Training Alters the Resolution of Lexical Interference: Evidence for Plasticity of Competition and Inhibition

Efthymia C. Kapnoula and
Department of Psychological and Brain Sciences, Delta Center, University of Iowa

Bob McMurray
Department of Psychological and Brain Sciences, Department of Communication Sciences and Disorders, and Department of Linguistics, Delta Center, University of Iowa

Abstract

Language learning is generally described as a problem of acquiring new information (e.g., new words). However, equally important are changes in *how* the system processes known information. For example, a wealth of studies has suggested dramatic changes over development in how efficiently children recognize familiar words, but it is unknown what kind of experience-dependent mechanisms of plasticity give rise to such changes in real-time processing. We examined the plasticity of the language processing system by testing whether a fundamental aspect of spoken word recognition, *lexical interference*, can be altered by experience. Adult participants were trained on a set of familiar words over a series of 4 tasks. In the high-competition (HC) condition, tasks were designed to encourage coactivation of similar words (e.g., *net* and *neck*) and to require listeners to resolve this competition. Tasks were similar in the low-competition (LC) condition, but did not enhance this competition. Immediately after training, interlexical interference was tested using a visual world paradigm task. Participants in the HC group resolved interference to a fuller degree than those in the LC group, demonstrating that experience can shape the way competition between words is resolved. TRACE simulations showed that the observed late differences in the pattern of interference resolution can be attributed to differences in the strength of lexical inhibition. These findings inform cognitive models in many domains that involve competition/ interference processes, and suggest an experience-dependent mechanism of plasticity that may underlie longer term changes in processing efficiency associated with both typical and atypical development.

Keywords

competition; lateral inhibition; plasticity; spoken word recognition; eye movements

Language learning (and, by extension, development) is commonly seen as a process of acquiring novel linguistic knowledge (e.g., words, syntactic rules). However, less commonly considered is the fact that learners must also build an underlying system that supports the efficient use and processing of this information in real-time (cf., McMurray, Kapnoula, &

Correspondence concerning this article should be addressed to Efthymia C. Kapnoula, E11 Seashore Hall, Iowa City, IA 52242-1409. ; Email: ekapnoula@iowa.uiowa.edu

Gaskell, in press; McMurray, Horst, & Samuelson, 2012). These two aspects of language development are closely linked to and shaped by one another, yet most research on language learning and development has focused mostly on the first kind of change (i.e., acquiring knowledge). We examined the second kind of change, change that takes place in the structure of the system to support real-time processing. We examined this in the domain of lexical processing, asking whether the dynamics of online competition between known words can be shaped by experience. Studying how the language processing system is being shaped by the information we process may allow us to achieve a more comprehensive understanding of the interactive dynamics between real-time processing, short-term learning, and long-term language development.

Two Timescales of Lexical Development: Acquiring Versus Using Words

Research on word learning has shown that acquiring a novel word entails more than just knowing the word form and its meaning; it also requires the listener to embed this information into the broader network that forms the mental lexicon in a way that supports efficient real-time processing (Coutanche & Thompson-Schill, 2014; Gaskell & Dumay, 2003; Gupta & Tisdale, 2009; Kapnoula, Packard, Gupta, & McMurray, 2015; Leach & Samuel, 2007; Mani & Plunkett, 2011; McMurray et al., in press; Storkel, 2002; Swingley & Aslin, 2007; Swingley, Pinto, & Fernald, 1999). Such processing entails a variety of phenomena—building activation of a word form from the input, activating the semantic network, inhibition among words during recognition, feedback from lexical to perceptual representations, and so forth. But, fundamentally, these processes require listeners to rapidly and flexibly access stored knowledge from auditory inputs.

In the context of learning and development, this distinction highlights the difference between two types of changes in the system: (a) the learner must acquire information about a word (e.g., its sound pattern, meaning) and (b) the learner must adjust the processing system so that he or she can use words more efficiently (e.g., lexical access). Research has mostly focused on the first kind of learning—the acquisition of novel information/knowledge. However, there is also evidence for significant changes in processing over development (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Marchman & Fernald, 2008; Rigler et al., 2015; Swingley et al., 1999) and aging (Mattys & Scharenborg, 2014; Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014), and also as a result of learning (Magnuson, Tanenhaus, Aslin, & Dahan, 2003) and adaptation to specific properties of the input, for example, accented speech (Poellmann, Bosker, McQueen, & Mitterer, 2014; Witteman, Weber, & McQueen, 2010; Witteman, Bardhan, Weber, & McQueen, 2015).

For example, a number of studies by Fernald and colleagues have documented that, between 18 and 36 months of age, the speed at which children recognize highly familiar words increases dramatically over development (Fernald et al., 1998; Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008), and these changes have been documented through adolescence (Rigler et al., 2015). These are not mere maturational changes—research on development during toddlerhood, for example, has suggested that such changes in processing correlate with vocabulary knowledge over and above age (Borovsky, Elman, & Fernald, 2012; Marchman & Fernald, 2008; see also Werker, Fennell, Corcoran, & Stager,

2002), while efficiency of lexical access also predicts later language learning (Fernald et al., 2006; Marchman & Fernald, 2008). Moreover, work with older children has suggested these changes are not simply changes in speed of processing—development qualitatively alters the process by which words compete with each other during recognition (Rigler et al., 2015; Sekerina & Brooks, 2007).

A critical aspect of this competition is that it is linked to the degree of phonological overlap or similarity among words; competition helps sort out ambiguous inputs by suppressing alternative candidate words (i.e., competitors). At the same time, this network of competitors changes over development as children acquire more words, resulting in both more neighbors for a given word and, likely, greater overlap between them. This is thought to impact speech perception (Garlock, Walley, & Metsala, 2001; Metsala & Walley, 1998; Stager & Werker, 1997; Storkel, 2002; Walley, 1993; Walley, Metsala, & Garlock, 2003; Werker et al., 2002), forcing children to develop more precise phonological and lexical representations. However, such growth may also place increasing demands on real-time processing, requiring learners to acquire more robust processing strategies that resolve competition more efficiently in the face of a changing lexicon.

Consequently, it is perhaps not surprising that differences in such real-time competition dynamics among well-known words have been associated with cognitive aging (Revill & Spieler, 2012; Sommers & Danielson, 1999), as well as a variety of developmental communicative disorders like specific language impairment (Dollaghan, 1998; Mainela-Arnold, Evans, & Coady, 2008; McMurray, Samelson, Lee, & Tomblin, 2010) and dyslexia (Desroches, Joanisse, & Robertson, 2006; Huettig & Brouwer, 2015). This again suggests that these differences are important not just for overall efficiency of processing, but also as determinants of language outcomes.

All of these studies employed highly familiar words and, nonetheless, observed complex changes in real-time processing over development (or with growth of the lexicon). Consequently, these studies suggest that, over development, it is not only the amount of acquired information that changes, but also the way listeners use this information during real-time language processing. This tuning of how words are processed likely occurs over the life span, and via various learning mechanisms (cf., Ramscar et al., 2014).

Taken together, these studies suggest that the way in which we use words is constantly shaped by experience. However, the exact locus of the plasticity underlying these changes is not known. To address this question, we need to study not only the developmental trajectory, but also the plasticity of targeted components of the spoken word recognition system within shorter time windows. That is, long-term changes in spoken word recognition processes are likely the cumulative result of many short-term changes. By studying such short-term learning in the lab, we can identify which components of the lexical processing system may be plastic, and the kinds of experiences that can shape them.

The present study focused on a crucial component of real-time lexical processing: how listeners cope with *lexical interference*. This aspect of word processing is evidenced by the fact that recognizing a spoken word is slower when it has neighbors, that is, recognizing *cat*

is slower because of the presence of *cap*, *hat*, and *cut* in the lexicon (Luce & Pisoni, 1998). As we describe, such interference is common in word recognition and constitutes a major problem that listeners need to solve. It is usually thought to derive from inhibition among words (Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Luce & Pisoni, 1998; McClelland & Elman, 1986), a key component of the real-time dynamics of word recognition. Here we test whether systematic exposure to different patterns of lexical coactivation can affect how listeners recover from interference. If plausible, this kind of change could theoretically drive changes in real-time processing at a longer time scale (see also Protopapas & Kapnoula, 2015, for an analogous argument for visual word recognition).

Interference Between Words

Spoken words are never available in their entirety, and listeners must accumulate information over time to recognize a word. It is now well known that listeners do not wait for the entire word before activating potential candidates (Allopenna, Magnuson, & Tanenhaus, 1998; McMurray, Clayards, Tanenhaus, & Aslin, 2008; but see Galle & McMurray, 2012). As a result, listeners activate (in parallel) multiple words that are consistent with the (partial) input that has been received up to that moment. These lexical candidates then compete with each other until one is left (Marslen-Wilson, 1987). For example, while hearing *sandwich*, when only /sæ-/ has been heard, listeners also activate words like *sack*, *sandal*, and *Santa.* Later, when the /n/ arrives, *sack* can be ruled out, and at /w/, *sandal* and *Santa* can, and so forth, until only one word remains under consideration.

Early models argued that this competition can be implemented without any interaction between words. In these models, as time passes and more information becomes available, previously compatible candidates become *in*compatible and thus fall out of competition (Cutler & Norris, 1979; Marslen-Wilson, 1987). However, later evidence for some form of *interference* between active lexical candidates suggested that this kind of race process is not sufficient for describing how words are recognized—there is an additional need for some form of interaction or inhibition among words. Evidence for interference comes from studies examining words that are disambiguated at the same point in time. In such words, the number of potential competitors affects the speed of recognition (Luce & Pisoni, 1998; Vitevitch, Luce, Pisoni, & Auer, 1999; Vitevitch & Luce, 1998). This suggests that the degree to which competition is resolved is a function of both the unfolding input and the presence of competing representations that can interfere with the target.

Dahan et al. (2001) offered the strongest evidence that these interference effects derive from interactions among specific words (e.g., lateral inhibition) using a stimulus manipulation introduced by Marslen-Wilson and Warren (1994; based on work by Streeter & Nigro, 1979). In this paradigm, auditory stimuli are manipulated to temporarily boost the activation of a competitor word. Specifically, a target word (*net*) is cross-spliced with the onset portion of a competitor (*neck*) so that the coarticulatory information in the onset (particularly in the vocoid) temporarily misleads listeners as to the upcoming consonant (*ne$_{ck}$t*) cued by the release burst. This, in turn, temporarily misleads listeners as to the identity of the target word (they briefly lean toward *neck*), which is resolved when the release burst (/t/) is heard. Thus, this splicing manipulation results in a temporary boost of activation for a competitor. If there

is inhibition between the competitor and the target, this should lead to the initial suppression of the latter by the former. Consequently, when the final, disambiguating /t/ is heard, it is more difficult to fully activate the target (*net*) than if the competitor (*neck*) was not as active.

Dahan et al. (2001) presented stimuli like these to participants while monitoring the activation of the target word using the visual world paradigm (VWP). Fixations to the referent of the target word were slowed by the mismatching coarticulation (when it activated a competitor) relative to an intact word. They were also slowed relative to a control condition in which the coarticulatory mismatch did not activate another word (e.g., *ne$_p$t*, where *nep* is not a word), suggesting that the interference effect was specifically due to the activation of the competitor, and not because of the coarticulatory mismatch per se. Moreover, the fact that the competitor picture (e.g., the *neck*) was not displayed on the screen suggests that this interference must be deriving from more general lexical properties, not from a decision-level process driven by the specific task demands. Thus, this provides strong evidence for interlexical inhibition as a component of the observed interference effect (see also Kapnoula, Packard, Gupta, & McMurray, 2015).

While interference as a behavioral phenomenon is usually considered an impediment to word recognition, the Dahan studies have suggested that it is a marker of lateral inhibition, which is likely a highly useful component of word recognition. In interactive models like TRACE (McClelland & Elman, 1986), lateral inhibition is crucial for dealing with embedded words (see also Gow & Gordon, 1995) and coping with segmentation ambiguity. Moreover, under typical circumstances, inhibition among words can greatly speed word recognition of even isolated words; when one word accumulates more evidence than neighbors, inhibition can help it suppress them more rapidly, speeding lexical access. Consequently, it is quite plausible that changes in lexical inhibition could underlie phenomena like the increase in lexical processing efficiency over development (Fernald et al., 1998; Marchman & Fernald, 2008; Rigler et al., 2015), or that changes in inhibition may underlie individual differences in how lexical competition is resolved among people with specific language impairment (Dollaghan, 1998; McMurray et al., 2010) or dyslexia (Desroches et al., 2006; Huettig & Brouwer, 2015). If such pathways are plastic and can be modified by experience, it would suggest that both typical and atypical development of lexical processing dynamics can be tuned by experience (either adaptively or maladaptively), as well as point the way toward potential loci for intervention in the case of atypical development.

The plasticity of processes that drive the onset and resolution of interference (like interlexical inhibition) is a crucial issue in spoken word recognition given the increasing interest in the various loci over which plasticity occurs in this system (McMurray et al., 2012; Mirman, McClelland, & Holt, 2006; Norris, McQueen, & Cutler, 1995). However, even though there is substantial work examining the formation and acquisition of different aspects of lexical representation, much less is known about the plasticity of these aspects for familiar words, or about how such plasticity affects real-time processing. For instance, a number of studies have examined the circumstances under which newly learned words interfere with the recognition of a known word (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; Kapnoula et al., 2015). Earlier studies suggested that newly learned words require

sleep-based consolidation to interfere with familiar word recognition (Davis, Di Betta, Macdonald, & Gaskell, 2009; Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; Tamminen & Gaskell, 2008). This would seem to rule out any kind of immediate plasticity in inhibition among familiar words (though it would still be applicable to longer term, developmental change). However, recently Kapnoula et al. (2015) tested interference from newly learned words using the more sensitive Dahan et al. (2001) paradigm and found that inhibitory links can be formed within 20 min (see also Coutanche & Thompson-Schill, 2014; Fernandes, Kolinsky, & Ventura, 2009; Kapnoula & McMurray, 2015; Lindsay & Gaskell, 2013). While these word learning studies suggest that lexical inhibition from novel to familiar words can be established within a short time period, there is no evidence that the interaction between known words, once formed, can be strengthened or weakened. That is, it is possible that the inhibitory links (which drive interference effects) are formed initially during learning, but then remain fixed and are not plastic.

Addressing this question of plasticity bears broader theoretical importance, for a number of reasons. First, across domains, it is unknown whether competition, in general, and lateral inhibition, more specifically, is susceptible to change via short-term training. Second, and perhaps more important, it is unclear whether changes in inhibition may underlie the broad changes in lexical processing of familiar words that are observed over development and in developmental language disorders like specific language impairment. Finally, this may also have important applications; evidence for changes in the resolution of lexical interference may point to methods of remediating problematic patterns of interference in atypical populations, or improving cognitive function in the variety of domains in which interference is thought to play a role.

Lexical Interference Is Multiply Determined

In examining the plasticity of processes related to lexical interference, it is crucial to point out that interference (as a measurable phenomenon) is a product of multiple dynamic properties of word recognition. Clearly, *lateral inhibition* between words is a prerequisite for interference; it is what enables more active words to suppress less active ones (Dahan et al., 2001; Luce & Pisoni, 1998), and, without some form of inhibition, it is not clear that interference effects could be observed at all. However, in interactive models, lateral inhibition is only one route among several that could affect interference. Because the inhibition exerted by one word is a function of its degree of activation, anything that affects the activation of competing words may in turn alter interference. For example, the spread of activation from the perceptual or phonological representations to the lexical layer (i.e., bottom-up spread of activation) may affect the overall interference pattern (see McMurray et al., in press, for simulations). This could happen either by prematurely activating a competitor, leading to higher levels of interference, or by giving the target word a boost in activation, thus expediting the suppression of the competitor. Thus, altering specific bottom-up pathways via training may also alter how interference is established and/or resolved. Similarly, a word's activation in interactive models is also a function of its ability to maintain activation over time or to resist decay. For example, McMurray et al. (2010) conducted simulations of their data on specific language impairment with the TRACE model and found that increased levels of decay offered the best fit. In this case, increased decay led

to less stable activation for the target word, which in turn exerted less inhibition to competitors.

Given these complex underpinnings of lexical interference effects and the challenges in isolating any single parameter in a behavioral task, it is likely impossible to attribute any effects of training exclusively to lateral inhibition. Moreover, changes in inhibition may have complex consequences for word recognition. Increasing inhibition may speed processing under some circumstances (when the input is clear), because it helps suppress competing alternatives faster; but, it may also lead to stronger interference, if activation of a competitor word is temporarily boosted early on (e.g., in the Dahan paradigm). In contrast, decreasing inhibition could either protect the target from initial interferences from a noisy input, but it may also weaken the target's ability to suppress competitors later.

Our primary goal was not to determine the exact parameter setting responsible for whatever changes/differences in lexical interference we may observe with training. Rather, our goal was to address the broader question of whether training regimes can change the way people deal with lexical interference, and to investigate the more specific question as to whether training can alter the efficiency of recovering from interference. In some ways, this is a broader issue than whether those changes come specifically from lexical inhibition or from other changes to the system. However, after documenting such effects, we turn to computational modeling to pinpoint their locus in inhibition.

The Present Study: Plasticity of Processes Underlying Lexical Interference

The present study addressed this question (i.e., whether the way in which listeners deal with lexical interference can be adjusted via experience) using a training-and-testing paradigm. Participants were assigned to a high-competition (HC) or low-competition (LC) group. Both groups were exposed to the same series of training tasks involving the same pairs of familiar competitor words (e.g., *net*, *neck*). The structure of the tasks was manipulated to induce three critical differences that we thought would promote changes in how lexical interference is resolved. First, tasks were designed to create a greater need for interference resolution between competing words for the HC group and to minimize it for the LC group; only the HC participants were required to fully resolve competition between the two words to accurately respond. This creates differential task demands for the two groups in a way that may lead participants in the HC condition to learn to resolve interference more effectively because their tasks demand it. Second, tasks in the LC group were manipulated to provide an external (top-down) source of bias on competition, so that one of the two competing words was favored over the other. This facilitates recognition of the target word and minimizes the need for inhibition.[1] Third, tasks in the HC group were designed to repeatedly coactivate both competitors in each word pair. By coactivating competing words, like *net* and *neck*, over and over again, we aimed to strengthen the links between the two words in a Hebbian-

---

[1]The crucial difference between the first and the second property of our task manipulation is that, according to the first account, resolving interference is *irrelevant* for the low-competition (LC) group to make an accurate response. That is, the competitor does not need to be adequately suppressed, and a slight advantage of the target over the competitor should suffice for a response to be made. In contrast, in the second account, the structure of the task facilitates the *resolution of the completion* in the LC group, by boosting the target's activation.

like manner (Mirman et al., 2006). This should, in turn, maximize interference between them in the future.

All three goals were achieved by changing the response set between groups. To minimize the need to resolve inhibition, the responses were chosen to create a situation in which the competitor's activation was not important for responding in the task. For example, if the participant has to decide if a word is *cat* or *dog*, he or she only needs to hear the initial /k/ to make an accurate decision; whereas, if the decision is *cat* or *cap*, the competitor must be adequately suppressed for the participant to make an accurate decision. Our hypothesis was that repeatedly exposing HC subjects to the latter kind of demands (adequate suppression of the competitor) would lead the system to develop better ways of dealing with interference.

Changing the response set also changes the external support for one word over another. This may bias activation for one word, creating more interference (if that word is the competitor). For example, in one task, participants matched a spoken word to one of two visually presented orthographic strings (i.e., spoken to written word matching task). Here, the visually presented strings may provide top-down support for one or both of the spoken words under competition. In our case, if the target word is *net* and the response options are *net* or *cap* (an LC version), the visual input provides clear top-down support for the target, biasing the competition without requiring inhibitory involvement. In contrast, if the response options are *net* and *neck*, both words receive top-down support and thus more time will be required for the target to fully suppress the competitor.

Lastly, according to a Hebbian account, one might assume that the strength of the (inhibitory) connection between two words is a function of how often they are coactive. In this case, the repetitive copresentation of *net* in the presence of *neck* may lead to an increase in the strength of the inhibitory connection between them.

Clearly, all three factors are closely related and may conspire to lead to plasticity in the system. However, independent of the exact way in which our tasks may shape interference, this manipulation allows us to use the exact same set of training tasks with the exact same set of words—the only difference between training groups are in these task demands.

Training consisted of a set of four different training tasks, all of which could be manipulated along similar lines. We used multiple tasks for three reasons. First, our manipulation depended on a detailed analysis of the logic of individual training tasks; given that many of these tasks have not been well studied, it seemed like a good idea to hedge our bets by implementing these ideas across multiple tasks rather than relying on one. Similarly, it seemed likely that each of these three sources of learning could be tapped differentially by different tasks; the use of multiple tasks might offer a way to capture all of them. Second, and most important, a host of research from automatic motor skill learning has suggested that variability in the training task typically (but not always) leads to more robust acquisition of automatic skills (Magill & Hall, 1990; Shea & Morgan, 1979; Wulf & Shea, 2002).

Critically, we asked if these effects of training would generalize to a different task, and specifically one directly targeting interference among words. Thus, after training, we evaluated inhibition between specific words using the subphonemic mismatch/VWP design

of Dahan et al. (2001). This was intended to test as closely as possible the specific interference between one word (whose activation has been temporarily raised) and another (the target word). As in Dahan et al. (2001), in this paradigm, the release burst of the target word (the /t/ in *cat*) was spliced with either the onset of the same word (*ca$_t$t*; *matching-splice*), that of a competitor (*ca$_p$t*; *word-splice*), or of a nonword (*ca$_{ck}$t*; *nonword-splice*).

Participants heard these stimuli and selected the referent from a set of four pictures that included the target and three other words. Looks to the target picture were used as an index of its activation. Critically, the competitor word was not present on the screen. The measure from this paradigm is the difference in fixations between the word-splice and matching-splice conditions, which is taken as evidence that the competitor word (*cap*) is temporarily boosted in the word-splice condition and inhibits the target word (*cat*). Relatedly, the difference between the nonword-splice and word-splice conditions can show that the decreased activation from the word-splice is not just due to subphonemic mismatch, but also reflects lateral inhibition between words (see also Kapnoula et al., 2015). Because the competitor was not present on the screen, any interference must derive from internal interference between two lexical items. Therefore, this VWP/splicing paradigm is an established measure of lexical inhibition/interference. Significant differences between the two groups in their pattern of interference would suggest that the training did in fact shape how previously known words interact with each other. Our hypothesis was that the HC training—which heightens demands on interference resolution—would lead learners to become better at coping with interference, by showing less of it, and/or recovering more quickly.

Method

Participants—Eighty monolingual English speakers participated in this experiment. This number is typical of between-subjects visual world studies (Farris-Trimble, McMurray, Cigrand, & Tomblin, 2014; McMurray et al., 2010). It was planned from the outset of the study, and no additional subjects were run. Participants were students at the University of Iowa and received course credit as compensation. Six of them were excluded from the analyses due to technical problems. This left 37 participants in each of the two experimental groups.

Design—During training, all participants were exposed to 28 pairs of familiar words (56 words in total; see Appendix A) over a set of four tasks. Each participant was randomly assigned to one of two training groups: HC or LC. The stimuli and tasks were identical across groups, however, the details of the individual tasks were manipulated such that in the HC group words from the same pair would be more likely to be coactivated (or, conversely, so that coactivation would lead to interference that would need to be resolved for participants to be able to perform the task accurately). In contrast, in the LC group, tasks were structured to either minimize coactivation and/or not to require its resolution to the same degree. Thus, the HC training was designed to induce higher demands for lexical interference resolution than the LC training.

Immediately after training, participants from both groups performed the same VWP task designed to assess lexical interference between specific pairs of words using the same three splice conditions as Dahan et al. (2001). This allowed us to compare the pattern of lexical interference between the two experimental groups using proportion of looks to the target as a measure of the lexical activation of the target word, and examining this as a function of splicing condition to assess interference. Note that we did not include a pretest evaluation of lexical interference because we wanted to minimize the exposure of LC participants to the spliced stimuli used in testing. Because these spliced stimuli are specifically made to increase lexical interference, such exposure could have acted as a form of competition-inducing training, thus jeopardizing our chance of finding a significant difference between the two experimental groups.

Training—During training, participants performed four different tasks exposing them to 28 pairs of known monosyllable words (see Appendix A). The four training tasks were: (a) phoneme monitoring, (b) same/different, (c) spoken to written word matching, and (d) primed production. Each task was manipulated across the two experimental groups (HC and LC) to create circumstances under which either there was a higher (or lower) demand for interference resolution. Specifically, for the HC group, all four tasks aimed to boost the activation of both words in a pair, and to place high demands on interference resolution for accurate performance. In contrast, in the LC group, the tasks were designed to boost the activation of the target and to minimize interference from competitors (see Table 1). According to our hypothesis, the HC tasks should create prolonged coactivation of the two competing words in the HC group along with the higher demands on interference resolution (compared with the LC group). The repetition of these circumstances may in turn shape the way in which interference is resolved in the long term, and these effects may generalize in a context outside that of training.

Stimuli: Each stimulus pair used in training consisted of two monosyllabic consonant-vowel-consonant (CVC), CCVC, or CVCC words that ended in a stop consonant (/b/, /d/, /g/, /p/, /t/, or /k/). The two words in each pair overlapped in all phonemes but the last consonant (e.g., *net-neck*), which shared voicing but differed on place of articulation. Critically (for the VWP design used at test), the third unused place of articulation did not result in a word (e.g., *cat* and *cap* were words, but there is no word, *cack*). Auditory stimuli for each word were recorded by a male native speaker of American English in a sound-attenuated room, sampling at 44100 Hz. Over the course of training, participants heard many repetitions of each word. Therefore, we were concerned that participants might learn and rely on small acoustic differences between the stimuli (e.g., differences in the pitch), minimizing the demands on lexical competition (because the words could be distinguished by spurious acoustic differences). To avoid this, for each word, we selected four different recordings that were presented during training as different tokens of the same word. Auditory stimuli were delivered over high-quality headphones.

Procedure: Participants were seated in a sound-attenuated room and received instructions about the training portion of the experiment. Visual stimuli were presented in Calibri 10-point font size and they appeared on the screen as black characters on a white background.

DMDX ([Forster & Forster, 2003]) was used for the stimulus presentation and the collection of motor and vocal responses. The first time a task was first introduced, participants performed a few practice trials to ensure they understood the instructions. Participants received immediate feedback at the end of every trial.

Phoneme monitoring: During the phoneme monitoring task, participants were instructed to select the letter that matched the last sound in a target word (from two options). On each trial, participants saw a cross at the center of the screen for 200 ms, followed by the two letters (e.g., *k* and *t*) on the two sides of the screen for 750 ms. Then the two letters disappeared for 150 ms, followed by the auditory presentation of the target word (e.g., *neck*). Next, participants were reminded of the two options and the cross changed into a question mark, prompting them to respond. Participants had 5 s to press one of two keys indicating which of the two letters on the screen corresponded to the last sound of the word they heard (see Figure 1). They received immediate feedback on every trial. On phoneme monitoring trials, each word was presented twice and, thus, there were 112 total trials in this task (56 words × 2 presentations).

The pair of letters that participants chose from was fixed for each word, but manipulated between groups to enhance or reduce the demand for resolving competition between the two words. Participants in the HC group had to choose between the correct phoneme and the phoneme corresponding to the last sound of the other word in that pair. For example, when the word was *bait* (with a close competitor, *bake*), the two options were always /t/ and /k/ for the HC group. However, in the LC group, the second option never corresponded to the ending of the other word in the pair. For example, when the word was *bait*, the options were /t/ and /p/; and when the word was *bake*, the options were /k/ and /p/.

Same/different: During the same/different task, participants had to report whether two auditorily presented words ended in the same (e.g., *net* and *cat*) or a different sound (e.g., *net* and *neck*). Each trial started with a cross at the center of the screen for 150 ms, followed by the auditory presentation of the two words separated by 200 ms. Finally, participants were visually presented with the two options (*same* and *different*) on the two sides of the screen and had 5 s to respond by pressing one of two keys (see Figure 2). They received immediate feedback on every trial. Each word was presented four times: twice within a *same* trial and twice in a *different* trial. Because two words were presented in each trial, the total number of trials was 112 (56 words × 2 same/different trials × 2 presentations / 2 stimuli per trial).

The word pairs presented in each trial were fixed for each participant, but they differed between the two groups. In the HC group, same trials always consisted of the same word (e.g., *net–net*), whereas different trials always included two words from the same pair (e.g., *net–neck*). In contrast, in the LC group, same trials consisted of two different words that ended in the same sound (e.g., *net–cat*), whereas different trials included two words from different pairs that ended in a different sound (e.g., *net–park*). To keep the tasks as similar as possible between groups, the same/different pairs were fixed in the LC group as well (e.g., for a given participant *net* was paired with *cat* in same trials and with *park* in different trials throughout the task).

Spoken to written word matching: In the spoken to written word matching task, participants saw two words on the screen, then heard one of them, and responded by pushing a key to indicate which they heard. Each trial started with a cross at the center of the screen for 150 ms, followed by the visual presentation of the two words (e.g., *neck* and *net*) on the two sides of the screen for 750 ms. Then the two words disappeared for 200 ms, followed by the auditory presentation of the target word (e.g., *neck*). Next participants were reminded of the two options and the cross changed to a question mark prompting them to respond. Participants had 5s to press the key indicating which of the two words on the screen corresponded to the word they heard (see Figure 3). They received immediate feedback on every trial. Each word was presented twice, yielding 112 experimental trials in this task (56 words × 2 presentations).

The pairs of words appearing together in each trial were fixed; for the HC group the two words were always from the same pair (e.g., *net* and *neck*), whereas for the LC group, the two pairs were those used in the different trials in the same/different task. Because the words in the LC pairs were so different (e.g., *net–park*), we were concerned that participants would not wait until the end of the auditory stimulus to make a decision. To avoid this issue (and keep the two groups as similar as possible), we included 56 catch trials in which the target word was followed by a tone and participants were instructed to hit a separate key in that case, yielding 168 total trials in this task.

Primed production: In the primed production task, participants were asked to say out loud either a word they just saw or one they just heard. In each trial, participants first saw a cross at the center of the screen for 150 ms, followed by the visual presentation of a word for 750 ms. Then the cross appeared again for 200 ms, followed by the auditory presentation of a word. After the auditory presentation, the cross remained on the screen for another 150 ms. Then the cross disappeared and participants received a visual prompt indicating that they should respond orally with either the word they saw (prompted by an eye symbol, 👁), or the word the heard (prompted by an ear symbol, 👂; see Figure 4). Participants had 5 s to respond, and their response was recorded by a microphone located over the computer screen. The recording was initiated by the DMDX voice key and a "Recording response …" message appeared on the screen. If no response was recorded during that time, participants were notified by a "No response" message. In this task, each word was presented in both a matching and a mismatching condition, depending on whether the visually and auditorily presented words were the same or different. In addition, trials varied on whether participants were cued to repeat the word they saw or the one they heard. Each word was presented four times (2 × 2 [Matching Conditions × Modality Conditions]) yielding a total of 224 trials in this task (56 words × 4 presentations).

In this task, the critical manipulation occurred on the mismatching trials. For the HC group, the visually and auditorily presented words in the mismatching trials were from the same pair (e.g., *net–neck*) and thus would have engaged in competition (which would need to be resolved to identify the correct word and repeat it). In contrast, for the LC group, the word pairs in the mismatching trials did not overlap (e.g., *net–park*). Instead, the same pairs were used as those used for the LC group in the different trials in the same/different task.

Structure of tasks: The total set of trials for each task was randomized, split into four blocks, and presented in a mixed manner (see Table 2). To minimize task order effects, the order of tasks within a block was selected for each participant from one of four possible task orders. The order of the tasks was held constant across blocks within participant.

Testing—The purpose of the testing phase was to evaluate the strength and time course of lexical interference between the two words in each pair, and to determine whether this differed between training groups. Following Dahan et al. (2001), we created auditory stimuli in which the onset of the word (from the initial consonant through the vowel and closure) was either consistent or inconsistent with the release burst of the word. There were three conditions, all of which ultimately led to a percept of the same target word (e.g., *net*). In the word-splice condition, we spliced the onset from a recording of the competing word onto the release burst of the target word. For example, we took what would commonly be described as the *ne-* from *neck* and spliced onto the *-t* from *net* to create a stimulus ($ne_{ck}t$) in which the competitor word (*neck*) would be overactivated, and temporarily inhibit the target (*net*). This was compared with the matching-splice condition in which both the onset and the release burst of the auditory item came from different recordings of the same word (ne_tt). To control for the fact that the word-splice condition was also a poorer acoustic instantiation of the target word, we also compared it with the nonword-splice condition (ne_pt) in which the onset mismatched the release, but did not activate another word.

Lexical activation of the target was measured using the VWP; participants saw four pictures, including a picture of the target (*net*), plus three semantically unrelated words. Each word in the set of four was presented as the auditory stimulus three times, once for each of the three splice conditions. As a result, each set appeared on 12 trials (4 pictures × 3 splice conditions), and all four pictures in a given set had an equal probability of being the referent of the auditory stimulus. Based on the findings reported by Dahan et al. (2001), our prediction was that all participants would look less to the target in the mismatching ($ne_{ck}t$) than the matching (ne_tt) condition, while a smaller difference (or no difference) should be found between the matching and the control conditions (see also Kapnoula et al., 2015). In addition, we expected to find a difference in the pattern of interference between the two experimental groups either in terms of strength (e.g., stronger interference for the HC group) or in terms of time course (e.g., faster resolution of the competition in the HC group).

Stimuli: On each testing trial, the participant saw four pictures, one of which corresponded to a word from the 28 pairs used in the training portion (e.g., *neck* from the *neck-net* pair). The other word from that pair was never played and no picture of it was ever shown. The other three pictures were semantically unrelated to the target word, but one of them had an initial phoneme overlap with the target word (see Appendix B for a list of the visual stimuli used in testing).

We used a total of 112 pictures (28 target words × 4 pictures in each set). All pictures were developed using a standard lab procedure (Apfelbaum, Blumstein, & McMurray, 2011; McMurray et al., 2010); for each word, 5–10 candidate images were downloaded from a commercial clipart database and viewed by a group of 3–5 undergraduate and graduate lab members. One image was selected and subsequently edited to ensure a prototypical

depiction of the target word. The final images were approved by a lab member with extensive experience using the VWP.

Auditory stimuli were constructed by cross-splicing the onset portion and release burst from recordings of different items. Specifically, we took the release burst from each target word (e.g., *net*), starting at the onset of the release burst and until the end of the word, and spliced it onto the onset portion from three other recordings. This onset portion was taken from the beginning of each recording and included everything up to the onset of the release, including the closure. The onset came either from a separate recording of the same word ($ne_t t$), the competing word as assigned in training ($ne_{ck} t$), or a nonword ($ne_p t$; see Appendix C for a full list of the spliced stimuli). All stimuli were recorded by the same speaker in a sound-attenuated room at 44100 Hz. For each word, multiple tokens were recorded from which the tokens with the strongest coarticulation were selected for splicing. All stimuli were spliced at the zero crossing closest to the onset of the release. The three filler items in each picture set were also spliced using a similar procedure. The only difference was that there was no word-splice condition; each target word was spliced with itself or one of two nonwords.

Procedure: Immediately after training, participants were familiarized with the 112 pictures used in the VWP task; they saw each picture along with its orthographic label. They were then fitted with an SR Research EyeLink II head mounted eye-tracker. After calibration, participants were given instructions for the testing phase, and testing began. At the beginning of each trial, participants saw four pictures in the four corners of a 19-in. monitor operating at $1,280 \times 1,204$ resolution. A small red circle was also present at the center of the screen. After 500 ms, the circle turned blue, which cued the participant to click on it to start the trial. This delay allowed us to minimize eye movements due to visual search (rather than lexical processing), because participants could take a brief look at the pictures before hearing anything. The blue circle disappeared after participants clicked on it and an auditory stimulus was played. Participants then clicked on the picture that matched the word they heard, and the trial ended as soon as they clicked on one of the pictures. There was no time limit and subjects were encouraged to take their time and to perform the task as naturally as possible. They typically responded in less than 2 s ($M = 1,295$ ms, S$D = 151$ ms).

Eye-tracking recording and analysis: We recorded eye movements at 250 Hz using an SR Research EyeLink II head-mounted eye-tracker. Both corneal reflection and pupil were used whenever possible. Participants were calibrated using the standard 9-point EyeLink procedure. The EyeLink II yields a real-time record of gaze in screen coordinates while compensating for head movements. This was automatically parsed into saccades and fixations using the default psychophysical parameters, and adjacent saccades and fixations were combined into a single "look" that began at the onset of the saccade and ended at the offset of the fixation (see also McMurray et al., 2010; McMurray, Tanenhaus, & Aslin, 2002).

Eye movements were recorded from the onset of the trial (the blue circle) to participants' response (mouse click). This variable time offset made it difficult to analyze results late in the time course. To address this issue, we adopted the approach of many prior studies (Allopenna et al., 1998; McMurray et al., 2002) by setting a fixed trial duration of 2,000 ms

(relative to stimulus onset). For trials that ended before this point, we extended the last eye movement; trials that were longer than 2,000 ms were truncated. This decision assumes that the last fixation reflects the word the participant "settled on," and particularly very late portions of the time course should be interpreted as an estimate of the final state of the system, not necessarily what the subject was fixating at a particular point in time.[2] The coordinates of each look were used to obtain information about which object was being fixated. For assigning fixations to objects, boundaries around the objects were extended by 100 pixels to account for noise and/or head drift in the eye-track record. However, this did not result in any overlap between objects; the neutral space between pictures was 124 pixels vertically and 380 pixels horizontally.

Results

Training—The accuracy of vocal responses in the primed production task was extracted offline using CheckVocal (Protopapas, 2007). For the rest of the tasks, accuracy was automatically output by DMDX. Across tasks and conditions, participants' accuracy was 94.1% (phoneme monitoring: $M = 97.1$; spoken to written word matching: $M = 92.5\%$; primed production: $M = 93.3\%$; same/different: $M = 95.0\%$). A series of independent-samples t tests assessed task performance as a function of condition in each task. Because our dependent variable was a proportion, raw accuracy data was transformed using the empirical logit transformation.

No significant differences were found between the two experimental conditions for the phoneme monitoring and spoken to written word matching tasks. However, participants in the HC condition performed marginally better than participants in the LC group on the same/different task ($M_{HC} = 97.3\%$, $M_{LC} = 92.5\%$), $t(72) = 1.92$, $p = .058$, and in the primed production task ($M_{HC} = 96.1\%$, $M_{LC} = 91.0\%$), $t(68) = 1.93$, $p = .057$. This difference was expected given the way these tasks were manipulated between the experimental conditions. For example, in the same/different task, participants reported whether the last sounds of the two words were the same or different; in the HC condition, this meant that they had to report whether they heard the same word twice or not (e.g., *net–net* vs. *net–neck*), whereas in the LC condition the words were always different and participants needed to break down the words into their sounds (e.g., *net–cat* vs. *net–park*). Similarly, in the primed production task, it is broadly accepted that lexical neighbors (like the prime) facilitate production (Vitevitch, 2002; but see Sadat, Martin, Costa, & Alario, 2014). Intriguingly, because in these cases the HC condition was apparently easier, any differences found in the test portion cannot be attributed to just the difficulty of the tasks (because HC training would be assumed to be harder in the other two).

Reaction times (RTs) were, overall, similar in the HC and the LC groups (phoneme monitoring: $M_{HC} = 447$ ms, $M_{LC} = 440$ ms; spoken to written word matching: $M_{HC} = 574$ ms, $M_{LC} = 524$ ms; primed production task: $M_{HC} = 672$ ms, $M_{LC} = 622$ ms; same/different:

---

[2]This is also why we used asymptotic functions like the logistic function to fit these data (as in McMurray et al., 2010) rather than polynomial approaches. Because at the asymptote the function does not change further, estimates of the asymptote can be seen as estimates of the final degree of settling, regardless of the actual time.

M_{HC} = 380 ms, M_{LC} = 550 ms). RTs did not differ significantly between groups, except in the same/different task, $t(72)$ = 3.6, $p < .01$.

Testing—Both motor responses (i.e. mouse clicks) and eye-movements during the testing portion were analyzed.

<u>Accuracy and RTs:</u> At test, participants were highly accurate in selecting the correct picture (99.2%, SD = 0.017%). Similar to training, we transformed the raw accuracy using the empirical logit transformation. We conducted a two-way, repeated-measures analysis of variance (ANOVA) to examine accuracy as a function of splicing condition (matching-splice, word-splice, nonword splice) and training group (HC, LC) on just the experimental trials. Accuracy differed significantly between splicing conditions, $F(2, 144)$ = 6.501, η^2 = .083, $p < .01$, but not between experimental groups ($F < 1$). The Splicing Condition × Group interaction was not significant ($F < 1$). Planned comparisons showed that participants were slightly, but significantly, more accurate in the matching-splice (99.35%) than in the nonword-splice (99.03%), $F(1, 72)$ = 11.39, η^2 = .137, $p < .01$, and the word-splice (99.18%), $F(1, 72)$ = 4.34, η^2 = .057, $p < .05$, conditions.

A similar ANOVA on RT showed that RTs differed significantly between splicing conditions, $F(1.8, 144)$ = 79.7, η^2 = .525, $p < .001$, but not between experimental groups, $F(1, 72)$ = 1.02, η^2 = .014, $p = .316$, and the Splicing × Group interaction was not significant, $F(1.8, 144)$ = 1.46, η^2 = .020, $p = .237$. Planned comparisons revealed that participants were significantly faster in the matching-splice (M = 1,234 ms) than in the nonword-splice (M = 1,322 ms), $F(1, 72)$ = 80.6, η^2 = .528, $p < .001$, and the word-splice (M = 1,330 ms), $F(1, 72)$ = 164.9, η^2 = .696, $p < .001$, conditions.

<u>Analysis of fixations: Documenting lexical interference:</u> We began our analysis of the fixations by collapsing across experimental groups to determine if there was evidence for interference as a whole (replicating the analyses of Dahan et al., 2001). To do this, we computed the proportion of trials on which participants were fixating the target for each of the three splicing conditions as a function of time (see Figure 5). In line with previous findings (Dahan et al., 2001, Kapnoula et al., 2015), we see a clear effect of splicing condition, with matching-splice showing the highest proportion of looks, followed by the nonword-splice, and word-splice conditions.

To evaluate this statistically, we computed the average proportion of fixations to the target between 600 ms and 1,600 ms poststimulus onset[3] and then transformed it using the empirical logit transformation. This was compared across splicing conditions using a series of linear mixed-effects models implemented in R and using the lme4 (Version 1.1–6; Bates, Maechler, & Dai, 2009) and lmerTest (Version 2.0 – 6; Kuznetsova, Brockhoff, & Christensen, 2013) R packages (R Development Core Team, 2011). Our three splice

[3]We chose this time window (600–1,600 ms) to be consistent with previous studies from our lab using the subphonemic mismatch paradigm (Kapnoula et al., 2015). The 600-ms onset was chosen because the stem duration (i.e., presplice sequence) lasts approximately 400 ms (plus the 200 ms needed to plan an eye movement). The 1,600-ms onset was chosen based on the range of participants' reaction times (about 1,000–1,600 ms); the broader time window ensured we captured differences in both fast and slow participants.

conditions were coded using two contrast codes (a) matching-versus nonword-splice (−.5/+.5) and (b) nonword-versus word-splice (−.5/+.5). Before focusing on the fixed effects, we evaluated models with various random-effect structures and found that the most complex model supported by the data was the one with random intercepts for subjects and items, and random slopes of the two contrast codes for items.

The difference between the matching-splice and the nonword-splice conditions was significant, $F(1, 26.8) = 18.9$, $p < .001$, suggesting that participants were sensitive to the subphonemic mismatch. More importantly, the difference between the word- and the nonword-splice conditions was also significant, $F(1, 26.7) = 19.1$, $p < .001$. These results showed a clear effect of splicing manipulation, and thus provided evidence of interference between the target and the competitor over and above any effect due to the subphonemic mismatch.

<u>Analysis of fixations: Between-group differences:</u> We next assessed our primary question: Whether training group influenced these interference effects. A comparison of Figure 5a (HC group) and Figure 5b (LC group) suggests that there may be a small difference in the effect of splicing between training groups. It also seems that differences are larger at specific times, suggesting that an analytic approach based on area under the curve may not detect such differences. Thus, to evaluate the effect of training group more precisely, we computed the effect of lexical interference over time by subtracting the fixations to the target in the word-splice condition from those in the matching-splice condition. Figure 6 displays this effect as a function of time and training group. It suggests a larger effect in the LC group than in the HC group, particularly at the end of the time course.

To evaluate this statistically, we examined the time course of the interference effect using a nonlinear curve-fitting approach (Farris-Trimble & McMurray, 2013; McMurray et al., 2010). By fitting nonlinear functions to the data as a function of time, we examined how lexical interference changes in time, thus avoiding the limitations set by restricting our analyses to a specific time window. Moreover, by using a curve that has a defined asymptote, we could estimate the "final" level of interference across variation in the RT of a trial, addressing whether each group fully resolved the interference (e.g., the interference effect asymptoted at 0). To do so, we fit an asymmetric Gaussian used by McMurray et al. (2010, Equation 1 below), the time course of the interference effect (e.g., Figure 6):

$$p\left(competitor\right) = \begin{cases} exp\left(\frac{(t-\mu)^2}{-2\sigma_1^2}\right)(p-b_1)+b_1 & if\, t \leq \mu \\ exp\left(\frac{(t-\mu)^2}{-2\sigma_1^2}\right)(p-b_2)+b_2 & if\, t > \mu \end{cases} \quad (1)$$

This function models the rise and fall of a function as two Gaussians (respectively) with separate lower asymptote and slopes, but constrained to have the same peak location (μ,) and height (p). The other four parameters, specified independently for each Gaussian, are: the onset baseline (b_1), the onset slope (u_1), the offset slope (u_2), and the offset baseline (b_2). Asymmetric Gaussians were fit to the interference and coarticulatory mismatch effect curves of each subject and each item separately using a constrained gradient descent method that

minimized the least-squares error and constrained the baseline parameters to be nonzero, and the peak to be higher than the baselines. The fitted parameters were then compared between groups to precisely characterize differences between them in the pattern of interference over time

Because the interference effects of individual subjects did not always conform to this function, we jackknifed the data before curve fitting. In this procedure, the raw time course functions are averaged for every subject except one. The function is then fit to this data and the parameters are extracted. This is repeated, excluding each subject in turn to yield a dataset of similar size to the original. The resulting parameter estimates are then compared using modified statistics that account for the much lower variance between subjects (cf., Miller, Patterson, & Ulrich, 1998; $_{see}$ Apfelbaum et al., 2011; McMurray et al., 2008, $_{for}$ examples in the VWP). Fits were good, with an average R^2 of 0.99 both for subjects and items.

We found that the offset baseline (b_2) differed significantly between groups, $t_{Jackknifed1}(72) = 2.34$, $p < .05$, $t_{Jackknifed2}(27) = 2.24$, $p < .05$, revealing a significantly lower asymptote (−.02) for the HC than the LC (.034) group. No other parameter differed significantly (see Table 3). Moreover, two one-sample t tests found that the offset baseline differed significantly from 0 for the LC, $t_{Jackknifed1}(72) = 3.16$, $p < .01$, $t_{Jackknifed2}(27) = 2.85$, $p < .01$, but not the HC group ($t_{Jackknifed11} < 1$, $t_{Jackknifed2} < 1$). This suggests that only the latter fully recovered from the competitor's inhibition effect on the target.

One alternative interpretation of our results could be that the HC tasks did not change the interlexical inhibition pattern, but, rather, they altered the amount of attention listeners paid to the fine-grained details of the auditory signal (which could help them resolve the competition more efficiently without requiring them to alter inhibition). To address this possibility, we compared the coarticulatory mismatch effect (matching-minus nonword-splice conditions) between the two groups following the same jackknifing and curve-fitting approach as in the main analyses. Because this effect does not reflect interlexical inhibition, we did not expect groups to differ. As shown in Table 4, none of the parameters differed significantly between groups, suggesting that both HC and LC participants were affected by the mismatching splice to a similar degree. This provides converging evidence that the crucial difference between the two groups has to do with the way in which they deal with the temporary boost of the competitor activation (i.e., word-splice).

A second alternative explanation is that participants in the HC group adopted a wait-and-see strategy in response to the competition demands of the HC training. That is, HC listeners might not have activated any lexical items until they reached the end of the word. Such an approach would minimize the competition between words, because participants would activate the target word at the end of the utterance (when it was unambiguous), without considering the competitor at all. To examine this possibility, we focused our analyses on the looks to the target in the matching-splice condition. If the HC participants adopted a wait-and-see strategy, then we would expect to see a slower pattern of looks to the target. To test this, we fitted a four parameter logistic (with parameters for minimum and maximum asymptote, slope, and crossover) to the looks-to-the-target curves (for the matching-splice

condition). This was done separately for each subject (and each item), using the logistic function used by McMurray et al. (2010). Because these functions were much more regularly shaped for individual subjects, the data were not jackknifed. Fits were good, with an average R^2 of 0.99 for both subjects and items. After fitting each participant's and each item's curve using this function, we compared the two relevant parameters (crossover and slope) between the two groups. No significant differences were found between the two groups in the crossover, $t_1 < 1$, $t_2(27) = 1.712$, $p = .098$, or the slope ($t_1 < 1$, $t_2 < 1$), suggesting very similar timing of target fixations between the two groups (see Figure 7a). Similarly, the offset baseline did not differ significantly, $t_1 < 1$, $t_2(27) = 1.72$, $p = .096$, suggesting similar degrees of fixation/activation. Overall, the difference between the groups seems to have been driven from the increased looks to the target in the word-splice condition for the HC compared with the LC participants (see Figure 7a).

Discussion

Our results suggested that, across both groups, the competitor interfered with target word recognition. However, at the same time, participants in the HC group were able to recover fully from this lexical competition while the LC group was not. Further, training did not appear to affect the general speed of processing, or sensitivity to coarticulatory mismatch, suggesting its locus is largely in specifically interference processes.

The timing of the effect, however, may seem counterintuitive, because the group effect was only observed quite late in the time course of processing. If the critical change (as a result of training) is at the level of lexical inhibition between target and competitor, then why did we not observe a between-groups difference earlier, when the coarticulatory mismatch was first encountered? One possibility is that the late timing of the effect suggests the training effect may lie not in the overall magnitude of interference, but in the way in which listeners *resolve* interference. It is unclear whether such a change specifically in the resolution of competition can derive from basic changes in things like inhibition or if it derives from change in a secondary mechanism within the lexical system or even a decision-stage process. To examine these possibilities, we conducted analyses with the TRACE model (McClelland & Elman, 1986) to determine whether a change in lexical inhibition alone could produce specifically the late differences we observed between groups.

TRACE Simulations

We modeled our results using a model that explicitly incorporates lexical inhibition, the TRACE model (McClelland & Elman, 1986). We conducted a series of TRACE simulations focusing on the parameter that directly controls the strength of interlexical inhibition. Our goal was to test whether changes in interlexical inhibition would give rise to late effects in lexical competition resolution, such as those described here, or if they would only affect early portions of processing. If so, this raises the possibility that these late effects can emerge within the lexical architecture; if not, this may require a separate (possibly postlexical) mechanism for resolving interference.

Method

We used the jTRACE implementation of TRACE (Strauss, Harris, & Magnuson, 2007) to model the activation of one item in our set (*cat*) in both the matching- and word-splice conditions.

Modeling splicing and measuring real-time inhibition—To model our splicing manipulation, we took advantage of the splicing option provided by jTRACE. Specifically, to model our matching-splice condition, we presented the first phoneme (/k/) from cycles 1-12, the second phoneme (the vowel, /ae/) from cycles 7-18, and the third phoneme (/t/) from cycles 13-24. All phonemes had a linear six-cycle ramp-up/ramp-down as is standard in TRACE. For the word-splice condition, everything was the same, except for the first five cycles of the target offset (cycles 13-17; see Figure 8), which now carried mismatching coarticulatory information about the offset consonant. As in the behavioral analyses, we calculated the inhibition effect by subtracting the activation of the target word in the word-splice from that in the matching-splice in each cycle.

Parameterization—Most parameters were kept at their default values. The only exceptions were (a) the inhibition between phonemes, which was set to a lower value of 0.01 (default value: 0.04) so to allow the model to recover from the splicing manipulation (cf., McMurray, Tanenhaus, & Aslin, 2009, for evidence that in general this parameter should be very low); and (b) acute and burst feature spread, which were raised to 7 (default value: 6) to better model our splicing manipulation. Lastly, lateral lexical inhibition, being the focus of our design, was systematically manipulated from the default value of .03-.06 (in steps of .005) to model different degrees of interlexical inhibition.

Rationale—We examined the effects of manipulating lexical inhibition in two steps. First, we directly examined the inhibition effect (i.e., target activation for matching-minus word-splice) over time as a function of the degree of lexical inhibition. Second, because we found that the critical difference between experimental groups lies in the late activation of the target in the word-splice condition (see Figure 7b), we examined the target activation in the matching- and the word-splice conditions as a function of lexical inhibition.

Results

As shown in Figure 9, the magnitude and timing of the peak of the interference effect differed across different levels of lexical inhibition. In contrast to what one might expect, little effect was observed during the coarticulatory mismatch, whereas an effect on the peak of the inhibition effect (both in terms of timing and strength) was observed approximately between the 40th and 60th cycle. This suggests that these effects may take time to build (and/or resolve). However, the crucial effect was observed much later, after the 60th cycle, where higher lexical inhibition (darker color) led to faster competition resolution than low inhibition (lighter color). While the early effect was not observed in our data, we note that there were several values of heightened inhibition where there were small effects on the peak, but much larger effects on the duration of the effect. Given uncertainty in mapping TRACE activation directly to eye movements, the more important finding thus appears to be that changes in inhibition can in fact prolong (or shorten) the period of interference.

Next we looked at the effect of lexical inhibition on the activation of the target when it was spliced with the competitor word (word-splice) as a function of group (which we modeled as differences in inhibition).

As shown in Figure 10, greater lexical inhibition leads to slightly lower activation of the target word early on, starting right after the onset of the coarticulatory mismatch (13th cycle) and lasting until a little after the end of the word (24th cycle). However, we see a much larger effect in the *opposite* direction much later (starting around the 40th cycle), where greater lexical inhibition leads to higher activation for the target. This pattern of results closely resembles the late effect reported above, according to which LC participants showed lower activation of the target compared with HC participants when it was spliced with a different word. Most importantly, this late effect was most prominent in the word-splice (panel b) rather than the matching-splice condition (panel a), as we also observed behaviorally (Figure 7a for matching-splice, Figure 7b for word-splice).

Discussion

Our simulations asked whether we can simulate the pattern of differences between the two experimental groups (HC vs. LC) by manipulating the lexical inhibition parameter in TRACE. The simulations confirmed that different degrees of lexical inhibition lead to different patterns of competition resolution, closely matching the reported behavioral differences between our two groups of participants: higher levels of lexical inhibition leads to more efficient activation of the target word in the word-splice condition. Even though the simulations did not fully capture the behavioral data (see the differences in peak of interference effect between Figures 6 and 9), they do show that differences in inhibition strength can have a significant effect on the resolution of the competition (see offset baseline in Figures 6 and 9), shown in both the measures of the interference effect and in the raw activation to the target. Crucially, however, what clearly stands out from our simulation results is that, as in our behavioral results, the most significant difference was observed *late* in processing, much after the offset of the target word.

One caveat to this modeling is that the object of our manipulation was a global inhibition parameter that affected inhibition between all the words, not just inhibition between the pairs of competitors. As a result, we cannot rule out that some of our modeling results may have also been affected by changes in inhibition between nontargeted words (e.g., in the inhibition between our target *cat* and words like *cab* and *hat*). However, given that our splicing manipulation boosts activation for the specific competitor, it seems likely that this word would exert the strongest effect. Moreover, because our behavioral results cannot pinpoint our effects to global or local changes, the demonstration that an increase in inhibition (of either kind) improves the recovery from interference is quite valuable.

These simulations provide useful insight into the underlying mechanism that gives rise to the pattern of behavioral results reported earlier (see ***Analysis of fixations: Between-group differences***). As seen in Figure 9 (which is the modeling equivalent of Figure 6), the primary effect of lexical inhibition is not observed at the time of the coarticulatory mismatch (cycles 13-17), but quite a bit later in processing (after the 40th cycle), and altering inhibition has a large effect on the duration of the interference effect. The timing of this effect suggests that

it most likely reflects differences in the degree to which the target inhibits the competitor. Therefore, lexical inhibition seems to have its greatest effect on the *resolution* of the competition between the target and the competitor. To examine this further, we plotted the activation of the *competitor* (e.g., *cap* when the target was *cat*) as a function of lexical inhibition (see Figure 11). Critically, the overall degree of competitor activity (the peak) did not appear to be affected by variation in inhibition; rather, the largest effects of inhibition appeared to be in the duration of its activity. This supports our interpretation of the empirical results suggesting that the primary effect of changing the strength of inhibition is in the (later) suppression of the competitor.

Also of note is that, to model the HC group (see Figures 5a and 6 in **Analysis of fixations: Between-group differences**), inhibition must be *increased* (rather than decreased). The fact that this appears to help fully suppress the competitor suggests that, when considering the role of higher or lower inhibition, we must consider both interference from the competitors and the ability of the target to suppress competitors to lead to accurate word recognition.

General Discussion

This study asked whether the way that listeners recover from lexical interference can change with training. Training was structured to create low versus high levels of interlexical competition, but with the same tasks and words in both conditions. The subsequent visual world task revealed that participants in the HC group were significantly better at recovering from the interference of a lexical competitor's than those in the LC group (see Figures 5–7). HC participants fully resolved the interlexical competition by the end of the trial (their inhibitory effect reached 0), whereas LC participants did not (see Figure 6). This suggests that our training modulated the degree of interference between words.

Three alternative explanations were ruled out. First was the possibility that the HC task variants were more difficult than the LC ones, prompting greater attention to the experiment in general. This seems unlikely because two of the HC tasks were marginally easier than the LC versions. Second, it was possible that the HC tasks did not alter inhibition at all, but rather prompted listeners to pay closer attention to the auditory signal, noticing the coarticulatory mismatch more. This should have resulted in a larger coarticulatory mismatch effect (the difference between the matching-splice and nonword-splice conditions), which was not observed. A third possibility was that subjects responded to the HC training by waiting until the word was fully heard to activate lexical entries (and therefore not activating the competing word). However, the time course of target fixations showed no difference between groups. This leaves changes in inhibitory processing as the most likely explanation.

An intriguing aspect of our findings is how late the effect appeared to be, compared with the presence of the coarticulatory mismatch (see Figure 6). It was not intuitively obvious that such a late effect was consistent with changes in lexical inhibition per se—one might have predicted more robust early effects from a modification of inhibition due to the early interference between the competitor and the target. However, our TRACE simulations showed that, when we manipulated lexical inhibition, there was only a small effect during the coarticulation, but robust late effects in processing (see Figures 9–11). Specifically, the

critical change seems to be on the degree to which the target word successfully suppressed the competitor (see Figure 11). The most likely explanation for this is that the interference effect we observed largely depended on how active the competing words were. Early in the time course of processing, no words were highly active and, consequently, there could not be strong inhibitory effects (see cycles 1–30 in Figures 9 and 10 in **TRACE Simulations**). However, later when both the competitor and the target had been partially activated, an increase in inhibition allowed the target to suppress the competitor more efficiently (see Figure 11 in **TRACE Simulations**). In other words, when both the competitor and the target had been partially activated, the measurable consequences of inhibition were much stronger. Consequently, a small change in the degree of inhibition had its largest effects late in processing, when the words were more active (see Figure 6 in **TRACE Simulations**). Thus, the simulation results not only showed that differences in lexical inhibition may affect largely the timing of interference resolution (e.g., the eventual target's ability to suppress competitors), but they also offered a mechanistic explanation of how this may take place.

An important question is whether the changes in lexical interference we observed with training are the result of global or local changes. That is, did HC training increase inhibition among all words, or just among pairs of words that were present during training? Given the high degree of overlap among words, this may be difficult to pinpoint—for example, training on *cat/cap* may also affect *cab* and *hat* because these competitors would also be partially active during training. However, either way, this kind of experience-dependent plasticity may be important for enhancing interference resolution, either over development (e.g., as the lexicon grows) or in challenging listening circumstances.

An important remaining question is the nature of the learning that gives rise to this effect. We outlined three ways in which our manipulation could lead to learning. First, our manipulation may set up task demands that require the resolution of interference (or which minimize the need for it). These demands, persisting over a period of 40 min (training duration), could have prompted a change in the system. Second, the boost in target activation afforded by the response options (in the LC training) could have reduced the need for resolving interference, leading to changes via a similar mechanism to the first. Finally, if inhibitory connections are modified by Hebbian type processes, coactivation of the words could lead to changes in their inhibitory links. These may all conspire as well.

The crucial difference between the first two and the third account is that the former two imply a global-level change in the system, which is focused on interference resolution. Such a change would have an effect on the interference resolution across all competing items (i.e., both trained and untrained word pairs). As we described in *Analysis of fixations: Between-group differences*, in that case the training effect could derive from decision-stage processes or secondary processes intended for recovering from an incorrect interpretation. However, as our TRACE simulations showed, such an effect may also arise from simply turning up the degree of inhibition across all words, which seems like the more parsimonious explanation. In contrast, the Hebbian account, by definition, operates on a local level (i.e., between specific words of pairs) because it relies on coactivation between specific words to increase their inhibitory connection. While we cannot make conclusive claims about which mechanism(s) is responsible (and they are not mutually exclusive), two lines of converging

evidence support a Hebbian account. First, the modeling results showed that the observed advantage for the HC group is due to their having *higher* lexical inhibition. This allows HC listeners to suppress the competitor more efficiently as soon as they have accumulated enough information for the target. Because HC training was developed to encourage coactivation of target and competitor, the fact that stronger inhibition was needed is consistent with a Hebbian account. Second, research on interference deriving from newly learned words has generally shown that when interference is measured with relatively coarse measures (e.g., lexical decision), newly learned words do not interfere with known word recognition unless there is a night of sleep (Dumay & Gaskell, 2007). However, there are two notable exceptions to this (Fernandes et al., 2009; Lindsay & Gaskell, 2013), both featuring training in which coactivation of multiple words would be enhanced. Lindsay and Gaskell (2013) used a training paradigm in which familiar and novel words were interleaved; and Fernandes et al. (2009) used an unsegmented stream of speech in which all of the to-be-learned words were closely interleaved. Taken together, this might suggest that coactivation of words leads to stronger inhibitory links between them via simple Hebbian learning, and that this is perhaps the same mechanism used to establish those links in the first place (see McMurray et al., in press, for a more in depth discussion).

As a whole, these results are significant for four reasons. First, language learning is often discussed as a process of establishing input/output relations or mappings. Here we showed that learning can also affect real-time processes by which representations interact with each other. This finding may have significant implications for the way in which we conceptualize developmental changes in general; if specific aspects of the real-time processes that underlie spoken word recognition are amenable to change, then developmental changes in processing could be viewed as the long-term outcome of these sorts of short-term changes in the system. This could offer a clear mechanism for developmental changes in the efficiency of familiar word recognition (Fernald et al., 1998; Marchman & Fernald, 2008; Rigler et al., 2015). It may also offer a mechanism by which developmental adversity could lead to individual differences and, perhaps, maladaptive outcomes. For example, in visual word recognition, Protopapas and Kapnoula (2015) recently examined how the effect of word length is modulated by both short-term (the presentation of a word for the second time) and long-term experiences (education). The presence of both short- and long-term modulations of the well-established effect of length is consistent with the idea that short-term learning is likely one of the driving forces behind individual differences in online processing.

This is not to say that changes in inhibition are the only mechanism by which real-time processing can change over development. For example, changes in the quality of the bottom-up mappings between input representations and words may also alter speed of processing (McMurray et al., 2012). Changes in the use of top-down knowledge or attentional processes may also improve processing speed. The bottom line, however, is that our study provides strong evidence of a clear form of experience-dependent plasticity that may be essential for developing real-time processing abilities.

Second, this finding is relevant to a wide variety of domains in which interference has been invoked, such as decision making (Ratcliff & McKoon, 2008; Usher & McClelland, 2001), attentional selection (Mordkoff & Yantis, 1993), visual search (Duncan & Humphreys, 1989;

Spivey, 2007), music perception (Bharucha, 1987), speech production (Dell, 1986), sentence processing (MacDonald, Pearlmutter, & Seidenberg, 1994), and word recognition (Dahan et al., 2001; Norris, McQueen, & Cutler, 1995; Vitevitch & Luce, 1998). Therefore, it will be important to determine if a similar form of plasticity can be observed in these other domains. In this regard, it is important to distinguish between lateral inhibition (between similar representations) and the top-down inhibition that is considered a core aspect of executive function. In fact, Burke and Shafto (2008) referred to the former as "automatic" and "obligatory" inhibition (e.g., McClelland & Rumelhart, 1981), to distinguish it from an attention-based form of inhibition (e.g., inhibition deficit theory; Hasher & Zacks, 1988). It is likely that training in attention-based forms of inhibition may also be relevant for language and may also be amenable to training (Novick, Hussey, Teubner-Rhodes, Harbison, & Bunting, 2014). However, our work is, to our knowledge, the first to show that this automatic inhibition is amenable to training.

Third, models of spoken word recognition do not usually discuss whether lexical inhibition is a fixed or a malleable aspect of processing (Gaskell & Marslen-Wilson, 1999; McClelland & Elman, 1986; Mirman et al., 2006; Norris & McQueen, 2008). Our findings clearly support the latter; interlexical inhibitory links are susceptible to change via experience, and are therefore likely to differ among individuals, or even among words within an individual. Therefore, this needs to be taken into account as a crucial feature of spoken word recognition and incorporated perhaps explicitly into models of word learning (e.g., Gaskell & Marslen-Wilson, 1997; McMurray et al., 2012; McMurray et al., in press).

Finally, this finding may be applicable to understanding and treating impairments related to disrupted lateral inhibition like specific language impairment (McMurray et al., 2010), dyslexia (Ziegler & Muneaux, 2007), or cognitive deficits arising from aging (Sommers, 1996; Sommers & Danielson, 1999). By demonstrating that inhibition can change with experience, it raises the possibility that adverse experiences can give rise to maladaptive patterns of lexical inhibition. Moreover, although our training paradigm was not designed to be directly applicable as a treatment regime, our strong evidence of experience-dependent plasticity suggests that it may be possible to design treatments that train individuals to cope with competition better. This may be particularly amenable to specific language impairment, given that their increased competition primarily appears at the end of processing (McMurray et al., 2010; Dollaghan, 1998)—exactly where we find the most robust effects of HC training.

What goal may this plasticity serve? We believe that there are two, not necessarily mutually exclusive, accounts why the system is susceptible to experience-based changes. The first is that this plasticity may not be necessary in adults, but may instead be the residue of development. As children's vocabularies expand, this plasticity may allow for similar sounding words (like *cat* and *cut*) to inhibit each other because they are added to the lexicon, tuning the degree and nature of the inhibition to the particular structure of a child's lexicon at a given point in development. A flexible, plastic system for establishing inhibitory links would enable the strength of these inhibitory processes to be adjusted, depending on (a) the degree of similarity between words and (b) the number of similar words in the lexicon. As the network of lexical representations (i.e., the system consisting of the representations and

the relations between them) is gradually stabilized, this plasticity is no longer necessary, but it remains as an internal feature of the system. Alternatively, the adaptive nature of inhibition may play a helpful role throughout the life span in strengthening inhibition in cases where it is most needed. For example, in contexts where pairs of similar sounding words are likely to appear, or for particular pairs of similar sounding words that co-occur frequently, inhibition can be strengthened to support perception. Similarly, increased inhibition may also help in adverse listening conditions (i.e., background noise, accented speech, or fast speech). Therefore, this property of inhibitory plasticity may allow the system to be constantly updated and adjusted according to the needs of the most commonly encountered listening situations.

More broadly, these results illustrate plasticity not only in the way information is represented in the language system, but also in the pathways that support various forms of real-time processing. This finding reveals, for the first time, that the way in which inhibitory processes unfold in time is susceptible to experience. This has important implications for the way we conceptualize lateral inhibition between representations across a variety of cognitive processes.

Acknowledgments

Appendix A Appendix A

Word Pairs Used in Training

Target word	Competitor
bait	bake
bat	back
bride	bribe
bug	bud
carp	cart
cat	cap
chick	chip
dart	dark
dot	dock
fork	fort
grad	grab
heap	heat
hub	hug
job	jog
knot	knock
leap	leak
mug	mud

Target word	Competitor
net	neck
part	park
pick	pit
pope	poke
rod	rob
shake	shape
steak	state
suit	soup
tarp	tart
web	wed
zip	zit

Appendix B   Appendix B

List of Images Used in the Visual World Paradigm Task

Target word	Cohort	Unrelated 1	Unrelated 2
bait	boot	jug	wet
bat	boat	street	drug
bride	bread	feet	yacht
bug	bark	dead	gap
cart	kid	snake	lid
cat	cord	blood	beard
chick	chart	hook	pig
dart	dog	ride	feed
dock	date	step	bulb
fork	fog	side	god
grad	gripe	stork	drop
heat	hood	maid	yard
hub	head	wreck	crib
job	jet	book	duck
knot	knight	rag	bead
leak	lark	peg	wig
mug	mit	spark	truck
net	nut	red	goat
part	pad	black	trout
pit	plug	luck	sweat
pope	plate	cube	dad
rod	root	bet	vote
shake	shed	choke	keg
steak	stick	check	milk
suit	sword	reed	flake

Target word	Cohort	Unrelated 1	Unrelated 2
tarp	toad	jeep	vet
web	wood	cook	shout
zip	zap	cloud	raid

Appendix C   Appendix C

Triplets (in International Phonetic Alphabet) Used in Testing to Construct the Spliced Stimuli

Matching-splice (ne_tt condition)	Word-splice ($ne_{ck}t$ condition)	Nonword-splice (ne_pt condition)
beɪt (bait)	beɪk (bake)	beɪp
bæt (bat)	bæk (back)	bæp
brɑɪd (bride)	brɑɪb (bribe)	brɑɪg
bʌg (bug)	bʌd (bud)	bʌb
kɑrp (carp)	kɑrp (cart)	kɑrp
kæt (cat)	kæp (cap)	kæk
tʃɪk (chick)	tʃɪp (chip)	tʃɪt
dært (dart)	dærk (dark)	dærp
dæt (dot)	dæk (dock)	dæp
fɔrk (fork)	fɔrt (fort)	fɔrp
græd (grad)	græb (grab)	græg
hip (heap)	hit (heat)	hik
hʌb (hub)	hʌg (hug)	hʌd
dʒæb (job)	dʒæg (jog)	dʒæd
næt (knot)	næk (knock)	næp
lip (leap)	lik (leak)	lit
mʌg (mug)	mʌd (mud)	mʌb
nɛt (net)	nɛk (neck)	nɛp
pært (part)	pærk (park)	pærp
pɪk (pick)	pɪt (pit)	pɪp
poʊp (pope)	poʊk (poke)	poʊt
ræd (rod)	ræb (rob)	ræg
ʃeɪk (shake)	ʃeɪp (shape)	ʃeɪt
steɪk (steak)	steɪt (state)	steɪp
sut (suit)	sup (soup)	suk
tærp (tarp)	tært (tart)	tærk
wɛb (web)	wɛd (wed)	wɛg
zɪp (zip)	zɪt (zit)	zɪk

References

Allopenna PD, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. Journal of Memory and Language. 1998; 38:419–439. http://dx.doi.org/10.1006/jmla.1997.2558.

Apfelbaum KS, Blumstein SE, McMurray B. Semantic priming is affected by real-time phonological competition: Evidence for continuous cascading systems. Psychonomic Bulletin & Review. 2011; 18:141–149. http://dx.doi.org/10.3758/s13423-010-0039-8. [PubMed: 21327343]

Bates D, Maechler M, Dai B. Lme4: Linear mixed-effects models using s4 classes. 2009 Retrieved from https://cran.r-project.org/web/packages/lme4/index.html.

Bharucha J. Music cognition and perceptual facilitation: A connectionist framework. Music Perception. 1987; 5:1–30. http://dx.doi.org/10.2307/40285384.

Borovsky A, Elman JL, Fernald A. Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. Journal of Experimental Child Psychology. 2012; 112:417–436. http://dx.doi.org/10.1016/j.jecp.2012.01.005. [PubMed: 22632758]

Burke, D.; Shafto, M. Language and aging. In: Craik, FIM.; Salthouse, TA., editors. The handbook of aging and cognition. Psychology Press; New York, NY: 2008. p. 373-443.

Coutanche MN, Thompson-Schill SL. Fast mapping rapidly integrates information into existing memory networks. Journal of Experimental Psychology: General. 2014; 143:2296–2303. http://dx.doi.org/10.1037/xge0000020. [PubMed: 25222265]

Cutler, A.; Norris, D. Monitoring sentence comprehension. In: Cooper, WE.; Walker, ECT., editors. Sentence processing: Psycholinguistic studies presented to Merrill Garrett. Erlbaum; Hillsdale, NJ: 1979. p. 113-134.

Dahan D, Magnuson JS, Tanenhaus MK, Hogan EM. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. Language and Cognitive Processes. 2001; 16:507–534. http://dx.doi.org/10.1080/01690960143000074.

Davis MH, Di Betta AM, Macdonald MJE, Gaskell MG. Learning and consolidation of novel spoken words. Journal of Cognitive Neuroscience. 2009; 21:803–820. http://dx.doi.org/10.1162/jocn.2009.21059. [PubMed: 18578598]

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychological Review. 1986; 93:283–321. http://dx.doi.org/10.1037/0033-295X.93.3.283. [PubMed: 3749399]

Desroches AS, Joanisse MF, Robertson EK. Specific phonological impairments in dyslexia revealed by eyetacking. Cognition. 2006; 100:B32–B42. http://dx.doi.org/10.1016/j.cognition.2005.09.001. [PubMed: 16288732]

Dollaghan C. Spoken word recognition in children with and without specific language impairment. Applied Psycholinguistics. 1998; 19:193–207. http://dx.doi.org/10.1017/S0142716400010031.

Dumay N, Gaskell MG. Sleep-associated changes in the mental representation of spoken words. Psychological Science. 2007; 18:35–39. http://dx.doi.org/10.1111/j.1467-9280.2007.01845.x. [PubMed: 17362375]

Duncan J, Humphreys GW. Visual search and stimulus similarity. Psychological Review. 1989; 96:433–458. http://dx.doi.org/10.1037/0033-295X.96.3.433. [PubMed: 2756067]

Farris-Trimble A, McMurray B. Test-retest reliability of eye tracking in the visual world paradigm for the study of real-time spoken word recognition. Journal of Speech, Language, and Hearing Research. 2013; 56:1328–1345. http://dx.doi.org/10.1044/1092-4388(2012/12-0145).

Farris-Trimble A, McMurray B, Cigrand N, Tomblin JB. The process of spoken word recognition in the face of signal degradation. Journal of Experimental Psychology: Human Perception and Performance. 2014; 40:308–327. http://dx.doi.org/10.1037/a0034353. [PubMed: 24041330]

Fernald A, Perfors A, Marchman VA. Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. Developmental Psychology. 2006; 42:98–116. http://dx.doi.org/10.1037/0012-1649.42.1.98. [PubMed: 16420121]

Fernald A, Pinto JP, Swingley D, Weinberg A, McRoberts GW. Rapid gains in speed of verbal processing by infants in the 2nd year. Psychological Science. 1998; 9:228–231. http://dx.doi.org/10.1111/1467-9280.00044.

Fernandes T, Kolinsky R, Ventura P. The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. Cognition. 2009; 112:349–366. http://dx.doi.org/10.1016/j.cognition.2009.05.002. [PubMed: 19545864]

Forster KI, Forster JC. DMDX: A windows display program with millisecond accuracy. Behavior Research Methods, Instruments, & Computers. 2003; 35:116–124.

Galle ME, McMurray B. The waiting is the hardest part: How asynchronous acoustic cues are integrated for fricative voicing and place of articulation. The Journal of the Acoustical Society of America. 2012; 132:1966. http://dx.doi.org/10.1121/1.4755244.

Garlock V, Walley A, Metsala J. Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. Journal of Memory and Language. 2001; 45:468–492. http://dx.doi.org/10.1006/jmla.2000.2784.

Gaskell MG, Dumay N. Lexical competition and the acquisition of novel words. Cognition. 2003; 89:105–132. http://dx.doi.org/10.1016/S0010-0277(03)00070-2. [PubMed: 12915296]

Gaskell MG, Marslen-Wilson WD. Integrating form and meaning: A distributed model of speech perception. Language and Cognitive Processes. 1997; 12:613–656. http://dx.doi.org/10.1080/016909697386646.

Gaskell MG, Marslen-Wilson WD. Ambiguity, competition, and blending in spoken word recognition. Cognitive Science. 1999; 23:439–462. http://dx.doi.org/10.1207/s15516709cog2304_3.

Gow DW Jr. Gordon PC. Lexical and prelexical influences on word segmentation: Evidence from priming. Journal of Experimental Psychology: Human Perception and Performance. 1995; 21:344–359. http://dx.doi.org/10.1037/0096-1523.21.2.344. [PubMed: 7714476]

Gupta P, Tisdale J. Word learning, phonological short-term memory, phonotactic probability and long-term memory: Towards an integrated framework. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences. 2009; 364:3755–3771. http://dx.doi.org/10.1098/rstb.2009.0132. [PubMed: 19933144]

Hasher L, Zacks RT. Working memory, comprehension, and aging: A review and a new view. Psychology of Learning and Motivation. 1988; 22:193–225. http://dx.doi.org/10.1016/S0079-7421(08)60041-9.

Huettig F, Brouwer S. Delayed anticipatory spoken language processing in adults with dyslexia—Evidence from eye-tracking. Dyslexia: An International Journal of Research and Practice. 2015; 21:97–122. http://dx.doi.org/10.1002/dys.1497.

Kapnoula EC, McMurray B. Newly learned word-forms are abstract and integrated immediately after acquisition. Psychological Bulletin & Review. 2015 Advance online publication. http://dx.doi.org/10.3758/s13423-015-0897-1.

Kapnoula EC, Packard S, Gupta P, McMurray B. Immediate lexical integration of novel word forms. Cognition. 2015; 134:85–99. http://dx.doi.org/10.1016/j.cognition.2014.09.007. [PubMed: 25460382]

Kuznetsova A, Brockhoff P, Christensen R. LmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package) (R package Version 2.0-6). 2013 Retrieved from http://CRAN.R-project.org/package=lmerTest.

Leach L, Samuel AG. Lexical configuration and lexical engagement: When adults learn new words. Cognitive Psychology. 2007; 55:306–353. http://dx.doi.org/10.1016/j.cogpsych.2007.01.001. [PubMed: 17367775]

Lindsay S, Gaskell MG. Lexical integration of novel words without sleep. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2013; 39:608–622. http://dx.doi.org/10.1037/a0029243.

Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear and Hearing. 1998; 19:1–36. http://dx.doi.org/10.1097/00003446-199802000-00001. [PubMed: 9504270]

MacDonald MC, Pearlmutter NJ, Seidenberg MS. The lexical nature of syntactic ambiguity resolution. Psychological Review. 1994; 101:676–703. http://dx.doi.org/10.1037/0033-295X.101.4.676. [PubMed: 7984711]

Magill R, Hall K. A review of the contextual interference effect in motor skill acquisition. Human Movement Science. 1990; 9:241–289. http://dx.doi.org/10.1016/0167-9457(90)90005-X.

Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. The time course of spoken word learning and recognition: Studies with artificial lexicons. Journal of Experimental Psychology: General. 2003; 132:202–227. http://dx.doi.org/10.1037/0096-3445.132.2.202. [PubMed: 12825637]

Mainela-Arnold E, Evans JL, Coady JA. Lexical representations in children with SLI: Evidence from a frequency-manipulated gating task. Journal of Speech, Language, and Hearing Research. 2008; 51:381–393. http://dx.doi.org/10.1044/1092-4388(2008/028).

Mani N, Plunkett K. Phonological priming and cohort effects in toddlers. Cognition. 2011; 121:196–206. http://dx.doi.org/10.1016/j.cognition.2011.06.013. [PubMed: 21831366]

Marchman VA, Fernald A. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. Developmental Science. 2008; 11:F9–F16. http://dx.doi.org/10.1111/j.1467-7687.2008.00671.x. [PubMed: 18466367]

Marslen-Wilson W, Warren P. Levels of perceptual representation and process in lexical access: Words, phonemes, and features. Psychological Review. 1994; 101:653–675. http://dx.doi.org/10.1037/0033-295X.101.4.653. [PubMed: 7984710]

Marslen-Wilson WD. Functional parallelism in spoken word-recognition. Cognition. 1987; 25:71–102. http://dx.doi.org/10.1016/0010-0277(87)90005-9. [PubMed: 3581730]

Mattys SL, Scharenborg O. Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors. Psychology and Aging. 2014; 29:150–162. http://dx.doi.org/10.1037/a0035387. [PubMed: 24660803]

McClelland JL, Elman JL. The TRACE model of speech perception. Cognitive Psychology. 1986; 18:1–86. http://dx.doi.org/10.1016/0010-0285(86)90015-0. [PubMed: 3753912]

McClelland JL, Rumelhart D. An interactive activation model of context effects in letter perception: I. An account of basic findings. Psychological Review. 1981; 88:375–407. http://dx.doi.org/10.1037/0033-295X.88.5.375.

McMurray B, Clayards MA, Tanenhaus MK, Aslin RN. Tracking the time course of phonetic cue integration during spoken word recognition. Psychonomic Bulletin & Review. 2008; 15:1064–1071. http://dx.doi.org/10.3758/PBR.15.6.1064. [PubMed: 19001568]

McMurray B, Horst JS, Samuelson LK. Word learning emerges from the interaction of online referent selection and slow associative learning. Psychological Review. 2012; 119:831–877. http://dx.doi.org/10.1037/a0029872. [PubMed: 23088341]

McMurray, B.; Kapnoula, EC.; Gaskell, MG. Learning and integration of new word-forms: Consolidation, pruning and the emergence of automaticity. In: Gaskell, MG.; Mirkovi , J., editors. Speech perception and spoken word recognition. Psychology Press; London, England: in press

McMurray B, Samelson VM, Lee SH, Tomblin JB. Individual differences in online spoken word recognition: Implications for SLI. Cognitive Psychology. 2010; 60:1–39. http://dx.doi.org/10.1016/j.cogpsych.2009.06.003. [PubMed: 19836014]

McMurray B, Tanenhaus MK, Aslin RN. Gradient effects of within-category phonetic variation on lexical access. Cognition. 2002; 86:B33–B42. http://dx.doi.org/10.1016/S0010-0277(02)00157-9. [PubMed: 12435537]

McMurray B, Tanenhaus MK, Aslin RN. Within-category VOT affects recovery from "lexical" garden paths: Evidence against phoneme-level inhibition. Journal of Memory and Language. 2009; 60:65–91. http://dx.doi.org/10.1016/j.jml.2008.07.002. [PubMed: 20046217]

Metsala, J.; Walley, A. Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In: Metsala, JL.; Ehri, LC., editors. Word recognition in beginning literacy. Erlbaum; Mahwah, NJ: 1998. p. 89-120.

Miller J, Patterson T, Ulrich R. Jackknife-based method for measuring LRP onset latency differences. Psychophysiology. 1998; 35:99–115. http://dx.doi.org/10.1111/1469-8986.3510099. [PubMed: 9499711]

Mirman D, McClelland JL, Holt LL. An interactive Hebbian account of lexically guided tuning of speech perception. Psychonomic Bulletin & Review. 2006; 13:958–965. http://dx.doi.org/10.3758/BF03213909. [PubMed: 17484419]

Mordkoff JT, Yantis S. Dividing attention between color and shape: Evidence of coactivation. Perception & Psychophysics. 1993; 53:357–366. http://dx.doi.org/10.3758/BF03206778. [PubMed: 8483699]

Norris D, McQueen JM. Shortlist B: A Bayesian model of continuous speech recognition. Psychological Review. 2008; 115:357–395. http://dx.doi.org/10.1037/0033-295X.115.2.357. [PubMed: 18426294]

Norris D, McQueen JM, Cutler A. Competition and segmentation in spoken-word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21:1209–1228. http://dx.doi.org/10.1037/0278-7393.21.5.1209.

Novick JM, Hussey E, Teubner-Rhodes S, Harbison JI, Bunting MF. Clearing the garden-path: Improving sentence processing through cognitive control training. Language, Cognition and Neuroscience. 2014; 29:186–217. http://dx.doi.org/10.1080/01690965.2012.758297.

Poellmann K, Bosker H, McQueen J, Mitterer H. Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch. Journal of Phonetics. 2014; 46:101–127. http://dx.doi.org/10.1016/j.wocn.2014.06.004.

Protopapas A. CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. Behavior Research Methods. 2007; 39:859–862. http://dx.doi.org/10.3758/BF03192979. [PubMed: 18183901]

Protopapas A, Kapnoula EC. Short-term and long-term effects on visual word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2015 Advance online publication. http://dx.doi.org/10.1037/xlm0000191.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2011. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/

Ramscar M, Hendrix P, Shaoul C, Milin P, Baayen H. The myth of cognitive decline: Non-linear dynamics of lifelong learning. Topics in Cognitive Science. 2014; 6:5–42. http://dx.doi.org/10.1111/tops.12078. [PubMed: 24421073]

Ratcliff R, McKoon G. The diffusion decision model: Theory and data for two-choice decision tasks. Neural Computation. 2008; 20:873–922. [PubMed: 18085991]

Revill KP, Spieler DH. The effect of lexical frequency on spoken word recognition in young and older listeners. Psychology and Aging. 2012; 27:80–87. http://dx.doi.org/10.1037/a0024113. [PubMed: 21707175]

Rigler H, Farris-Trimble A, Greiner L, Walker J, Tomblin JB, McMurray B. The slow developmental timecourse of real-time spoken word recognition. Developmental Psychology. 2015 Advance online publication. http://dx.doi.org/10.1037/dev0000044.

Sadat J, Martin CD, Costa A, Alario FX. Reconciling phonological neighborhood effects in speech production through single trial analysis. Cognitive Psychology. 2014; 68:33–58. http://dx.doi.org/10.1016/j.cogpsych.2013.10.001. [PubMed: 24291531]

Sekerina IA, Brooks PJ. Eye movements during spoken word recognition in Russian children. Journal of Experimental Child Psychology. 2007; 98:20–45. http://dx.doi.org/10.1016/j.jecp.2007.04.005. [PubMed: 17560596]

Shea J, Morgan R. Contextual interference effects on the acquisition, retention, and transfer of a motor skill. Journal of Experimental Psychology: Human Learning and Memory. 1979; 5:179–187. http://dx.doi.org/10.1037/0278-7393.5.2.179.

Sommers MS. The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. Psychology and Aging. 1996; 11:333–341. http://dx.doi.org/10.1037/0882-7974.11.2.333. [PubMed: 8795062]

Sommers MS, Danielson SM. Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. Psychology and Aging. 1999; 14:458–472. http://dx.doi.org/10.1037/0882-7974.14.3.458. [PubMed: 10509700]

Spivey, MJ. The continuity of mind. Oxford University Press; New York, NY: 2007.

Stager CL, Werker JF. Infants listen for more phonetic detail in speech perception than in word-learning tasks. Nature. 1997; 388:381–382. http://dx.doi.org/10.1038/41102. [PubMed: 9237755]

Storkel HL. Restructuring of similarity neighbourhoods in the developing mental lexicon. Journal of Child Language. 2002; 29:251–274. http://dx.doi.org/10.1017/S0305000902005032. [PubMed: 12109371]

Strauss TJ, Harris HD, Magnuson JS. jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. Behavior Research Methods. 2007; 39:19–30. http://dx.doi.org/10.3758/BF03192840. [PubMed: 17552468]

Streeter LA, Nigro GN. The role of medial consonant transitions in word perception. The Journal of the Acoustical Society of America. 1979; 65:1533–1541. http://dx.doi.org/10.1121/1.382917. [PubMed: 489823]

Swingley D, Aslin RN. Lexical competition in young children's word learning. Cognitive Psychology. 2007; 54:99–132. http://dx.doi.org/10.1016/j.cogpsych.2006.05.001. [PubMed: 17054932]

Swingley D, Pinto JP, Fernald A. Continuous processing in word recognition at 24 months. Cognition. 1999; 71:73–108. http://dx.doi.org/10.1016/S0010-0277(99)00021-9. [PubMed: 10444905]

Tamminen J, Gaskell MG. Newly learned spoken words show long-term lexical competition effects. Quarterly Journal of Experimental Psychology. 2008; 61:361–371. http://dx.doi.org/10.1080/17470210701634545.

Usher M, McClelland JL. The time course of perceptual choice: The leaky, competing accumulator model. Psychological Review. 2001; 108:550–592. http://dx.doi.org/10.1037/0033-295X.108.3.550. [PubMed: 11488378]

Vitevitch MS. The influence of phonological similarity neighborhoods on speech production. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28:735–747. http://dx.doi.org/10.1037/0278-7393.28.4.735.

Vitevitch MS, Luce PA. When words compete: Levels of processing in perception of spoken words. Psychological Science. 1998; 9:325–329. http://dx.doi.org/10.1111/1467-9280.00064.

Vitevitch MS, Luce PA, Pisoni DB, Auer ET. Phonotactics, neighborhood activation, and lexical access for spoken words. Brain and Language. 1999; 68:306–311. http://dx.doi.org/10.1006/brln.1999.2116. [PubMed: 10433774]

Walley A. The role of vocabulary development in children= s spoken word recognition and segmentation ability. Developmental Review. 1993; 13:286–350. http://dx.doi.org/10.1006/drev.1993.1015.

Walley A, Metsala J, Garlock V. Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. Reading and Writing. 2003; 16:5–20. http://dx.doi.org/10.1023/A: 1021789804977.

Werker J, Fennell C, Corcoran K, Stager C. Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. Infancy. 2002; 3:1–30. http://dx.doi.org/10.1207/S15327078IN0301_1.

Witteman MJ, Bardhan NP, Weber A, McQueen JM. Automaticity and stability of adaptation to a foreign-accented speaker. Language and Speech. 2015; 58:168–189. http://dx.doi.org/10.1177/0023830914528102. [PubMed: 26677641]

Witteman MJ, Weber A, McQueen JM. Rapid and long-lasting adaptation to foreign-accented speech. The Journal of the Acoustical Society of America. 2010; 128:2486. http://dx.doi.org/10.1121/1.3508921.

Wulf G, Shea CH. Principles derived from the study of simple skills do not generalize to complex skill learning. Psychonomic Bulletin & Review. 2002; 9:185–211. http://dx.doi.org/10.3758/BF03196276. [PubMed: 12120783]

Ziegler JC, Muneaux M. Orthographic facilitation and phonological inhibition in spoken word recognition: A developmental study. Psychonomic Bulletin & Review. 2007; 14:75–80. http://dx.doi.org/10.3758/BF03194031. [PubMed: 17546734]
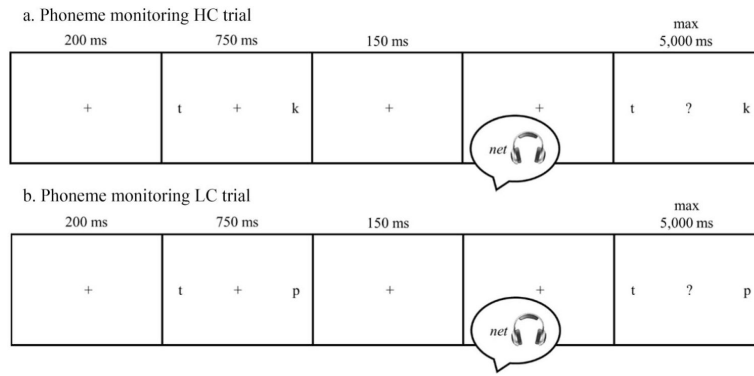
Figure 1.
Examples of trials for the phoneme monitoring task for the high-competition (HC) group (a) and low-competition (LC) group (b).
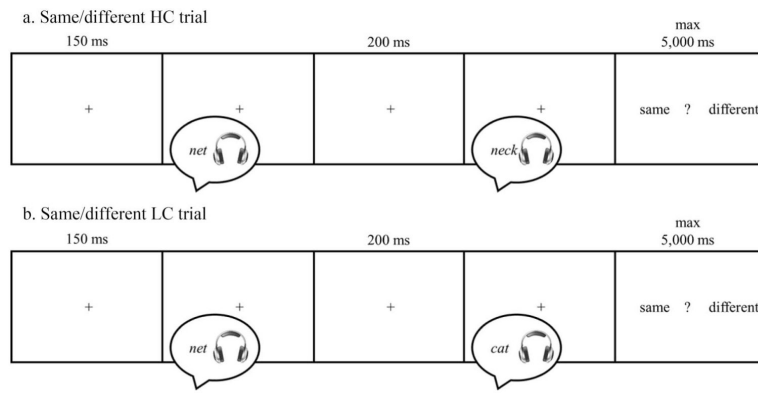
Figure 2.
Examples of trials for the same/different task for the high-competition (HC) group (a) and low-competition (LC) group (b).
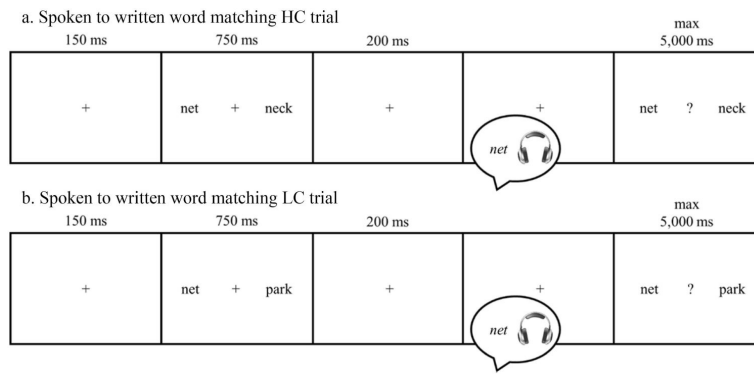
Figure 3.
Examples of trials for the spoken to written word matching task for the high-competition (HC) group (a) and low-competition (LC) group (b).
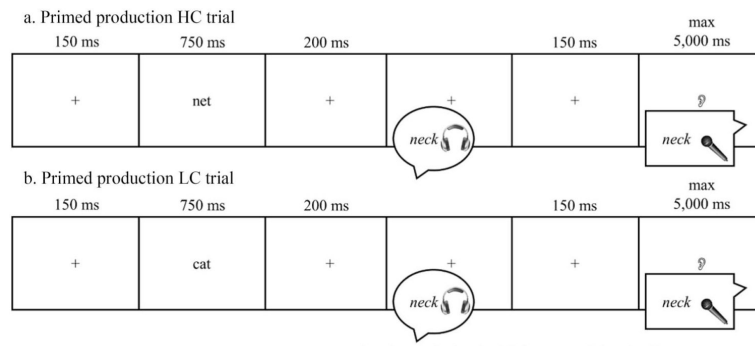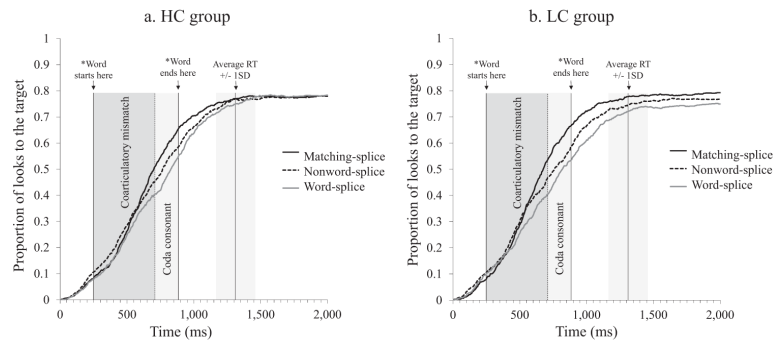
Figure 4.

Examples of trials for the primed production task for the high-competition (HC) group (a) and low-competition (LC) group (b).

Figure 5.
Proportion of looks to the target per splice condition for the high-competition (HC) group (a) and the low-competition (LC) group (b). Time 0 ms indicates the trial onset. RT = reaction time. * Marker on the *x*-axis adjusted by 200 ms to account for oculomotor delay plus 50 ms for silence added to stimulus onset.
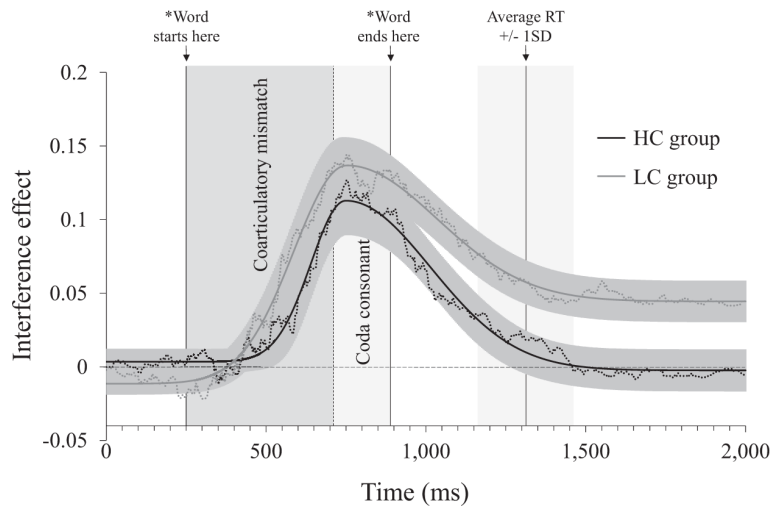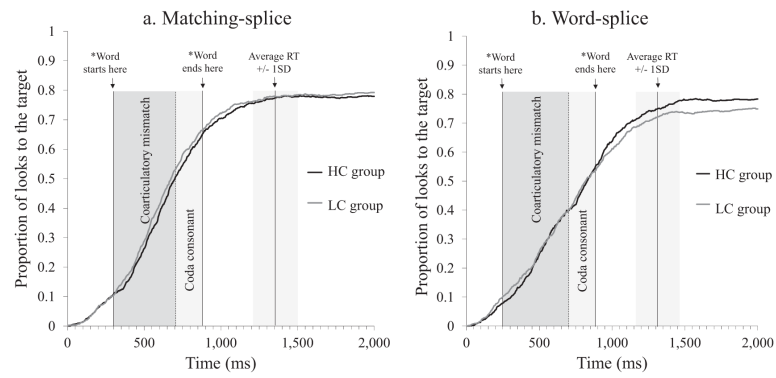
Figure 6.
Interference curves per group (solid lines: fitted data; dotted lines: observed data). Gray shadows represent error bars for the jackknifed/curve-fitted data. RT = reaction time; HC = high-competition; LC = low-competition. * Marker on the *x*-axis adjusted by 200 ms to account for oculomotor delay plus 50 ms for silence at stimulus onset.
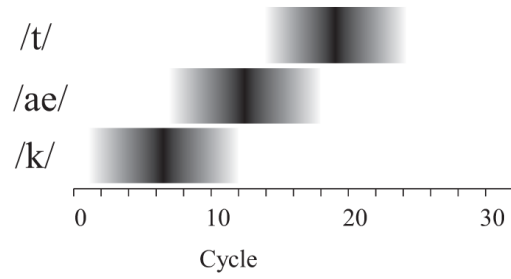
Figure 7.
Proportion of looks to the target in the matching-splice (a) and word-splice (b) conditions per group. RT = reaction time; HC = high-competition; LC = low-competition. [*] Marker on the *x*-axis adjusted by 200 ms to account for oculomotor delay plus 50 ms of silence at stimulus onset.
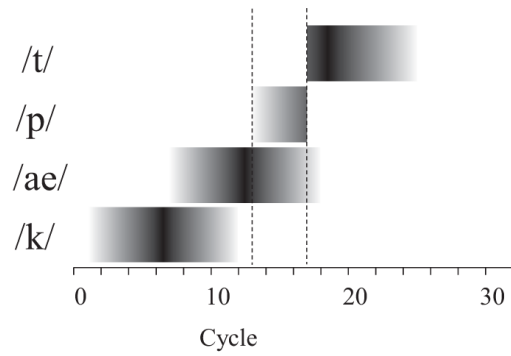
a. Matching-splice

b. Word-splice

Figure 8.
Matching-splice (a) and word-splice (b) as presented in TRACE. Dotted lines mark the beginning and end of coarticulatory mismatch.
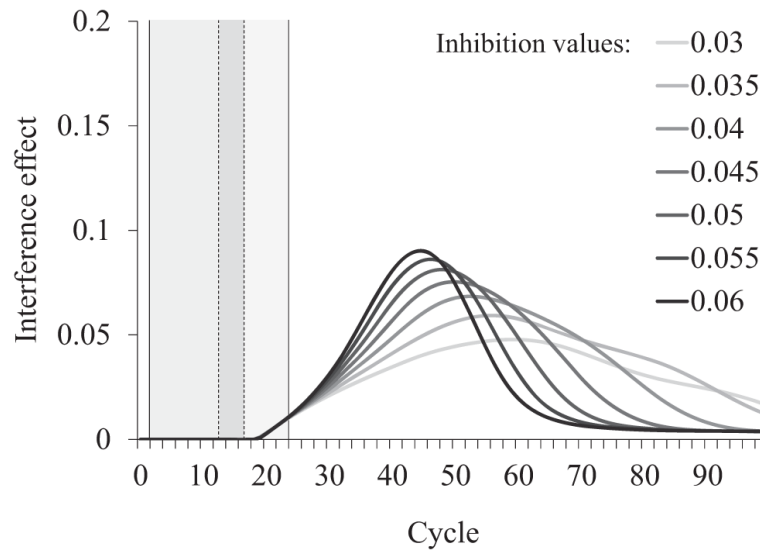
Figure 9.
Interference effect (i.e., target activation in matching-minus word-splice; see Figure 8) in jTRACE simulations as a function of lexical inhibition. Dotted lines mark the beginning and end of coarticulatory mismatch for the word-splice condition.
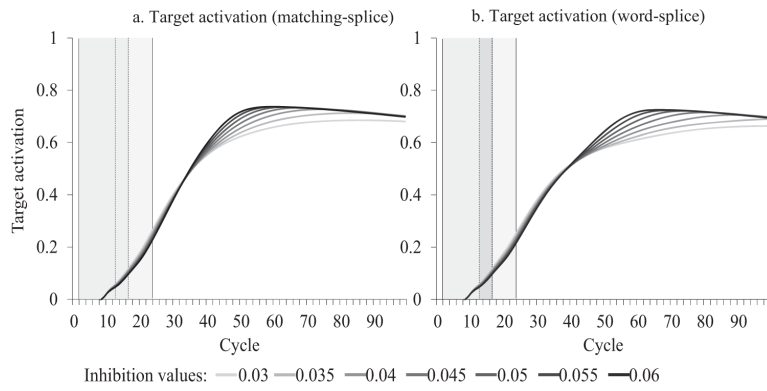
Figure 10.
Target activation, when it is spliced with itself (matching-splice; panel a) and when it is spliced with the competitor (word-splice; panel b) as a function of inhibition level. Dotted lines mark the beginning and end of coarticulatory mismatch for the word-splice condition.
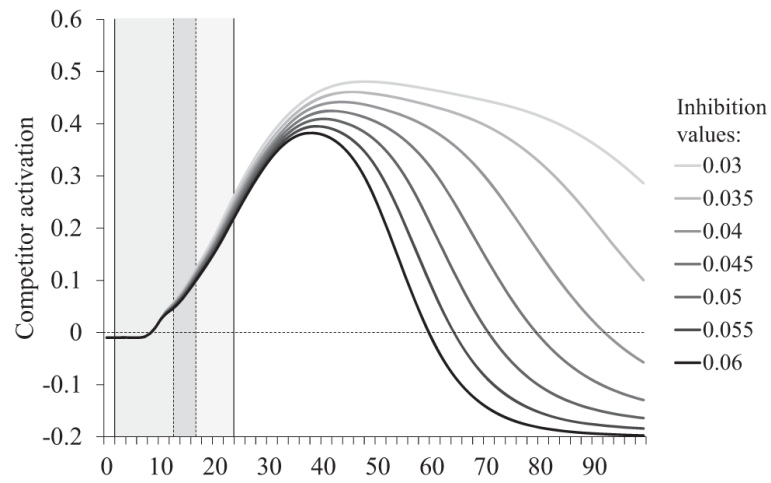
Figure 11.

Competitor activation when the target is spliced with the competitor (i.e., word-splice) as a function of lexical inhibition. Dotted lines mark the beginning and end of coarticulatory mismatch for the word-splice condition.

Table 1

Description of Training Tasks

Task	Description	HC group	LC group
Phoneme monitoring	See two letters; then hear one word; pick the letter that matches the last sound in the word	Both letters are possible endings of words with the same onset (e.g., net/neck)	Only one letter is a possible ending of a word with a given onset (e.g., net/nep)
Same/different	Hear two words one after the other; report whether they ended in the same or a different sound	Both words have same onset (e.g., cat/cap)	Words have different onsets (e.g., cat/neck)
Spoken to written word matching	See two words; then hear one word; report which of the two words on the screen they heard	Two words on screen have same onset	Two words on screen have different onset
Primed production	See a word; then hear the same/a different word; see an eye/ear; say out loud the word they saw/heard	When audio/visual words differ, they have same onset	When audio/visual words differ, they have different onsets

Note. HC = high competition; LC = low competition.

Table 2

Structure of the Training Trials

	Block 1	Block 2	Block 3	Block 4	Total
Phoneme monitoring	28 trials	28 trials	28 trials	28 trials	112 trials
Same/different	28 trials	28 trials	28 trials	28 trials	112 trials
Spoken to written word matching	42 trials	42 trials	42 trials	42 trials	168 trials
Primed production	56 trials	56 trials	56 trials	56 trials	224 trials

Table 3

Group Comparisons for the Six Parameters of the Asymmetric Gaussians (Fits of Interference Effect Curves)

Parameter	By subjects analyses		By items analyses (paired t)	
	$t(72)$	p	$t(27)$	p
Peak location (μ)	0.07	0.94	0.27	0.79
Peak height (p)	0.79	0.43	0.62	0.54
Onset baseline (b_1)	−1.30	0.20	−1.53	0.14
Onset slope (σ_1)	1.20	0.24	1.04	0.31
Offset baseline (b_2)	2.34	0.022	2.24	0.03
Offset slope (σ_2)	0.26	0.80	0.07	0.94

Note. All *t* tests reported here are adjusted for the jackknifing procedure (Miller et al., 1998).

Table 4

Group Comparisons for the Six Parameters of the Asymmetric Gaussians (Fits of Coarticulatory Mismatch Effect Curves)

Parameter	By subjects analyses		By items analyses (paired t)	
	$t(72)$	p	$t(27)$	p
Peak location (μ)	0.23	0.82	−0.07	0.94
Peak height (p)	0.78	0.44	0.54	0.59
Onset baseline (b_1)	0.18	0.86	0.11	0.91
Onset slope (σ_1)	0.54	0.59	−0.02	0.98
Offset baseline (b_2)	0.98	0.33	0.16	0.87
Offset slope (σ_2)	−0.16	0.87	0.01	1.00

Note. All t tests reported here are adjusted for the jackknifing procedure (Miller et al., 1998).