegal sequences in the stimuli; however, these differences were not expected to affect whether the measurements taken could index veridical, perceptual representations of illegal sound combinations.

Instead of an oddball paradigm, we opted to measure the effects of various types of primes on ERPs elicited by physically identical targets, in addition to behavioral similarity ratings of the prime–target pairs. This design allowed us to measure the influences of multiple levels of the representations of the primes, which had to be maintained across the 1,200-ms stimulus onset asynchrony (SOA), on processing of the targets. Previous auditory priming studies have shown that shared characteristics of prime–target pairs (e.g., rhyming) result in decreased amplitudes of the ERPs elicited by targets (Coch, Grossi, Coffey-Corina, Holcomb, & Neville, 2002; Coch, Grossi, Skendzel, & Neville, 2005; Cycowicz, Friedman, & Rothstein, 1996; Friedrich, Schild, & Röder, 2009; Weber-Fox, Spencer, Cuadrado, & Smith, 2003). Specifically, Coch et al. (2002) demonstrated that the posterior negativity between 200 and 500 ms after the onset

of targets was reduced in amplitude when the words were preceded by rhyming primes (e.g., *nail–mail*) as compared to nonrhyming primes (e.g., *paid–mail*). In addition, the targets preceded by rhyming primes elicited a larger negativity over anterior regions. Coch et al. (2005) observed similar effects for rhyming and nonrhyming pseudowords, which suggests that lexical status is not a critical factor for ERP priming effects in adults. Furthermore, in German listeners, the posterior negativity 300–400 ms after the onset of a target word (e.g., *Treppe*) was reduced in amplitude when the word was preceded by matching (e.g., *trep*) or similar (e.g., *krep*) onset fragments, as compared to control primes (e.g., *dra*); both matching and similar primes also resulted in a larger negativity over left anterior regions in response to targets (Friedrich et al., 2009). These results indicate that the effects of phonological similarity on auditory ERPs are not limited to rhyming. It is also important that the effects of phonological similarity on ERPs may be influenced by the task that listeners are engaged in and the time between presentation of the primes and targets. The reduced posterior negativity in response to rhyming prime–target pairs was observed even when participants were performing a semantic judgment on target words with a short SOA between the primes and targets (Praamstra & Stegeman, 1993), but not when performing the same task with a longer SOA (Perrin & García-Larrea, 2003). These data suggest that representations of the primes must be maintained in memory while targets are presented in order to observe effects of phonological similarity between the primes and targets on ERPs.

Thus, an ERP priming design offered an optimal way to address the first goal of the study, which was to determine whether any veridical representations of illegal syllables are maintained for long enough to influence the processing of subsequent speech sounds. To ensure that listeners maintained at least some representation of the prime syllables, participants were asked to make similarity ratings for the prime–target pairs presented on each trial. To be certain that any differences in ERPs elicited by the targets reflected stored representations of the primes rather than merely transient effects of acoustic similarity, the onset asynchrony for prime–target pairs was set at 1,200 ms. We assumed that the prime syllable was processed fully during this time, such that when the target was presented, the listener compared it to stored representations of the prime. ERPs to the target syllable should therefore vary depending on how effectively the representations of the prime syllable facilitated processing of the target. Moreover, the time course of ERP differences informed our understanding of whether perceived similarities between the prime and target existed for perceptual or higher-level representations.

To determine the influence of stored representations of the prime on processing of the target, we presented three different Trial Types, in which one of three possible primes

preceded the same target. The first Trial Type was a Control condition, in which the prime and target were two phonemically distinct and attested English consonant clusters (e.g., /kwa/–/gwa/). The second Trial Type was the Identity condition, in which the prime and target were different tokens of identical syllables (e.g., /gwa/–/gwa/). The third Trial Type was the Test condition, in which the prime was a syllable of interest that differed from the target by only one contrastive feature (e.g., /dwa/–/gwa/, /d/ and /g/ differ in place of articulation). Differences in the ERPs elicited by targets preceded by Control and Identity primes could be used to define the largest priming effects in both early and later portions of the ERP waveforms, since the Control condition included primes and targets that were maximally different, and the Identity condition included primes and targets that were maximally similar in the context of the experiment. The extent to which the perceptual and higher-level representations of the Test prime syllable were similar to the target could then be determined by comparing the ERPs elicited by targets in this condition to those in both the Control and Identity conditions. Specifically, if the Test prime syllable (e.g., /dwa/) was perceived as highly similar to the target (e.g., /gwa/), the ERPs elicited by the target in the Test condition should closely resemble those in the Identity condition. If, on the other hand, the Test prime syllable was perceived as distinct from the target, ERPs elicited by the target in the Test condition should instead closely resemble those in the Control condition.

The examples provided above correspond to the Trial Types in the Legal Prime Status condition. The Test syllables in this condition included consonant clusters of a coronal stop /d/ or /t/ followed by /w/, which occur as the onsets in words like *dwarf* and *twice*. We predicted that listeners would perceive the Legal Test prime syllables (e.g., /dwa/) as being different from the target syllables in this condition (e.g., /gwa/), and that those different percepts would be reflected in both similarity ratings and ERPs. Specifically, the ERPs elicited by targets (e.g., /gwa/) were expected to be more similar in the Test and Control conditions (e.g., primes /dwa/ and /kwa/), and both were expected to differ in the same way from those in the Identity condition (e.g., prime /gwa/).

Illegal Prime Status included the critical conditions of the experiment, the results of which could adjudicate between the interactive and autonomous hypotheses. In these and all other conditions, the onset of the target syllable was an attested English consonant cluster (e.g., /gla/). The Control prime included another attested English consonant cluster that was phonemically distinct from the target (e.g., /kla/); the Identity prime was again a different token of the same syllable as the target (e.g., /gla/). The Illegal Test prime included a consonant cluster that is illegal in English and differed from the target by only one feature (e.g., /dla/, /d/ and /g/ differ in place of articulation). If listeners mistakenly

categorized /d/ as /g/ before /l/, as the behavioral evidence previously cited predicts, then participants should rate the primes and targets as being similar in both the Identity condition (e.g., /gla/–/gla/) and the Test condition (e.g., /dla/–/gla/). To the extent that all of the representations of the Illegal Test primes were influenced by native-language experience, ERPs elicited by targets (e.g., /gla/) would be more similar in all time windows for the Test and Identity conditions (e.g., primes /dla/ and /gla/), and both should differ in the same way from the Control condition (e.g., prime /kla/). This pattern of ERP data would indicate that listeners' knowledge of the phonotactic prohibition against /dl/ and /tl/ as onset clusters in English results in categorical repair of the perceptual representations of these illegal clusters to /gl/ and /kl/ and would support the interactive hypothesis. In contrast, if the perceptual representations of the Illegal Test primes included the information needed to differentiate /d/ from /g/ in this context, the ERPs elicited by targets (e.g., /gla/) would be more similar, in at least the early portion of the waveforms, for the Test and Control conditions (e.g., primes /dla/ and /kla/), and both should differ in the same way from the Identity condition (e.g., prime /gla/). This pattern of results would instead support the autonomous hypothesis, by showing that at least some of the representations of illegal patterns are veridical.

A second goal of the experiment was to determine whether the perceptual representations of syllables prohibited by phonotactic constraints are qualitatively different from the representations of syllables that are legal but accidentally nonoccurring, because they never arose during the language's historical development. The distinction between such absent clusters and outright illegal ones is motivated by behavioral evidence that although /bw/ and /dl/ are both absent as syllable onsets in English, Moreton (2002) only found a perceptual bias against /dl/. Specifically, when a stop from a /b/–/g/ continuum was followed by a sonorant from an /l/–/w/ continuum, on trials in which listeners reported hearing /b/, they were at least as likely to report that the following sound was /w/ as they were on trials in which they reported /g/ as the first consonant. If absence in a native language drives the categorization decisions, English listeners should be just as reluctant to categorize the sonorant as /w/ after /b/ as they would be to categorize it as /l/ after /d/, since the frequency of both /bw/ and /dl/ as onset clusters in English is 0. The finding that listeners were willing to report hearing /bw/ suggests that this onset cluster is better described as accidentally absent rather than, as has been suggested for /dl/, as being barred by phonotactic constraints.

Further evidence that illegal onsets are actively prohibited is provided by their being consistently repaired when they enter a language (Berent, Lennertz, Smolensky, & Vaknin-Nusbaum, 2009; Berent, Steriade, Lennertz, & Vaknin, 2007; Davidson, 2006, 2010). For example, the

initial /t/ in the name of an indigenous people of the Northwest Coast, *Tlingit*, is repaired to /k/ by English speakers. Active prohibitions sometimes arise during a language's history and rule out formerly permitted sequences. For example, many English words that are spelled with a silent *k* or *g* in front of *n* are cognates with German words in which both sounds are still pronounced (e.g., English *knee*, *knob*, *knuckle*, *knead*, *knot*, *knave*, *gnome* cognates of German *Knie*, *Knopf*, *knöchel*, *kneten*, *knoten*, *Knabe*, *Gnom*). At some point after the ancestral languages of modern English and German separated, English introduced a prohibition against /kn/ and /gn/ onset clusters that German did not. In contrast, onset clusters that are accidentally nonoccurring are not excluded by the rules of a language. For example, the pseudoword *blick* does not exist in English, but there are no phonotactic constraints that bar it: All of its constituent sequences do occur in existing words (e.g., *blimp*, *click*, *black*). In English, absent onset clusters are exemplified by a bilabial stop /b, p/ followed by /w/, which occur only in loan words. The *Oxford English Dictionary* lists only *Puerto Rican*, *pueblo*, *puissant*, *puissantly*, *puissance*, and *puissancey* with initial /pw/ clusters, and only *Buen Retiro* and *buisson* with initial /bw/ clusters (see Moreton, 2002, for lexical statistics on such clusters in English). English speakers also occasionally repair these absent onsets. For example, some English speakers pronounce the first word in *Puerto Rico* as /porto/. However, this seems to be an exception; the place name *Pueblo* with the same onset is reliably pronounced /pwɛblo/, with the /pw/ intact.

From how speakers treat illegal and absent onsets, it is reasonable to expect that there would be differences both in the extent to which listeners would rate syllables including these types of onsets as being similar to targets, and in the patterns of electrophysiological priming effects. To test these predictions, we included a third Prime Status condition. In the Absent Prime Status condition, the targets, Control primes, and Identity primes were again syllables that contained attested English consonant clusters (e.g., /gwa/); the Control primes were phonemically distinct from the targets (e.g., /twa/). Test primes differed from the targets by one contrastive feature and included onset clusters that are absent from English (e.g., /bwa/, /b/ and /g/ differ in place of articulation). Moreton's (2002) behavioral data suggest that listeners would perceive the Absent Test prime syllables (e.g., /bwa/) as being different from the target syllables in this condition (e.g., /gwa/), which would be reflected in both similarity ratings and ERPs. Similar to the Legal Prime Status conditions, we predicted that the ERPs elicited by targets (e.g., /gwa/) would be more similar in the Test and Control conditions (e.g., primes /bwa/ and /twa/) and that both should differ in the same way from the Identity condition (e.g., prime /gwa/).

## Method

### Participants

Eighteen participants (four female), ranging in age from 20 to 30 years ( $M = 24.5$  years) provided data included in the analysis. All of the participants reported being right-handed, having normal hearing and normal or corrected-to-normal vision, taking no psychoactive medications, and having no neurological disorders. The participants were all monolingual native speakers of American English with no experience with Hebrew, the language from which the stimuli were drawn. Five additional participants contributed data to the experiment; the data from two were discarded due to recording errors, from two because of low-frequency drift caused by skin potentials, and from one because of high-frequency noise in the EEG data. All participants provided written consent and were compensated \$10/h for their participation.

### Stimuli

We manipulated four factors with three, three, two, and two levels, resulting in 36 conditions. Examples of each condition are shown in Table 1. The first factor was Trial Type, defined by the shared characteristics of the prime and target. The pairs were (1) Control syllables that were phonemically distinct and formed two legal sequences (e.g., /kla-/gla/), (2) Identity syllables that were different tokens of phonologically identical sequences (e.g., /gla-/gla/), and (3) Test syllables that differed by a single feature in the first consonant, such that the prime met specific phonotactic conditions (e.g., /dla-/gla/). The second factor, Prime Status, referred to the linguistic status of the prime syllable in the Test Trial Type. The Prime Status conditions were (1) Legal consonant clusters as the onsets of the test primes (e.g., /dw/), (2) Illegal clusters as the onsets of the Test primes (e.g., /dl/), and (3) onset clusters that are Absent in English for the Test primes (e.g., /bw/). The third factor was the Voicing of the first consonant of the target syllables; this stop consonant was either voiced (e.g., /gw/) or voiceless (e.g., /kw/). The fourth factor (Speaker) was whether the prime and target were produced by the same speaker or by different speakers. These third and fourth factors did not interact with any other factors, and so are not discussed further.

The stimuli consisted of CCV syllables, in which the consonant clusters were those described above and listed in Table 1, followed by one of the three vowels /a/, /e/, or /i/. Two native speakers of Hebrew, who also speak English, recorded the stimuli in a sound-attenuating room. Each syllable was produced multiple times, with different intonations corresponding to various emotional states (e.g., surprised, happy, sad). As such, at least 30 tokens (3 vowels  $\times$  2

**Table 1** Consonant clusters from nine prime–target conditions, arranged by Trial Type (rows), Prime Status (columns), and Voicing (rows)

Trial Type		Legal		Illegal		Absent	
		Prime	Target	Prime	Target	Prime	Target
Control prime	Voiced	kw	gw	kl	gl	tw	gw
	Unvoiced	gw	kw	gl	kl	gw	tw
Identity prime	Voiced	gw	gw	gl	gl	gw	gw
	Unvoiced	kw	kw	kl	kl	tw	tw
Test prime	Voiced	dw	gw	dl	gl	bw	gw
	Unvoiced	tw	kw	tl	kl	pw	tw

Each cluster was followed by /a/, /e/, or /i/, but vowel was held constant within every prime–target pair presented on the same trial

speakers  $\times$  5 intonations) were recorded for each consonant cluster. The stimuli were spoken into a head-mounted microphone and digitized at a sampling rate of 44.1 kHz with 16-bit resolution. The individual syllables were then cut out of the continuous recording using a waveform editor, down-sampled to 16 kHz, and equalized in peak intensity. From the initial recordings, five prime–target pairs were selected to exemplify each of the 36 conditions. As is shown in Fig. 1, pairs were selected such that there was high similarity within pairs (i.e., same intonation and vowel) and low similarity across the pairs (i.e., different intonations and vowels); furthermore, half of the trials included primes and targets recorded by different speakers. The primes and targets were paired in this manner to clearly define the maximum degrees of phonological similarity and difference possible in the context of the experiment. The final group of selected and edited primes had a mean duration of 297 ms (range = 197–388 ms,  $SD = 44$  ms); the targets had a mean duration of 301 ms (range = 198–388 ms,  $SD = 45$  ms).

### Procedure

Participants were seated in a comfortable chair located 150 cm from the computer monitor. Each trial began with the presentation of a fixation cross, which remained on the computer monitor until the response cue appeared. The prime syllable began 500 ms after the onset of the fixation point, and the target stimulus began 1,200 ms after the onset of the prime. Since the primes had durations of 197–388 ms, the resulting silent intervals between primes and targets were 812–1,003 ms. Sounds were presented from two loudspeakers on either side of the monitor at an average intensity of 57 dBA as measured at the location of the participants. The response cue (“How similar? 1 = most similar; 4 = most dissimilar”) appeared 500 ms after the offset of the target and remained on the screen until the participant responded or until 5 s had elapsed. Short breaks were offered approximately every 15 min. Participants were encouraged to ask for longer breaks as needed. Each of the five prime–target pairs from the 36 conditions was presented ten times over

the course of the experiment, resulting in a total of 1,800 trials. These trials were presented in random order in 15 blocks of 120 trials each. The entire experimental session took 3–3.5 h to complete.

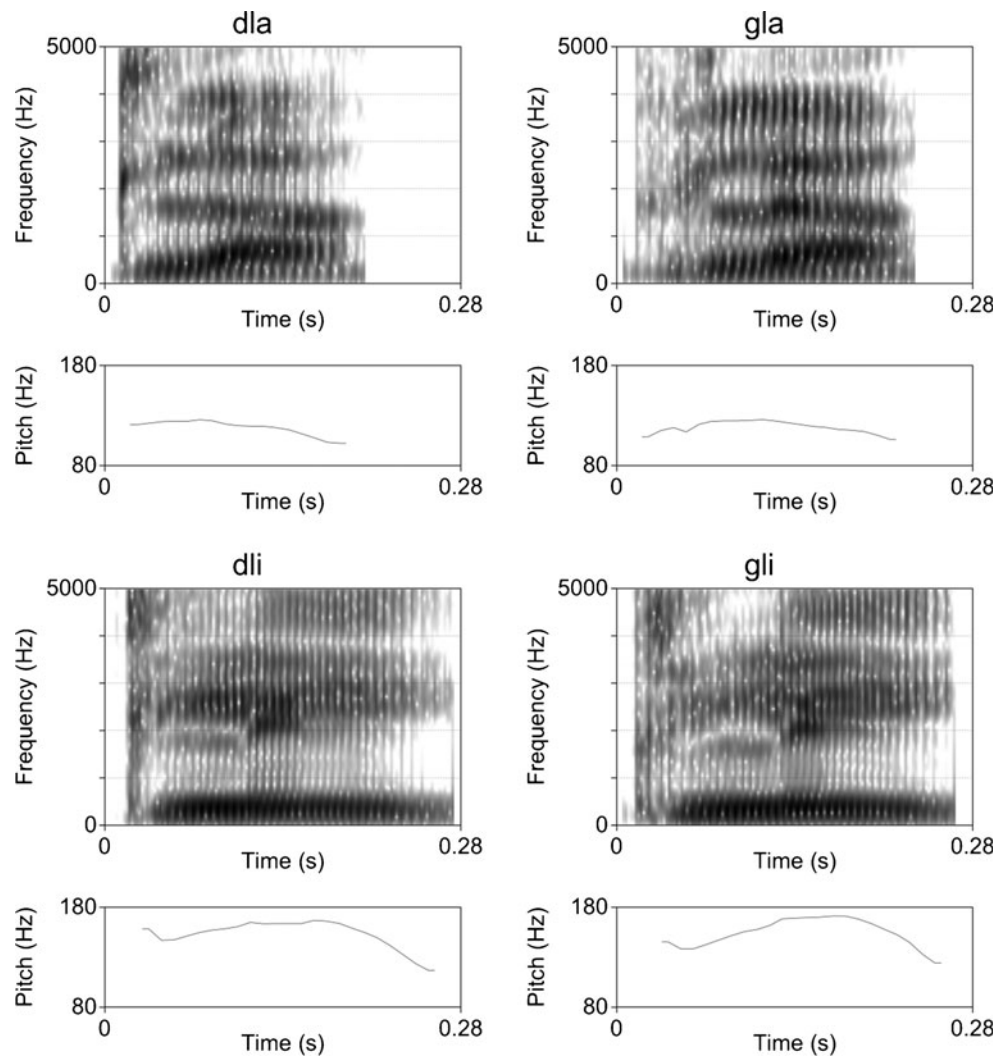
Continuous electroencephalogram (EEG) was sampled at 250 Hz and a bandwidth of 0.01–100 Hz throughout the duration of the experiment from 128-channel HydroCel Geodesic Sensor Nets (Electrical Geodesics Inc., Eugene, OR). Impedance was brought below 50 k $\Omega$  at every electrode at the beginning of the experiment and was maintained below 100 k $\Omega$  for the duration. The continuous EEG was divided into 700-ms epochs, from 100 ms before to 600 ms after the onset of a target. Trials that contained eye blinks, eye movements, or other identifiable artifacts were excluded from the analysis by setting participant-specific amplitude thresholds on the basis of visual inspection of the raw data. Data from the remaining trials were averaged by participants and conditions, re-referenced to the average of the two mastoid electrodes, and corrected to the 100-ms pretarget baseline. Participants included in the analysis contributed data from 83 to 198 ( $M = 176$ ) out of a possible 200 trials in each Trial Type by Prime Status condition.

### Analysis

The similarity ratings for prime–target pairs were the dependent variable in ordinal logistic mixed-effects regression models (Christensen, 2011) in the R statistical programming language (R Development Core Team, 2007). Specifically, the intercepts were calculated as the thresholds for multiple binary partitions of the responses, with the fixed effects of Trial Type and Prime Status and the random effects of participants and items.<sup>2</sup> Initial models were used to determine the effects of Trial Type within each Prime Status condition separately. Pair-wise comparisons between Trial Types were then repeated in models that included their possible interactions with Prime Status. Thresholds were

<sup>2</sup> The ordinal package in R does not yet permit inclusion of random effects on the slopes of the fixed effects.

**Fig. 1** Spectrograms and pitch contours from two representative Illegal Test prime–target pairs. A comparison of the primes and targets demonstrates the similarity of the syllables within a trial, including those with Test or Control primes. A comparison of the spectrograms and pitch contours in the top and bottom panels demonstrates variability between the pairs presented on separate trials. All syllables shown here were produced by the same speaker



calculated for the three possible partitions of the four response values, which ranged from 1 (i.e., *most similar*) to 4 (i.e., *most dissimilar*). The threshold estimates reflected the cumulative probabilities of ratings of 1 as compared to 2 or higher, ratings of 1 or 2 as compared to 3 or 4, and ratings of 3 or lower as compared to 4. The thresholds were not required to be equidistant, since participants did not distribute their ratings evenly across the four values.

ERP measurements taken at 72 electrodes across the scalp were first grouped into 12 regions by averaging recordings made at six nearby electrodes. The regions were focused over anterior, anterior/central, posterior/central, and posterior regions of the left and right hemispheres, as well as medial areas. The approximate locations of the electrodes included in each region are shown in Fig. 2. Mean amplitude measurements were taken in four time intervals: 50–80 ms (P1), 90–140 ms (N1), 200–350 ms, and 350–600 ms after target onset. The measurements taken in each time interval were entered into a 3 (Trial Type)  $\times$  3 (Prime Status)  $\times$  4 (Anterior–Posterior electrode position)  $\times$  3 (Laterality of

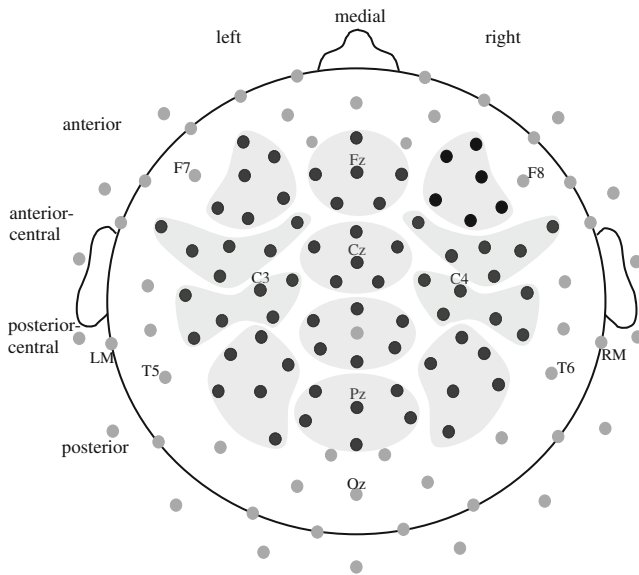
electrode position) repeated measures ANOVA to determine the distribution of any main effects. Greenhouse–Geisser adjustments were applied to all  $p$  values. All significant ( $p < .05$ ) interactions of Trial Type and electrode position factors were followed up by additional ANOVAs at subsets of electrodes. Trial Type was examined separately for each level of Prime Status, to test the specific hypotheses concerning whether data from the Test and Control Trial Types or the Test and Identity Trial Types would pattern together. Main effects of Trial Type were followed up with planned comparisons including only two levels of this factor.

## Results

### Similarity ratings

Histograms of the similarity ratings are shown in Fig. 3. Ratings of 4 (i.e., *most dissimilar*) predominated for all



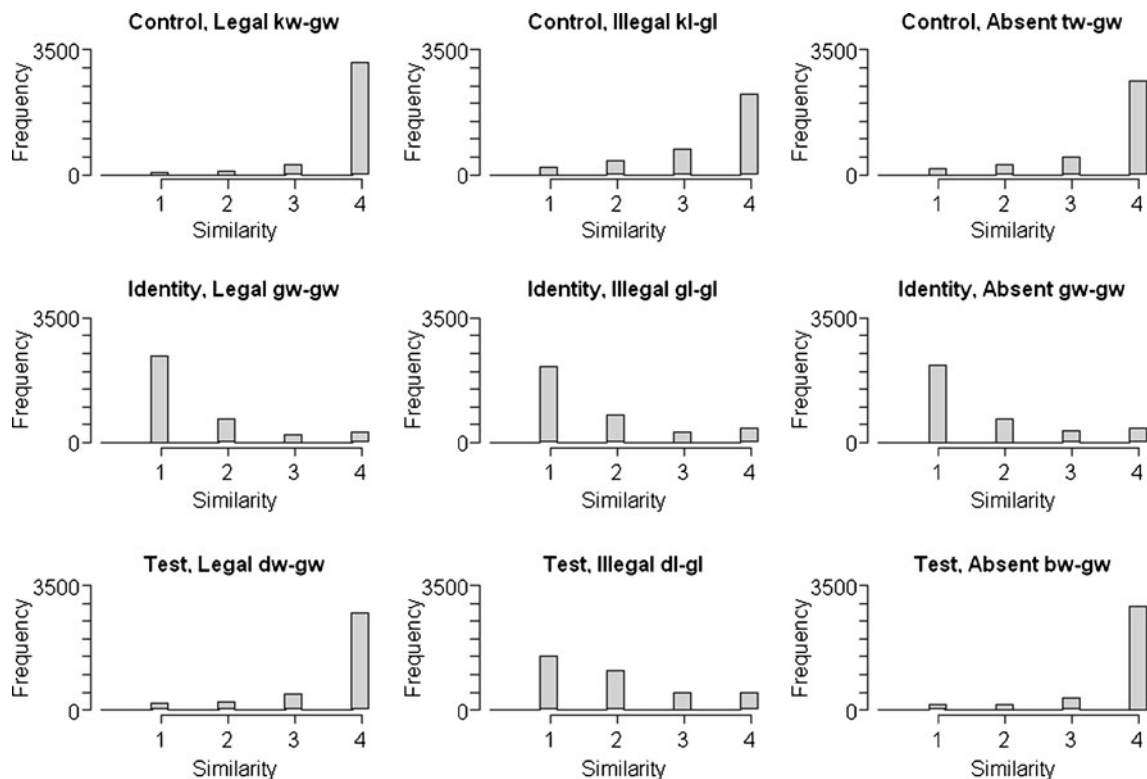


**Fig. 2** Approximate locations of the 72 electrodes that contributed data to the event-related potential analyses. Measurements from six nearby electrodes were averaged together. Electrode position was included as two factors in omnibus ANOVAs: Anterior–Posterior (anterior, anterior/central, posterior/central, and posterior) and Laterality (left, medial, right)

Control Trial Types, regardless of Prime Status (top row), while ratings of 1 (i.e., *most similar*) predominated for all Identity trials, regardless of Prime Status (middle row).

Ratings of 4 also predominated for Test Trial Types (bottom row) when the Prime Status was either Legal (left column) or Absent (right column). In contrast, for Test Trials with an Illegal prime (center column), the majority of ratings were 1s and 2s. These distributions indicated that on Test trials, participants consistently reported that Legal primes differed from targets (e.g., /dw/–/gw/) and that Absent primes differed from targets (e.g., /bw/–/gw/), but they did not report that Illegal primes differed from targets (e.g., /dl/–/gl/). This result was predicted from previous demonstrations of English listeners’ inability to distinguish /d/ from /g/ and /t/ from /k/ before /l/ (Hallé & Best, 2007; Hallé et al., 2003; Moreton, 2002).

In the Legal Prime Status condition, the parameter estimates (logs of the odds ratios, or logits) indicated that participants gave more of the higher ratings (i.e., *more dissimilar*) in the Control than in the Identity Trial Type condition,  $\beta = 5.50, z = 64.99, p < .001$ , and also in the Test than in the Identity condition,  $\beta = 4.51, z = 62.47, p < .001$ . There were also fewer of the highest ratings in the Test than in the Control condition,  $\beta = -1.12, z = -15.13, p < .001$ . In the Illegal Prime Status condition, participants also gave more of the higher ratings (i.e., *more dissimilar*) in the Control than in the Identity condition,  $\beta = 3.32, z = 59.34, p < .001$ , and in the Test than in the Identity condition,  $\beta = 0.64, z = 13.74, p < .001$ . However, unlike in the Legal Prime Status condition, the difference in the pattern of



**Fig. 3** Similarity ratings pooled across Participants, Voicing, and Items, and shown separately by Trial Type and Prime Status

responses was much smaller for the Test and Identity conditions: A logit of 0.64 corresponds to an odds ratio of just 1.89, which is more than an order of magnitude smaller than the odds ratio of 27.66 that corresponds to a logit of 3.32. Furthermore, the number of higher ratings was much smaller for Test than for Control Trial Types,  $\beta = -2.83$ ,  $z = -50.68$ ,  $p < .001$ . As was true in the other Prime Status conditions, in the Absent condition participants gave more of the higher ratings (i.e., *more dissimilar*) in the Control than in the Identity Trial Type condition,  $\beta = 3.73$ ,  $z = 59.38$ ,  $p < .001$ , and in the Test than in the Identity condition,  $\beta = 4.26$ ,  $z = 62.27$ ,  $p < .001$ . There were also more of the highest ratings in the Test than in the Control condition,  $\beta = 0.61$ ,  $z = 9.49$ ,  $p < .001$ , but this difference was again small.

To test the prediction that participants would rate primes and targets as more similar in the Identity Trial Type conditions and as more dissimilar in the Control Trial Type conditions across Prime Status, this pairwise comparison between Trial Types was repeated in a model that included interactions with Prime Status. As is shown in Table 2, we found a large effect of Trial Type across Prime Status conditions,  $\beta = 5.18$ ,  $z = 73.67$ ,  $p < .001$ . However, there were also significant interactions, indicating that the difference in the numbers of higher ratings for Control and Identity Trial Types (Fig. 3, top and middle rows) was smaller in the Illegal than in the Legal Prime Status condition,

$\beta = -2.00$ ,  $z = -24.63$ ,  $p < .001$ , and in the Absent than in the Legal condition,  $\beta = -1.46$ ,  $z = -17.54$ ,  $p < .001$ .

The patterns of similarity ratings for Test and Identity Trial Types were expected to differ across Prime Status. As is shown in Table 3, participants gave more of the higher ratings (i.e., *more dissimilar*) on Test than on Identity Trial Types across Prime Status conditions,  $\beta = 4.20$ ,  $z = 71.08$ ,  $p < .001$ . There was no difference in the patterns of responses for these two Trial Types in the Absent and Legal Prime Status conditions ( $p > .75$ ). However, when comparing the ratings for Test and Identity Trial Types for the Illegal relative to the Legal Prime Status condition, the significant interaction,  $\beta = -3.59$ ,  $z = -48.43$ ,  $p < .001$ , reflected far fewer of the higher ratings (i.e., *more dissimilar*) on the Illegal Test Trials.

The patterns of similarity ratings for Test and Control Trial Types were also expected to differ across Prime Status. Participants gave more of the higher ratings (i.e., *more dissimilar*) on both of these Trial Types in the Legal and Absent Prime Status conditions. However, as is shown in Table 4, the difference in the numbers of higher ratings for Test and Control Trial Types was larger in the Absent than in the Legal Prime Status condition,  $\beta = 1.57$ ,  $z = 16.95$ ,  $p < .001$ . Only on the Illegal Test trials did this pattern reverse, such that participants gave far fewer of the higher ratings (i.e., *more dissimilar*) on Test than on Control Trial Types for the Illegal as compared to the Legal Prime Status conditions,  $\beta = -1.92$ ,  $z = -22.41$ ,  $p < .001$ .

**Table 2** Estimates of fixed effects and thresholds for a mixed effects ordinal logistic regression model of similarity ratings comparing Control and Identity Trial Types in a model that includes interactions between Trial Type and Prime Status

	Estimate	SE	z	p
Fixed Effects Estimates				
Parameter:				
Control vs. Identity Trial Type	5.18	0.07	73.67	<.001
Illegal vs. Legal Prime Status	0.44	0.05	8.78	<.001
Absent vs. Legal Prime Status	0.39	0.05	7.56	<.001
Trial Type $\times$ Illegal vs. Legal Prime Status	-2.00	0.08	-24.61	<.001
Trial Type $\times$ Absent vs. Legal Prime Status	-1.46	0.08	-17.54	<.001
Threshold Estimates				
Threshold:				
1 vs. 2–4	0.83	0.22	3.85	
1, 2 vs. 3,4	1.94	0.22	8.95	
1–3 vs. 4	2.91	0.22	13.38	

SE = standard error. The thresholds show the logs of the cumulative odds ratios (logits) for lower ratings as compared to higher ones for three binary splits: 1 vs. 2 or more, 1 and 2 vs. 3 and 4, and 3 or less vs. 4. The predicted probabilities are entirely analogous to intercepts in binomial logistic regression

**Table 3** Estimates of fixed effects and thresholds for a mixed effects ordinal logistic regression model of similarity ratings comparing Identity and Test Trial Types in a model that includes interactions between Trial Type and Prime Status

	Estimate	SE	z	p
Fixed Effects Estimates				
Parameter:				
Test vs. Identity Trial Type	4.20	0.06	71.08	<.001
Illegal vs. Legal Prime Status	0.44	0.05	8.76	<.001
Absent vs. Legal Prime Status	0.38	0.05	7.62	<.001
Trial Type $\times$ Illegal vs. Legal Prime Status	-3.59	0.07	-48.43	<.001
Trial Type $\times$ Absent vs. Legal Prime Status	-0.02	0.08	-0.28	.78
Threshold Estimates				
Threshold:				
1 vs. 2–4	0.80	0.21	3.82	
1, 2 vs. 3,4	1.97	0.21	9.38	
1–3 vs. 4	2.85	0.21	13.54	

SE = standard error. The thresholds show the logits for lower ratings as compared to higher ones

**Table 4** Estimates of fixed effects and thresholds for a mixed effects ordinal logistic regression model of similarity ratings comparing Control and Test Trial Types in a model that includes interactions between Trial Type and Prime Status

	Estimate	SE	z	p
<b>Fixed Effects Estimates</b>				
Parameter:				
Test vs. Control Trial Type	-0.99	0.07	-14.43	<.001
Illegal vs. Legal Prime Status	-1.70	0.07	-25.81	<.001
Absent vs. Legal Prime Status	-1.19	0.07	-17.56	<.001
Trial Type × Illegal vs. Legal Prime Status	-1.92	0.09	-22.41	<.001
Trial Type × Absent vs. Legal Prime Status	1.57	0.09	16.95	<.001
<b>Threshold Estimates</b>				
Threshold:				
1 vs. 2–4	-4.89	0.29	-16.94	
1, 2 vs. 3,4	-3.62	0.29	-12.61	
1–3 vs. 4	-2.47	0.29	-8.64	

SE = standard error. The thresholds show the logits for lower ratings as compared to higher ones

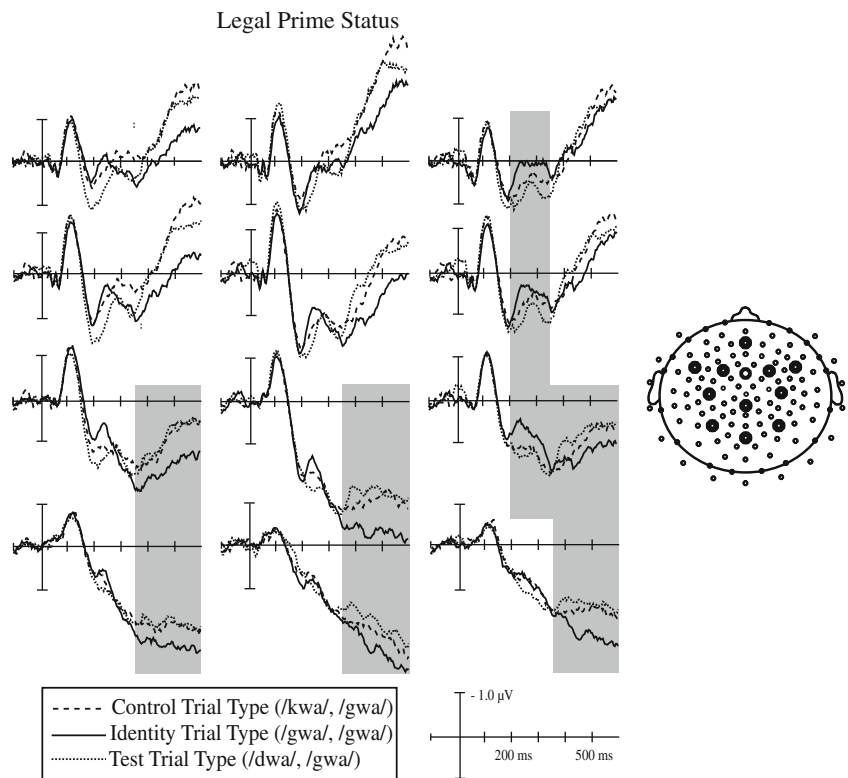
ERPs time-locked to targets

Figure 4 shows the grand average ERPs elicited by target syllables for all Trial Types in the Legal Prime Status condition; Fig. 5 shows the waveforms in the Illegal Prime Status

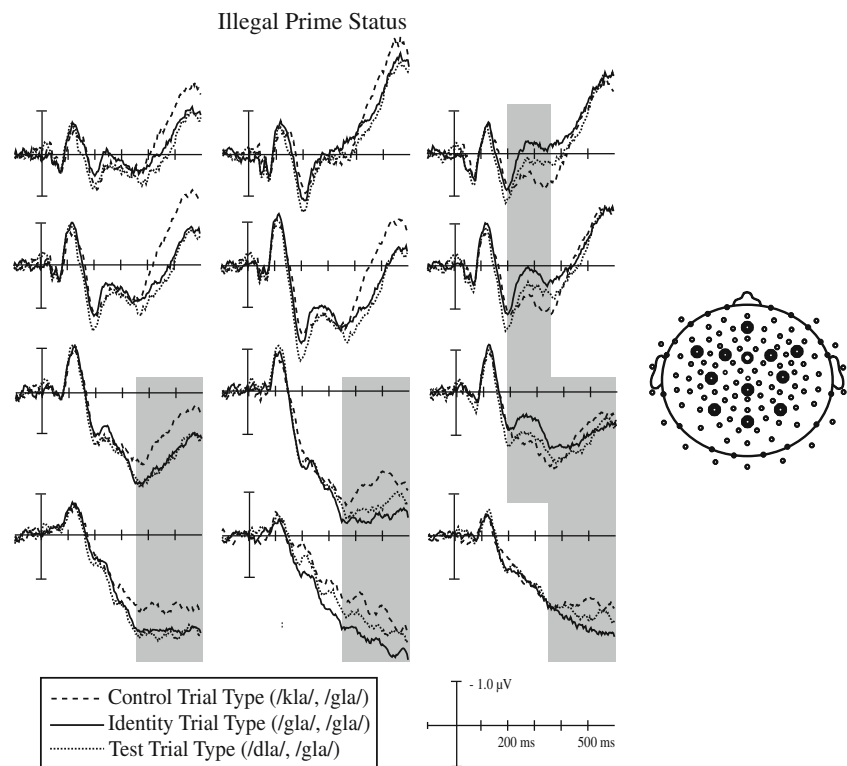
condition, and Fig. 6 in the Absent Prime Status condition. Target sounds elicited a typical positive–negative–positive series of peaks that were largest over anterior and medial regions. Targets elicited a first positive deflection (P1) that peaked at 70 ms post-stimulus-onset and a first negative deflection (N1) that peaked at 103 ms. There was no evidence of effects of Trial Type or Prime Status on the amplitude or latency of the P1 or N1 across the scalp ( $p > .50$ ).

*Mean amplitude 200–350 ms* Between 200 and 350 ms after onset, all targets elicited a positivity that varied in amplitude by condition. Greater amounts of priming, reflecting similarity between the representations of the primes and targets, were evident as a reduction in the amplitude of this positive deflection. The omnibus ANOVA revealed significant interactions between Trial Type, Anterior–Posterior electrode position, and Laterality of the electrode position,  $F(12, 204) = 2.98, p < .05$ , and between Trial Type, Prime Status, and Laterality,  $F(8, 136) = 3.62, p < .001$ . The differences in mean amplitudes in this time window were largest over the right hemisphere; we followed up the omnibus ANOVA with a  $3 \times 3 \times 4$  ANOVA with the factors Trial Type, Prime Status, and Anterior–Posterior electrode position on mean amplitudes measured at the right-lateralized electrodes only. This analysis revealed an interaction of Trial Type and Anterior–Posterior electrode position,  $F(6, 102) = 18.49, p < .001$ , driven by larger differences for the three Trial Types over right anterior, anterior/central, and posterior/central regions. The interactions between the Trial Type and

**Fig. 4** Event-related potentials elicited by target syllables preceded by Control primes (dashed lines), Identity primes (solid lines), and Test primes (dotted lines) in the Legal Prime Status condition. Negative is plotted up. Shading indicates the time range and scalp regions at which statistically significant differences between at least two conditions were observed. The electrode map indicates the approximate locations where the data in the figure were recorded



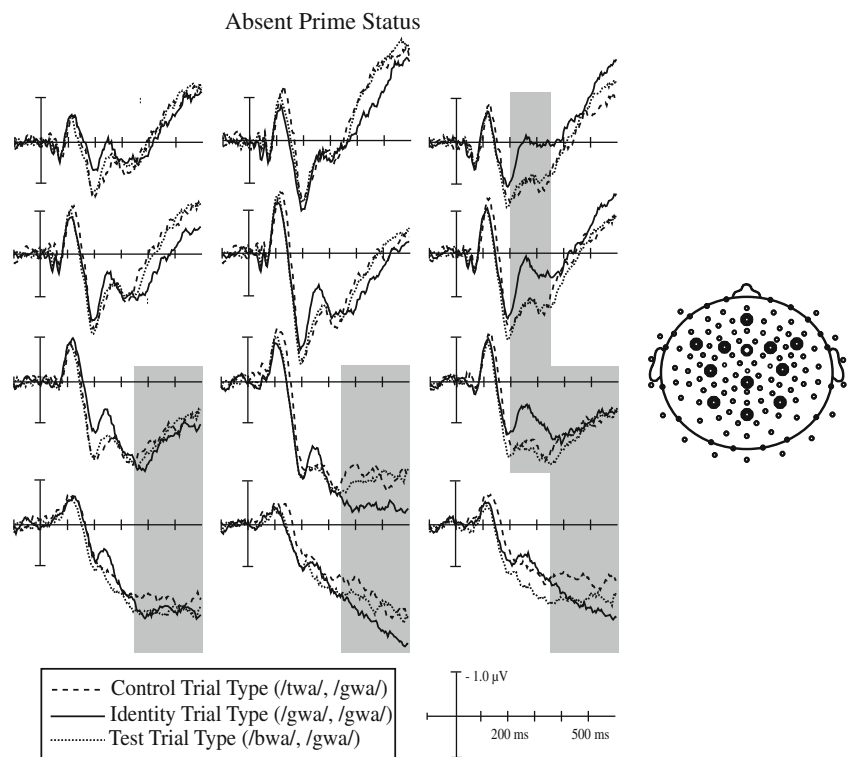
**Fig. 5** Event-related potentials elicited by target syllables preceded by Control primes (dashed lines), Identity primes (solid lines), and Test primes (dotted lines) in the Illegal Prime Status condition. Negative is plotted up. Shading indicates the time range and scalp regions at which statistically significant differences between at least two conditions were observed

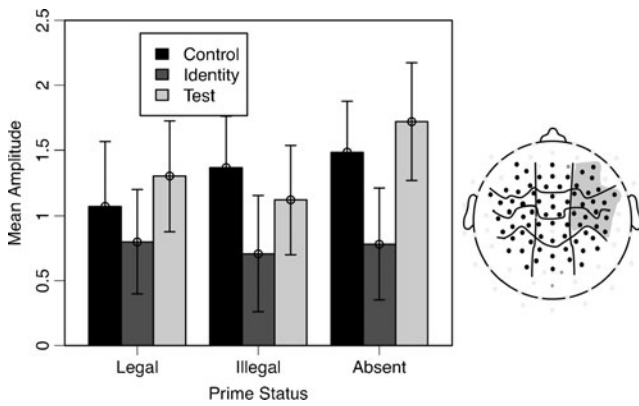


electrode position factors on measurements taken across the entire scalp motivated further analysis of the mean amplitudes 200–350 ms after target onset measured at all but the most posterior regions over the right hemisphere, using one-way ANOVAs with the factor Trial

Type for each Prime Status condition separately. The mean amplitudes measured at these electrodes (Fig. 7) revealed more priming—a less positive waveform—for Identity primes than for Test primes for the Legal, Illegal, and Absent Prime Status conditions.

**Fig. 6** Event-related potentials elicited by target syllables preceded by Control primes (dashed lines), Identity primes (solid lines), and Test primes (dotted lines) in the Absent Prime Status condition. Negative is plotted up. Shading indicates the time range and scalp regions at which statistically significant differences between at least two conditions were observed





**Fig. 7** Average amplitudes between 200 and 350 ms after target onset, measured over right anterior, anterior/central, and posterior/central regions. Error bars show standard errors

For the Legal Prime Status condition, a one-way ANOVA on mean amplitudes measured 200–350 ms after target onset at the right anterior, anterior/central, and posterior/central subsets of electrodes revealed a main effect of Trial Type (Control, Identity, Test),  $F(2, 34) = 4.08, p < .05$ . Paired  $t$  tests showed that the ERPs elicited by targets that followed Control primes and Identity primes did not differ,  $t(17) = 1.28, p = .22$ . Likewise, there was no difference for the Control and Test Trial Types,  $t(17) = 1.66, p = .12$ . However, relative to the ERPs for targets following Test primes, the positivity at 200–350 ms was reduced in amplitude for targets following Identity primes,  $t(17) = 2.95, p < .001$ .

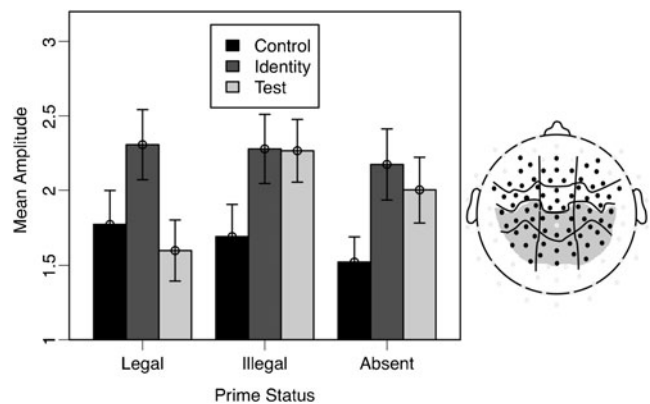
For the Illegal Prime Status condition, a one-way ANOVA on mean amplitudes measured 200–350 ms after target onset at the right anterior, anterior/central, and posterior/central subsets of electrodes revealed a main effect of Trial Type (Control, Identity, Test),  $F(2, 34) = 5.70, p < .01$ . Paired  $t$  tests showed that the positivity elicited by targets that followed Identity primes was reduced in amplitude as compared to that for targets that followed Control primes,  $t(17) = 3.04, p < .01$ . Similar to what was found for the Legal Prime Status conditions, in the Illegal Prime Status conditions there was no difference in the response to targets that followed Control and Test primes,  $t(17) = 1.48, p = .16$ . Again consistent with the Legal Prime Status conditions, in the Illegal Prime Status conditions, the positivity elicited by targets that followed Identity primes was marginally reduced in amplitude relative to that for targets that followed Test primes,  $t(17) = 2.03, p = .06$ .

For the Absent Prime Status conditions, a one-way ANOVA on mean amplitudes measured 200–350 ms after target onset at the right anterior, anterior/central, and posterior/central subsets of electrodes revealed a main effect of Trial Type (Control, Identity, Test),  $F(2, 34) = 12.70, p < .001$ . Paired  $t$  tests showed that the positivity elicited by targets that followed Identity primes was reduced in amplitude as compared to that for targets that followed Control primes,  $t(17) =$

4.74,  $p < .001$ . As was true for the other Prime Status conditions, there was no difference in the response to targets that followed Control and Test primes in this time window,  $t(17) = 1.28, p = .22$ . Also consistent with the other Prime Status conditions, the positivity elicited by targets that followed Identity primes was reduced in amplitude relative to that for targets that followed Test primes in the Absent Prime Status conditions,  $t(17) = 3.52, p < .005$ .

To determine whether any differences appeared in the patterns of Trial Type effects across the three Prime Status conditions, mean amplitudes 200–350 ms after the onset of targets measured at right anterior, anterior/central, and posterior/central electrodes were also entered into a Trial Type (Control, Identity, Test) by Prime Status (Legal, Illegal, Absent) ANOVA. The lack of interaction between these two factors when all three Trial Types were included ( $p = .09$ ), as well as when each of the possible combinations of two Trial Types was included ( $ps > .09$ ), supports the conclusion that the patterns of priming (i.e., Identity < Control = Test) did not differ for the measurements taken in this time window.

*Mean amplitude 350–600 ms* Between 350 and 600 ms after onset, all targets elicited a posterior positivity that was larger in amplitude for conditions in which participants reported the primes and targets to be highly similar. The omnibus ANOVA revealed a significant interaction between Trial Type, Prime Status, Anterior–Posterior electrode position, and Laterality of the electrode position,  $F(24, 408) = 2.49, p < .05$ . The differences between conditions were largest over posterior and posterior/central regions. For measurements taken at these posterior and posterior/central electrodes, there were no interactions between Trial Type or Prime Status and Laterality of electrode position ( $ps > .3$ ). The interactions on measurements taken across the entire scalp motivated further analysis of mean amplitudes 350–600 ms after target onset measured at the posterior and posterior/central electrodes, using one-way ANOVAs with the



**Fig. 8** Average amplitudes between 350 and 600 ms after target onset, measured over posterior and posterior/central regions. Error bars show standard errors

factor Trial Type for each Prime Status condition separately. The mean amplitudes measured at these electrodes are shown in (Fig. 8).

For the Legal Prime Status conditions, a one-way ANOVA on mean amplitudes measured 350–600 ms after target onset at the posterior and posterior/central subsets of electrodes revealed a main effect of Trial Type (Control, Identity, Test),  $F(2, 34) = 10.93$ ,  $p < .005$ . Paired  $t$  tests showed that ERPs elicited by targets that followed Identity primes elicited a larger positivity than did those that followed Control primes,  $t(17) = 3.12$ ,  $p < .01$ . There was no difference for the Control and Test Trial Types,  $t(17) = 1.19$ ,  $p = .25$ . Furthermore, ERPs 350–600 ms after target onset were also more positive for the Identity Trial Type than for the Test Trial Type,  $t(17) = 4.60$ ,  $p < .001$ . These results mirror the behavioral data, in that listeners rated primes and targets as being more similar in the Identity Trial Type condition than in both the Control and Test conditions, and the differences in ratings between the Control and Test conditions were small.

For the Illegal Prime Status conditions, a one-way ANOVA on mean amplitudes measured 350–600 ms after target onset at the posterior and posterior/central subsets of electrodes revealed a main effect of Trial Type (Control, Identity, Test),  $F(2, 34) = 8.85$ ,  $p < .005$ . Paired  $t$  tests showed that, similar to what was found for the Legal Prime Status conditions, in the Illegal Prime Status conditions targets that followed Identity primes elicited a larger positivity than did those that followed Control primes,  $t(17) = 3.20$ ,  $p < .01$ . Strikingly, in the Illegal Prime Status condition, targets that followed Test primes also elicited a larger positivity than did targets that followed Control primes,  $t(17) = 3.49$ ,  $p < .005$ , and amplitudes did not differ in the Identity and Test Trial Types in this time window,  $t(17) = 0.10$ ,  $p = .92$ . The finding that in the Illegal Prime Status condition targets that followed both Identity and Test primes elicited a larger positivity than did targets that followed Control primes also mirrors the pattern of the behavioral results: Primes and targets were rated as more similar to each other in the Identity and Test Trial Type conditions than in the Control Trial Type condition when the Test primes were Illegal.

For the Absent Prime Status conditions, a one-way ANOVA on mean amplitudes measured 350–600 ms after target onset at the posterior and posterior/central subsets of electrodes revealed a marginally significant effect of Trial Type (Control, Identity, Test),  $F(2, 34) = 3.30$ ,  $p = .06$ . Paired  $t$  tests showed that, as was true in the Legal Prime Status condition, in the Absent Prime Status condition, targets that followed Identity primes elicited a larger positivity than did targets that followed Control primes,  $t(17) = 3.02$ ,  $p < .01$ . However, there was some indication that targets that followed Test primes may have also elicited a

larger positivity than did targets that followed Control primes,  $t(17) = 1.93$ ,  $p = .07$ , and the responses to targets that followed Identity primes and Test primes did not differ in this time window,  $t(17) = 0.54$ ,  $p = .60$ .

To determine whether there were any differences in the patterns of Trial Type effects across the three Prime Status conditions, mean amplitudes 350–600 ms after the onsets of targets measured at posterior and posterior/central electrodes were also entered into a Trial Type (Control, Identity, Test) by Prime Status (Legal, Illegal, Absent) ANOVA. We found a marginally significant interaction between these two factors when all three Trial Types were included,  $F(4, 68) = 2.84$ ,  $p = .054$ . When only Identity and Control Trial Types were included in this analysis, there was no interaction with Prime Status,  $F(2, 34) = 0.12$ ,  $p = .88$ , indicating that the larger positivity in response to targets that followed Identity as compared to Control primes held across the three levels of Prime Status. However, when only Identity and Test Trial Types were included in this analysis, there was a marginal interaction with Prime Status,  $F(2, 34) = 3.56$ ,  $p = .058$ . These results support the conclusion that targets that followed primes that were reported to be similar (i.e., Identity primes in all Prime Status conditions, and Test primes in the Illegal Prime Status condition) elicited a larger posterior positivity 350–600 ms after onset than did the same targets that followed primes considered to be dissimilar.

## Discussion

The present experiment was designed to determine the extent to which at least some representations of nonnative speech sequences are immune to the application of listeners' knowledge of native-language phonotactics. Specifically, we tested the predictions of the autonomous and interactive hypotheses. The autonomous account predicts that speech sounds within phonotactically illegal clusters have veridical representations, along with the more abstract, categorical representations that are repaired to fit native-language phonotactics. The interactive hypothesis predicts that the phonotactic requirements of the listener's native language will influence all speech representations. In addition to testing the extremes of phonotactically legal and illegal clusters, we also measured the effects of hearing phonotactically absent primes on processing target syllables. The behavioral and ERP results supported two main conclusions. First, at least some of the representations of illegal sequences were veridical under conditions that replicated the effects of native-language phonotactic patterns on behavioral responses. Second, onset clusters that were legal but absent (i.e., accidentally missing rather than prohibited) had veridical representations at all levels, including those reflected in behavioral responses.

Replicating prior work (Hallé & Best, 2007; Hallé et al., 2003; Hallé et al., 1998; Massaro & Cohen, 1983; Moreton, 2002), we demonstrated that English listeners explicitly report hearing illegal onset strings as legal syllables that differ by one phonological feature. Specifically, participants rated phonotactically illegal syllables (e.g., /dla/) as being highly similar to legal syllables that differed only in place of articulation (e.g., /gla/). In contrast, phonotactically legal syllables (e.g., /kla-/gla/) were rated as highly dissimilar. The same was not true for the absent-but-legal prime syllables (e.g., /bwa/), which were rated as dissimilar from targets (e.g., /gwa/). These behavioral results inform our understanding of the differences between sequences that do not occur in a language because they are actively prohibited by that language's phonotactic constraints, and sequences that represent accidental gaps in the lexicon. Our results clearly demonstrate that illegal, prohibited sequences are repaired to conform to the language's phonotactic requirements, whereas absent-but-legal sequences are not. This evidence helps to explain why English speakers can successfully produce the absent clusters /bw/ and /pw/ in loan words from Spanish such as *bueno* and *pueblo*, but repair the prohibited cluster /tl/ to /kl/ in the loan word *Tlingit*. However, it is also important to note that defining the difference between the onset clusters /dl, tl/ and /bw, pw/ as being illegal or legal may not be the only important distinction. Since we used only two types of each class of item, it is possible that some other difference between these pairs of onsets drove the distinct patterns of behavioral results. Additional research including a greater number of illegal and absent sequences in languages other than English and with greater variability along other dimensions will be necessary to determine whether legality is the critical factor in observing behavioral effects of assimilation.

In contrast, the ERP results indicated that hearing a phonotactically illegal cluster affects early processing of a subsequent legal cluster in a manner that is not captured by behavioral responses. Neither Illegal primes (e.g., /dla/) nor Control primes (e.g., /kla/) facilitated processing of a target (e.g., /gla/) to the same extent as did different tokens of the same syllable (e.g., /gla/). The facilitation in processing of the legal, attested target syllables when preceded by the same syllables was observed as a reduction in the positivity elicited 200–350 ms after the onset of all targets. In contrast, between 350 and 600 ms after target onset, Illegal and Identity primes exerted a similar influence over target processing. These results are consistent with the autonomous model of speech perception, which predicts that at least some representations of speech sounds are veridical and that only categorical representations are regulated by language-specific phonotactic constraints. Importantly, the differences between the representations of illegal clusters and the legal sequences that they are reported to be similar to are not

ephemeral; they continue to influence the perceptual processing of other legal syllables presented 1,200 ms later.

Our results are inconsistent with those reported by Dehaene-Lambertz et al. (2000), who observed no differences in ERPs recorded from Japanese speakers dependent on whether there was or was not a switch between legal (e.g., /igumo/) and illegal strings (e.g., /igmo/). Our results may have differed for at least three reasons. First, we investigated a different phonotactic constraint. Participants in Dehaene-Lambertz et al.'s study had to detect whether a sound had been added or subtracted, while participants in our study had to detect a change in a sound's features. That their participants showed no evidence of representing the illegal sequence suggests, perhaps paradoxically, that a feature change is more salient than adding or subtracting an entire sound. To determine whether the critical factor is whether the phonotactic violation affects the prosodic shape of the whole word, future studies that include both types of manipulations will be necessary. Second, Dehaene-Lambertz et al. compared results across participants whose native languages differed, but we compared results within participants whose native language was the same. Our within-subjects measures may have allowed us to detect previously unreported differences in the representations of the illegal clusters and their legal counterparts. In particular, the French and Japanese speakers may have differed in how engaged they were in detecting a change between standards and deviants. The inability of Japanese speakers to hear differences in any of the critical stimuli may have led them to stop attending to and forming detailed representations of the sounds. In contrast, in the present experiment participants reported hearing differences between the prime and target more than 50 % of the time. Third, both the tasks and the ERP measures employed in the two studies were quite different, which could explain the conflicting results. Specifically, Dehaene-Lambertz et al.'s mismatch paradigm may have reflected neural activity during the formation of early representations; in the present study, ERP effects were driven by the relationship between phonotactically legal sequences and the stored representations of syllable strings presented 1,200 ms earlier. Differences in the initial and stored representations of speech sounds might explain the differences in the results of the two studies. The structures of the languages being tested and the position of the illegal sequences at the beginning rather than the middle of the stimuli may also have contributed to the early ERP effects reported in the two studies. However, the finding that listeners differentially processed legal sequences that were preceded by Identity and by Illegal Test primes in the present study indicates that there are veridical representations of at least some phonotactically illegal sequences that the paradigm and stimuli employed by Dehaene-Lambertz et al. failed to capture.

The ERP results also distinguish between perceptual, veridical representations and categorical ones influenced by language experience. Those perceptual representations may consist of the auditory qualities evoked by the stimuli, and the representations of /d/ or /t/ and /g/ or /k/ differ enough in their values for those auditory qualities before /l/ that they can be distinguished there, as they are in other contexts. The categorical representations would consist instead of the feature values that distinguish one phoneme from another, such that representations of /d/ or /t/ and /g/ or /k/ would no longer differ before /l/, as they do in other contexts, because the phonotactics prohibit stop consonants specified for the coronal place feature in that context. That is, representations consisting of auditory qualities are necessarily veridical, perhaps because they are produced at a prelinguistic or linguistically naïve stage in processing, but those consisting of feature values are not veridical if the grammar does not permit them to be. These characterizations of different types of representations are compatible with our ERP data, but our results do not require that the contents of the representations differ in this way. The perceptual representations could instead consist of language-specific phones (Norris et al., 2000) rather than auditory qualities, and the architecture of processing could delay the application of phonotactic constraints until later, perhaps because the evaluation of each phone is first context-free, and only later context-sensitive.

One way to determine what the contents of the representations are that influence the early and late processing of subsequent speech sounds would be to manipulate the discrimination task. Fujisaki and Kawashima (1969, 1970) showed that listeners discriminated within-category stimuli significantly better than was predicted by their categorization of those same stimuli. The researchers interpreted this result as evidence that listeners could use two kinds of memory of the stimuli to discriminate them. Using auditory representations would permit listeners to discriminate physically different stimuli that belong to the same category, while relying on phonetic representations would instead limit them to discriminating stimuli that belonged to different categories. In line with this distinction, Pisoni (1973) demonstrated that listeners' ability to discriminate stimuli to which they assigned the same label varied inversely with the interstimulus interval, a result that represents the fading of auditory but not of phonetic memory. The onset asynchrony between prime and target in our stimuli, 1,200 ms, was, however, far longer than Pisoni found to be sufficient for auditory representations to have faded. One possibility is that asking listeners to give similarity ratings in the present study encouraged longer maintenance of the auditory representations. However, it is not clear why providing similarity judgments on a 4-point scale, which could essentially be interpreted as a same/different decision with a two-level

confidence rating, would result in fundamentally different memory processes than would a discrimination task. Another possibility is that the order in which the veridical and more abstract categorical representations are used may be in the opposite direction. This possibility is supported by research on word recognition processes by McLennan, Luce, and Charles-Luce (2003) and McLennan and Luce (2005), who argued, within the adaptive resonance theory (ART) framework of Grossberg (1986), that more-frequent phonological features affect perception more quickly than infrequent features do. Under this view, abstract categorical (i.e., frequent) representations are accessed more quickly than veridical (i.e., less frequent) representations during word recognition. Initial representations, similar to those measured in Dehaene-Lambertz et al. (2000), may be more categorical than the longer-stored representations indexed in the present study. The time that elapsed between the prime and the target in the present study may have allowed listeners to use more detailed, exemplar-based representations rather than more general, categorical representations.

Whether it is based on initial representations stored for an extended period of time or less-frequent representations that require time to access, the difference between processing targets that followed Illegal Test and Identity primes may not reflect a raw auditory representation of the prime. Within-category discrimination depends on the format of the discrimination task (Carney, Widin, & Viemeister, 1977; Gerrits & Schouten, 2004; Pisoni & Lazarus, 1974). Furthermore, response times are systematically slower for stimuli near category boundaries than for those near the center of a category (Pisoni & Tash, 1974), and eye movements to incorrect competing words are more frequent for stimuli close to a boundary between categories (McMurray, Aslin, Tanenhaus, Spivey, & Subik, 2008). These results indicate that the representations on which listeners base manual responses and eye movements grade continuously with stimulus properties and are not reduced to unanalyzable, categorical representations. Our ERP results measured 200–350 ms after target onset further indicate that these detailed representations of the stimuli's auditory qualities are maintained, at least for a time, in a manner that modulates processing of subsequent signals.

A comparison of the early ERP effects observed in the present study with previously reported ERP effects can also inform the characterization of the representations that can be used to differentiate between illegal onset clusters (e.g., /dl/) and the legal clusters that they are typically repaired to (e.g., /gl/). The first possibility to consider is that, even though we did not use an oddball paradigm, the differences in ERPs observed 200–350 ms after the onset of targets could be related to a mismatch response such as that observed in Dehaene-Lambertz et al. (2000). However, in the present study the largest negativity (or reduced positivity) in this



time window was observed when the same syllable was presented as both prime and target (i.e., the Identity Trial Type). If it were the case that the single presentation of a prime was sufficient to serve as a standard, then targets that differed from those primes would be expected to elicit a larger MMN.

A second possibility is that the differences observed 200–350 ms after the onset of targets are in fact instances of the phonological mapping negativity (PMN), which is elicited between 200 and 350 ms after target onset when auditory input does not match phonological expectations (Connolly & Phillips, 1994; Desroches, Newman, & Joanisse, 2009; Newman & Connolly, 2009). For example, Newman and Connolly demonstrated a larger PMN when participants imagined a word or nonword (e.g., *snap* without the /s/ or *snoth* without the /s/), then heard a mismatching prompt (e.g., *tap*, *toth*), as compared to when the prompt matched the imagined word or nonword. Desroches et al. demonstrated a PMN to targets preceded by rhymes (e.g., *bone* preceded by *cone*) rather than by identical words (e.g., *cone* preceded by *cone*), but not to targets preceded by cohort pairs (e.g., *comb* preceded by *cone*). These results are interpreted as evidence that the PMN reflects a mismatch between top-down expectations and the phonological input. However, the present results are not consistent with a PMN interpretation, as the pattern of results that we observed is opposite in polarity. Specifically, listeners could have generated expectations that targets would be similar to the primes, as is supported by the later ERP effects discussed below, or that targets would be different from the primes, which is consistent with the fact that the primes and targets did differ on a majority of trials. If the mismatch between expectations of targets that were similar to the primes (e.g., /gwa/–/gwa/) and actual targets that differed from them (e.g., /gwa/–/kwa/) had generated a PMN, the largest negativity would be observed in Control conditions (e.g., /gwa/–/kwa/). However, it was precisely in these conditions that we observed the largest *positivity*. Expectations that targets would differ from primes (e.g., /gwa/–/kwa/) could elicit a PMN on Identity trials when those expectations were wrong (e.g., /gwa/–/gwa/). However, if this were the case, a larger negativity should also have been observed when listeners were wrong about precisely how the targets would differ from the primes (e.g., /gwa/–/twa/), but this was not true in the present study.

A more likely explanation for the differences observed 200–350 ms after target onset is that all sounds typically elicit a positivity during this time window (P2, part of the positive–negative–positive oscillation), which is reduced in amplitude for sounds preceded by similar sounds. As such, the more similar the target is to the prime, the smaller the positivity in this time window would be. These results are consistent with those of previous ERP priming studies. For

example, Coch et al. (2005) observed a reduced posterior negativity in addition to a larger anterior negativity in response to nonword targets preceded by rhyming nonword primes (e.g., *blane*–*hane*), as compared to targets preceded by nonrhyming primes (e.g., *crute*–*hane*). A priming effect would not be expected in the present study if listeners failed to actively engage in a phonological processing task; rhyme priming effects were not observed when stimuli were presented with a long SOA and a semantic judgment task (Perrin & García-Larrea, 2003). In the present study, the fact that the acoustic similarity of primes and targets affected processing of the target after a long delay of 1,200 ms suggests that listeners can store an accurate perceptual representation over an extended period of time, and that the similarity rating task encouraged them to do so.

Additional studies will be needed to define the critical differences between experiments that result in a PMN and those that result in priming. At least three methodological distinctions may determine whether speech sounds that are incongruent with the previous context elicit a larger negativity or a larger positivity during the same portion of the ERP waveforms. First, the PMN is often observed when listeners can make predictions about the upcoming material. The PMN is largest when listeners' predictions about upcoming words or nonwords are not borne out when those predictions are based on alliteration (Diaz & Swaab, 2007), rhyme (Desroches et al., 2009), sound imagery (Newman & Connolly, 2009; Newman, Connolly, Service, & McIvor, 2003), or semantic context (Connolly & Phillips, 1994; Hagoort & Brown, 2000; van den Brink, Brown, & Hagoort, 2001). In the present study, the primes and targets differed from each other so frequently that listeners may have abandoned the process of making predictions in the face of frequent errors, allowing the priming effects to be observed. Second, PMN experiments and the present study differ in the amounts of difference between the congruent and incongruent items. In studies designed to elicit a PMN, typically larger differences are used between the incongruent items, including different words (Connolly & Phillips, 1994; Diaz & Swaab, 2007; Hagoort & Brown, 2000; van den Brink et al., 2001), words that have undergone phoneme deletion (Newman & Connolly, 2009; Newman et al., 2003), or a change of multiple phonetic features (Desroches et al., 2009). In experimental contexts in which differences are more pronounced, listeners may focus on those differences, resulting in a mismatch response. In experimental contexts in which all of the items are quite similar to each other, as in the present study, listeners may focus on similarity, resulting in a priming effect. A third difference between the present study and those reporting a PMN relates to the task that participants engage in. Priming effects are typically observed when listeners are required to make explicit same/

different comparisons between primes and targets. Studies showing the PMN, on the other hand, often rely on more implicit comparisons, including semantic acceptability (Connolly & Phillips, 1994), matching (Desroches et al., 2009; Newman & Connolly, 2009; Newman et al., 2003), target recall (Diaz & Swaab, 2007), and listening for comprehension (Hagoort & Brown, 2000; van den Brink et al., 2001).

The second ERP effect that we observed was a posterior positivity 350–600 ms after the onset of targets that were rated as being similar to their primes. This effect is likely an example of a P3, or P300. The P3 is typically described as indexing recognition that a stimulus is relevant to the task (Donald & Little, 1981; Squires, Hillyard, & Lindsay, 1973; Sutton, Braren, Zubin, & John, 1965). Furthermore, P3 amplitudes have been shown to be larger when participants are highly confident that a stimulus should be categorized as a target on the basis of the task requirements (Squires et al., 1973). As such, the amplitude of the P3 is expected to correspond to behavioral responses; in the present study, the posterior positivity closely mirrors the behavioral results. That is, prime–target pairs that were rated as being very similar—both Identity trials in all three Prime Status conditions and Test trials in only the Illegal Prime Status condition—elicited a larger P3 amplitude. This result is consistent with the idea that listeners considered similar prime–target pairs to be the stimuli that they were monitoring for, perhaps encouraged by the description of the task as being similarity rather than dissimilarity judgment.

The results of the present study are consistent with the autonomous hypothesis of speech perception, in which perceptual representations of speech are veridical, and only higher-level representations are mapped onto language-specific categories. This finding raises the question of why both behavioral responses and the measures employed in Dehaene-Lambertz et al. (2000) reflect categorical representations rather than veridical auditory representations or context-free, language-specific phones. It is likely far too easy for listeners' behavior to be driven by the assignment of speech sounds to linguistic categories. Most behavioral tests of speech perception explicitly or implicitly rely on eliciting categorical responses. More fundamentally, the information that a speech sound conveys is embodied in category values. Although our ERP results show that listeners' phonotactic knowledge is not applied to all of their representations of speech sounds, it is applied to some of the representations, resulting in effects in the later portions of the waveforms that mirror behavioral data. This pattern of data suggests that listeners can maintain the veridical representations of speech sounds even after other representations have been repaired to be phonotactically legal. However, access to the veridical representations may be denied at the stages of processing that are reflected in most behavioral and MMN responses. Denial

of access is expected in a strictly feed-forward autonomous model, because the content of earlier representations cannot be revised by referring to the content of later ones. As such, our results can inform the further specification of models of speech perception. Specifically, the data indicate that accurate models of speech processing must include veridical representations of speech sounds that can be maintained over time or graded, categorical representations that reflect the difference in the acoustics of typical and atypical utterances.

Furthermore, the behavioral and ERP data described here converge in distinguishing the representations of onset clusters that are accidentally absent from the language from those that are actively prohibited by a phonotactic constraint. Mere absence is not enough to force a revision of even the categorical representations of consonant clusters, reflected in both behavior and the later portions of the neurophysiological response. These findings support a model of language learning that is based on something in addition to detecting statistically robust generalizations. If such generalizations were all that were learned, the patterns of results would have been the same for Illegal and Absent primes. Instead, the active prohibition against onset clusters such as /dl/ and /tl/, observed across a large number of languages, must be driven by linguistic knowledge beyond statistical absence. The results reported by Berent et al. (2007) suggest that /dl/ and /tl/ may be disfavored because /l/ differs too little in sonority from /d/ and /t/. However, since the results for these specific clusters were not distinguished from the results for others (including /bn/, /tm/, /km/, and /bw/) in Berent et al.'s (2007) analysis, additional research will be necessary to determine whether the phonotactic learning that goes beyond statistical generalizations can be attributed to principles such as the pattern of sonority in onset clusters.

**Author note** The research reported here was partially funded by NIH Institutional Training Grant MH16745, which provided postdoctoral training for M.B., R01-DC006241 awarded to J.K., and a John Merck Fellowship in the Biology of Developmental Disabilities awarded to L.D.S. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH, NIDCD, NIMH, or the John Merck Foundation. We thank Michael Becker and Shai Cohen for assistance with stimulus preparation, Martin Nolet and Paul Dainton for assistance with data collection and analysis, and Lori Astheimer, Patrick Taylor, Benjamin Zobel, Nicolas Planet, and William Bush for comments on previous versions of the manuscript.

## References

- Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology*, 26, 75–108.
- Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we never heard: Evidence from perceptual illusions. *Cognition*, 104, 591–630.

- Brown, R., & Hildum, D. (1956). Expectancy and the perception of syllables. *Language*, 32, 411–419.
- Carney, A., Widin, G., & Viemeister, N. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, 62, 961–970.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, 1, 351–353. doi:10.1038/1561
- Christensen, R. (2011). Ordinal-regression models for ordinal data (R package version 2011.08-12) [Computer software manual]. Available from [www.cran.r-project.org/package=ordinal/](http://www.cran.r-project.org/package=ordinal/)
- Coch, D., Grossi, G., Coffey-Corina, S., Holcomb, P., & Neville, H. (2002). A developmental investigation of ERP auditory rhyming effects. *Developmental Science*, 5, 467–489.
- Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP non-word rhyming effects in children and adults. *Journal of Cognitive Neuroscience*, 17, 168–182.
- Connolly, J., & Phillips, N. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6, 256–266.
- Cycowicz, Y., Friedman, D., & Rothstein, M. (1996). An ERP developmental study of repetition priming by auditory novel stimuli. *Psychophysiology*, 33, 680–690.
- Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, 34, 104–137.
- Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics*, 38, 272–288.
- Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience*, 12, 635–647.
- Desroches, A., Newman, R., & Joanisse, M. (2009). Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity. *Journal of Cognitive Neuroscience*, 21, 1893–1906.
- Diaz, M. T., & Swaab, T. Y. (2007). Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Cognitive Brain Research*, 1146, 85–100.
- Donald, M., & Little, R. (1981). The analysis of stimulus probability inside and outside the focus of attention, as reflected by the auditory N1 and P3 components. *Canadian Journal of Psychology*, 35, 175–187.
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568–1578.
- Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language & Cognitive Processes*, 16, 491–505.
- Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, 67, 224–238. doi:10.3758/BF03206487
- Elman, J., & McClelland, J. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27, 143–165.
- Friedrich, C., Schild, U., & Röder, B. (2009). Electrophysiological indices of word fragment priming allow characterizing neural stages of speech recognition. *Biological Psychology*, 80, 105–113.
- Fujisaki, H., & Kawashima, T. (1969). On the modes and mechanisms of speech perception. *Annual Report of the Engineering Research Institute*, 28, 67–73.
- Fujisaki, H., & Kawashima, T. (1970). Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute*, 29, 207–214.
- Ganong, W. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 110–125.
- Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, 66, 363–376. doi:10.3758/BF03194885
- Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. 1. Speech perception* (pp. 187–294). New York, NY: Academic Press.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, 38, 1518–1530. doi:10.1016/S0028-3932(00)00052-X
- Hallé, P., & Best, C. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop + /l/ clusters. *Journal of the Acoustical Society of America*, 121, 2899–2914.
- Hallé, P., Best, C., & Bachrach, A. (2003). Perception of /dl/ and /tl/ clusters: A cross-linguistic perceptual study with French and Israeli listeners. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2893–2896). Barcelona, Spain: Universitat Autònoma de Barcelona.
- Hallé, P., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, 24, 592–608.
- Hansen, J., Dickstein, P., Berka, C., & Hillyard, S. (1983). Event-related potentials during selective attention to speech sounds. *Biological Psychology*, 16, 211–224.
- Hillyard, S. A. (1981). Selective auditory attention and early event-related potentials: A rejoinder. *Canadian Journal of Psychology*, 35, 159–174.
- Hink, R., & Hillyard, S. A. (1976). Auditory evoked potentials during selective listening to dichotic speech messages. *Perception & Psychophysics*, 20, 236–242.
- Horváth, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., & Winkler, I. (2008). MMN or no MMN: No magnitude of deviance effect on the MMN amplitude. *Psychophysiology*, 45, 60–69.
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51, 141–178.
- Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, 13, 262–268. doi:10.3758/BF03193841
- Kraljic, T., & Samuel, A. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, 56, 1–15.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Magnuson, J., McMurray, B., Tanenhaus, M., & Aslin, R. (2003a). Lexical effects on compensation for coarticulation: The ghost of Christmas past. *Cognitive Science*, 27, 285–298.
- Magnuson, J., McMurray, B., Tanenhaus, M., & Aslin, R. (2003b). Lexical effects on compensation for coarticulation: A tale of two systems? *Cognitive Science*, 27, 801–805.
- Massaro, D. W., & Cohen, M. M. (1983). Phonological context in speech perception. *Perception & Psychophysics*, 34, 338–348.
- Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, 32, 543–562. doi:10.1080/03640210802035357
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1–86. doi:10.1016/0010-0285(86)90015-0
- McClelland, J. L., Mirman, D., & Holt, L. (2006a). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10, 363–369.

- McClelland, J. L., Mirman, D., & Holt, L. (2006b). Response to McQueen et al.: Theoretical and empirical arguments support interactive processing. *Trends in Cognitive Sciences*, *10*, 534.
- McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 306–321. doi:10.1037/0278-7393.31.2.306
- McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 539–553. doi:10.1037/0278-7393.29.4.539
- McMurray, B., Aslin, R., Tanenhaus, M., Spivey, M., & Subik, D. (2008). Gradient sensitivity to within-category variation in speech: Implications for categorical perception. *Journal of Experimental Psychology. Human Perception and Performance*, *34*, 1609–1631.
- McQueen, J. (2003). The ghost of Christmas future: Didn't Scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus, and Aslin (2003). *Cognitive Science*, *27*, 795–799.
- McQueen, J., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30*, 1113–1126.
- McQueen, J., Jesse, A., & Norris, D. (2009). No lexical-prelexical feedback during speech perception or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language*, *61*, 1–18.
- McQueen, J., Norris, D., & Cutler, A. (2006a). Are there really interactive processes in speech perception? *Trends in Cognitive Sciences*, *10*, 533.
- McQueen, J., Norris, D., & Cutler, A. (2006b). The dynamic nature of speech perception. *Language and Speech*, *49*, 101–112.
- Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*, 55–71.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *The Behavioral and Brain Sciences*, *13*, 201–288.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21.
- Näätänen, R. (2008). Mismatch negativity (MMN) as an index of central auditory system plasticity. *International Journal of Audiology*, *47*, S16–S20.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434. doi:10.1038/385432a0
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.
- Newman, R., & Connolly, J. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom-up and top-down effects on spoken word processing. *Biological Psychology*, *80*, 114–121.
- Newman, R. L., Connolly, J. F., Service, E., & McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potential. *Psychophysiology*, *40*, 640–647.
- Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *The Behavioral and Brain Sciences*, *23*, 299–370.
- Norris, D., McQueen, J., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204–238.
- Pakarinen, S., Takegata, R., Rinne, T., Huottilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology*, *118*, 177–185.
- Perrin, F., & García-Larrea, L. (2003). Modulation of the N400 potential during auditory phonological/semantic interaction. *Cognitive Brain Research*, *17*, 36–47.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Jr Johnson, R., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152. doi:10.1111/1469-8986.3720127
- Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, *13*, 253–260. doi:10.3758/BF03214136
- Pisoni, D. B., & Lazarus, J. (1974). Categorical and non-categorical modes of perception along the voicing continuum. *Journal of the Acoustical Society of America*, *55*, 328–334.
- Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, *15*, 285–290. doi:10.3758/BF03213946
- Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, *39*, 347–370. doi:10.1006/jmla.1998.2571
- Pitt, M. A., & Samuel, A. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 699–725.
- Praamstra, P., & Stegeman, D. (1993). Phonological effects on the auditory N400 event-related brain potential. *Cognitive Brain Research*, *1*, 73–86.
- R Development Core Team. (2007). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from [www.R-project.org](http://www.R-project.org)
- Samuel, A., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*, 416–434.
- Schröger, E., & Eimer, M. (1997). Endogenous covert spatial orienting in audition: “Cost-benefit” analyses of reaction times and event-related potentials. *Quarterly Journal of Experimental Psychology*, *50A*, 457–474.
- Squires, K. C., Hillyard, S. A., & Lindsay, P. H. (1973). Vertex potentials evoked during auditory signal detection: Relation to decision criteria. *Perception & Psychophysics*, *14*, 265–272. doi:10.3758/BF03212388
- Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*, 1187–1188.
- van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, *13*, 967–985.
- Weber-Fox, C., Spencer, R., Cuadrado, E., & Smith, A. (2003). Development of neural processes mediating rhyme judgments: Phonological and orthographic interactions. *Developmental Psychobiology*, *43*, 128–145.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., & Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, *36*, 638–642. doi:10.1111/1469-8986.3650638