ELSEVIER

Computational and behavioral investigations of lexically induced delays in phoneme recognition

Daniel Mirman *, James L. McClelland, Lori L. Holt

Department of Psychology, Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Previous studies have failed to demonstrate lexically induced delays in phoneme recognition, casting doubt on interactive models of speech perception. We present TRACE simulations that explain these failures: previously tested conditions failed to produce lexically induced delay effects because the input was too unambiguous and the control condition was conflated with lexical status and neighborhood structure. Since between-layer connections are solely excitatory, between-layer delay effects can emerge only indirectly through facilitation of within-layer competition. If the lexically consistent phoneme partially matches the input acoustics, it will become partially active. Additional support from lexical feedback will extend the duration of competition between the acoustically present phoneme and the lexically consistent phoneme, thus delaying detection. This prediction holds across a range of relevant parameter values. Two behavioral experiments tested and confirmed this prediction. These results answer one of the challenges to the interactive view of speech perception.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Speech perception; Phoneme recognition; Lexical inhibition; Interactive processing; TRACE; Word recognition

Numerous studies have shown that phonemes are detected faster in words than nonwords (e.g., Rubin, Turvey, & Van Gelder, 1976) and faster in more word-like nonwords than less word-like nonwords (e.g., Connine, Titone, Deelman, & Blasko, 1997; see also Luce & Large, 2001; Newman, Sawusch, & Luce, 1997). This lexical facilitation of phoneme perception can be explained either by an interactive model in which lexical activation has a direct influence on phoneme processing (e.g., TRACE; McClelland & Elman, 1986), or by an autonomous model in which lexical and phonemic infor-

mation are combined at a later decision stage (e.g., Merge; Norris, McQueen, & Cutler, 2000).

Under the interactive view, activation of lexical items feeds back to the phoneme layer, providing additional activation of lexically consistent phoneme targets. This additional activation speeds resolution of competition at the phoneme layer. Thus, phonemes that are consistent with a lexical item are detected faster than phonemes in nonwords because phonemes in nonwords do not benefit from lexical feedback. Intuitively, lexical facilitation is just one side of the interactive coin; the other side is that lexical feedback should delay recognition of lexically inconsistent phonemes. For example, when an input such as /abɔlI/ is presented, the lexical representation of "abolish" should become active and support the phoneme /ʃ/ in the final

---

* Corresponding author. Fax: +1 412 268 2798.
   E-mail address: dmirman@andrew.cmu.edu (D. Mirman).

position, but if a lexically inconsistent phoneme (such as /s/ or /k/) occurs there, phoneme detection should be slowed due to within-layer competition with the lexically consistent phoneme (/ʃ/). Interactions in the TRACE model consist of within-layer competition (through mutually inhibitory connections) and between-layer excitation (through excitatory connections between consistent elements). Thus, a lexical item cannot directly inhibit a phoneme that is inconsistent with it, but it can delay recognition of this phoneme by exciting a lexically consistent phoneme and creating or extending competition among partially active phonemes.

Two sets of experiments (Frauenfelder, Segui, & Dijkstra, 1990; Wurm & Samuel, 1997) have failed to find evidence of lexically induced delays in phoneme recognition. In both cases, participants were asked to press a button when they heard the target phoneme (e.g., /t/; in each example the target phoneme is underlined). There were three important types of target-bearing tokens: "inhibiting nonwords" (INW) in which only the target had been changed (e.g., "vocabu<u>t</u>ary"), "control nonwords" (CNW) in which an additional phoneme had been changed (e.g., "socabu<u>t</u>ary") and "true nonwords" (TNW) in which one phoneme per syllable had been changed (e.g., "kigrona<u>t</u>y"). The crucial comparison was reaction time to detect the target phoneme in INW compared to CNW and TNW. The logic was that since INW tokens are more consistent with a particular lexical item, they should cause greater lexical activation, but since the target phoneme is inconsistent with the lexical item (e.g., /t/ in "vocabulary"), the greater lexical activation should cause slower target detection times. Thus, the lexicon would induce a greater delay in recognition of phoneme targets in INW compared to targets in CNW or TNW, because for the INW there would be more lexical activation to inhibit the phoneme target. Neither set of experiments found such an effect. Importantly, Frauenfelder et al. compared items with lexically consistent and lexically inconsistent initial phonemes and found that this difference was sufficient to produce a lexical facilitation effect (e.g., the /t/ in "gladia<u>t</u>eur"[1] was detected faster than in "bladia<u>t</u>eur") but not a lexically induced delay effect (e.g., no difference in detection of the /t/ in "vocabu<u>t</u>aire" and "socabu<u>t</u>aire"). This result shows that changing the initial phoneme influences lexical activation enough to cause differences in lexical facilitation, but does not cause the comparable lexical delay effect. Frauenfelder et al. noted that acoustic similarity is important to partial activation of phonemes. This point was

supported by simulations of the TRACE model reported briefly in Peeters, Frauenfelder, and Wittenburg (1989). Frauenfelder et al. therefore tested INW and CNW items in which the phoneme target was similar to the lexically consistent one (e.g., *simplici<u>d</u>é* and *fim-plici<u>d</u>é* from *simplicité*; their Experiment 4), but again found no evidence of lexically induced delays. Thus, in spite of some care in their efforts to uncover a possible inhibitory effect, the study of Frauenfelder et al. failed to find any evidence for it.

The Merge model (Norris et al., 2000) was proposed as an autonomous alternative to TRACE. One of the motivations for the Merge model was to account for the above failures to demonstrate lexically induced delays in phoneme recognition. Norris et al. raised a further concern: if direct lexical feedback can activate lexically consistent phonemes that are not present in the input, an interactive model can "run the risk of hallucinating" (Norris et al., 2000, p. 302). That is, if interactive models predict lexically induced delays in phoneme recognition (which have not been demonstrated), they may further predict lexically induced *errors* in phoneme recognition.

The discrepancy between reported simulations and the behavioral failures to demonstrate lexically induced delays in phoneme recognition are a challenge to the TRACE model. A complete answer to this challenge requires three steps. First, it is important to examine whether the TRACE model predicts delays under the tested conditions and explain why it does or does not. Second, it is necessary to elucidate a set of conditions under which the TRACE model predicts a robust lexical delay effect. Third, TRACE-predicted lexical delay effects must be tested behaviorally. It is important to answer this challenge. On the one hand, TRACE is consistent with a wide range of findings: a variety of direct and indirect lexical influences on phoneme perception in words (e.g., Ganong, 1980; Magnuson, McMurray, Tanenhaus, & Aslin, 2003a; Rubin et al., 1976; Samuel, 1997, 2001) and nonwords (Connine et al., 1997; Newman et al., 1997), dynamics of lexical activation and competition as demonstrated by eye-tracking (Allopenna, Magnuson, & Tanenhaus, 1998) and gating experiments (e.g., Tyler & Wessels, 1983), sensitivity both to phonotactic regularities and effects of individual items (e.g., Massaro & Cohen, 1983; McClelland & Elman, 1986), categorical perception and trading relations in the identification of phonemes (e.g., Liberman, Harris, Hoffman, & Griffith, 1957; Denes, 1955; see McClelland & Elman, 1986), and other findings. On the other hand, the lack of evidence for lexical delay effects and a few other issues have been proposed as reasons to prefer alternative accounts (Norris et al., 2000). Each of the raised objections deserves consideration; this report focuses on the issue of lexically induced delays in phoneme recognition.

---

[1] The Frauenfelder et al. (1990) experiments were carried out in French and the critical findings were replicated in English by Wurm and Samuel (1997).

Two characteristics of previous experiments may have obscured observation of lexical delay effects. First, lexical feedback may not have been able to extend competition at the phoneme layer because the lexically consistent phoneme was not active and lexical input alone was not sufficient to make it active enough to cause significant competition. TRACE simulations presented in this report show that lexical feedback can delay phoneme recognition through competition among phonemes if the lexically consistent phoneme is partially activated as a result of acoustical similarity to the target phoneme that is present in the input. A more general statement of the same principle is that the phoneme layer evaluates the evidence for each phoneme and the influence of any one source of evidence (e.g., lexical input) is dependent on the degree of ambiguity left open by the other sources. Thus, lexical influences on phoneme recognition will be strongest when the acoustic information is ambiguous. Second, previous experiments have conflated the manipulation designed to show lexical delay effects with the lexical status and neighborhood structure of the item at the target location. The INW items (e.g., "vocabutary") are words at the point of the target, but the CNW items (e.g., "socabutary") and TNW items (e.g., "kigronaty") are already nonwords well before the target. Some evidence indicates that detection of phoneme targets in nonwords is hindered by processes other than lack of lexical facilitation. For example, using a dual-task paradigm, Wurm and Samuel (1997) showed that processing nonwords may require greater attentional resources, thus causing slower responses in tasks such as phoneme monitoring. This is particularly problematic for the INW-CNW (or TNW) comparison where the prediction is that responses will be faster to phonemes in CNW (or TNW), but where the CNW (or TNW) becomes a nonword at an earlier point than the INW. Furthermore, recent evidence indicates that onset lexical neighborhood density (words matching the onset of the stimulus) may play a particularly important role in lexical processing (Vitevitch, Armbruster, & Chu, 2004; see also Allopenna et al., 1998).

The extra processing load that occurs when processing a nonword is not explained by either the TRACE model or the Merge model. This issue deserves to be resolved and a complete model of speech perception will have to offer an account of it. However, the present studies focus on lexically induced delay effects; thus, we simply note that CNW and TNW may be poor comparisons for INW given that processing of nonwords imposes an extra processing load. We address the problem by using a design that avoids the problem inherent in this comparison.

In this report, we provide an analysis of lexically induced delays of phoneme recognition in the TRACE model and behavioral tests of TRACE predictions. First, we present simulations of the TRACE model that address the experimental conditions used by Frauenfelder et al. (1990) and Wurm and Samuel (1997). Under these conditions, the TRACE model produces a small inhibitory effect, but variability in activation arising from variability in the lexical neighborhoods of paired items tends to obscure this small effect. The effect is also made more difficult to detect due to the generalized slowdown in processing nonwords that was noted above, indicating that the conditions used previously may not be conducive to finding an inhibitory effect even if such an effect actually exists. Second, we present TRACE simulations that produce lexically induced delays in phoneme recognition using a new control condition that avoids the problems associated with the use of CNW and TNW items as control stimuli. Finally, we present two experiments that provide behavioral evidence of an inhibitory effect with the new control condition, supporting the TRACE prediction.

Simulation 1: Past findings

Methods

As a first step, it was important to test the TRACE model under the conditions of previous experiments. The implemented version of TRACE has a restricted phonetic inventory, so it was not possible to test the model on the exact stimuli used in the previous experiments. Instead, inputs were designed according to the principles of the experiments carried out by Frauenfelder et al. (1990) and Wurm and Samuel (1997). Five conditions were tested: words (e.g., /sikrʌt/, "secret"); matched nonwords (MNW), in which the first phoneme has been replaced with a phoneme that differed by one feature (e.g., /ʃikrʌt/ from "secret"); inhibitory nonwords (INW), in which the target phoneme replaced a phoneme that differed by at least two features (e.g., /parʃʌt/ from "partial"); control nonwords (CNW), which were derived from INW by changing the first phoneme as with the MNW (e.g., /karʃʌt/); and true nonwords (TNW) in which at least two more phonemes were replaced (e.g., /bulsʌt/). In each of these five conditions the phoneme target was either /l/ or /t/. In addition, two more sets of nonword stimuli were designed to match Frauenfelder et al.'s Experiment 4, which was intended to address the issue of acoustic similarity. In this experiment, INW items contained a target phoneme that was acoustically similar to the lexically consistent phoneme (to distinguish them from the other INW items, we will refer to these as "near nonwords" or NNW) and CNW items ("control near nonwords" or CNNW, for consistency), which were

constructed the same way as items in their previous experiments. The phoneme targets in NNW and CNNW were either /k/ or /ʃ/. In the context of TRACE, we define acoustically similar phonemes as those that differ by one feature. Recall that Peeters et al. (1989) reported simulations that showed that these conditions should produce lexical inhibition but Frauenfelder et al. found no delay. The full set of test stimuli is in Appendix A. The simulations were carried out using the original implementation of the TRACE model with the parameter settings reported in McClelland and Elman (1986; see Appendix B for full set of parameters). The lexicon was a slightly enhanced version of the lexicon used by McClelland and Elman and contained 229 words (a few words were added to allow 10 items per condition in our simulations). The lexicon was not designed to mimic English lexical neighborhoods, but it has both dense and sparse neighborhoods to allow testing of neighborhood effects (see McClelland & Elman, 1986).

In TRACE, the probability of the model choosing a particular response i from a set of alternatives indexed by j was computed using the Luce (1959) choice rule:

$$p(R_i) = \frac{e^{ka_i}}{\sum_j e^{ka_j}},$$

where a_i is the activation level of phoneme i. Under the Luce rule, phonemes with higher activation are more likely to produce the response and all response probabilities add to 1.0. To compare simulation results to behav-ioral response time data, simulation RT was computed as the number of cycles required to reach a Luce proba-bility threshold of 0.95.

Results

Fig. 1 shows the Luce choice probability of choosing the target as a function of processing cycles in each of the five tested conditions. Targets were recognized fast-est in words, somewhat slower in matched nonwords, and even slower in the other three nonword conditions, which were approximately equal to each other. Fig. 2A shows the mean reaction times (in cycles required for Luce choice probability to reach threshold) for targets in each of the five conditions. Note that the consistency in Fig. 1 indicates that the same basic pattern would hold for a range of Luce probability thresholds. Paired samples t tests indicated that targets were detected faster in words (e.g., /sikrʌt/) than MNW (e.g., /ʃikrʌt/), (t(9) = 3.087, p = .013), which is consistent with the lex-ical facilitation effect. However, detection of targets in INW (e.g., /parʃʌt/) was not significantly slower than detection of targets in CNW (e.g., /karʃʌt/), (t(9) = 0.754, p = .470). Thus, there was no evidence of a lexically induced delay in phoneme recognition. These simulation results are consistent with the behav-ioral data of Frauenfelder et al.'s (1990) Experiment 3, which showed lexical facilitation when comparing target detection rates in words and MNW, but no lexical delay when comparing target detection rates in INW and CNW.
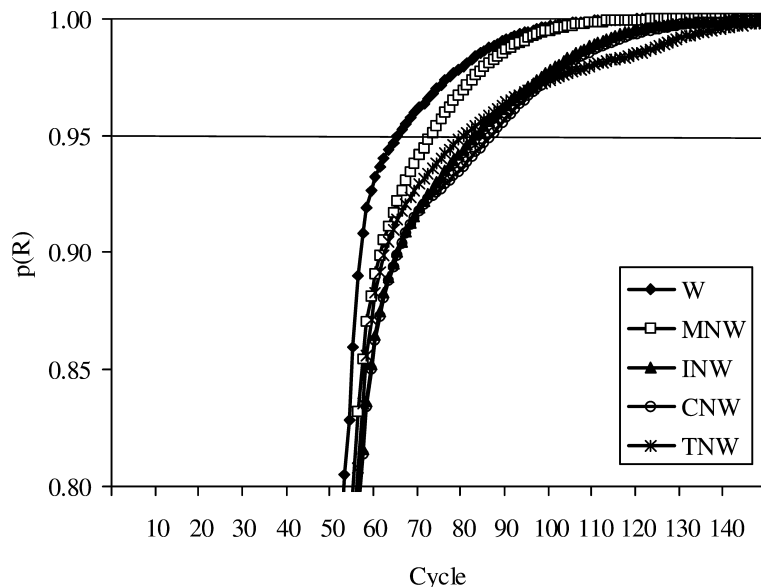

Fig. 1. Luce choice probability for targets in each of the five conditions tested in previous experiments. Phoneme targets are recognized fastest in words, somewhat slower in MNW, and slowest in the other three conditions.
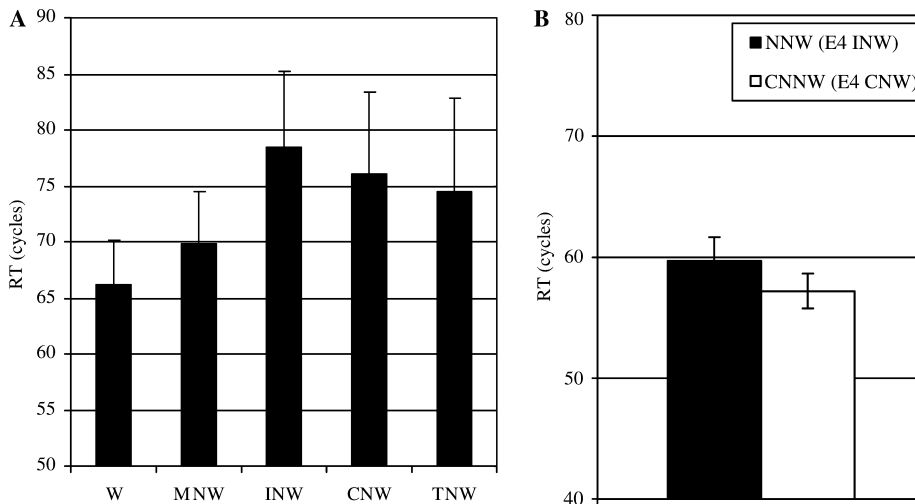
Fig. 2. (A) Simulated reaction time (RT) in each of the five conditions tested in previous experiments. (B) Simulated RT for the acoustically similar conditions tested in Frauenfelder et al.'s (1990) Experiment 4. RT is measured by number of processing cycles required to reach a Luce response threshold of 0.95. Error bars show standard error.

Detection of targets in INW was not significantly slower than detection of targets in TNW[2] (e.g., /bulsʌt/), ($t(8) = 0.249$, $p = .809$). In behavioral studies (Wurm & Samuel, 1997), targets in inhibitory non-words were detected *faster* than in true nonwords. This difference may be due to participants' additional difficulty processing nonwords items, for example if processing nonwords places greater attentional demands on the cognitive system (Wurm & Samuel, 1997). The pattern appears to suggest that the effect is a matter of degree, with items differing only slightly from words inducing less of a load than items that are less word-like. Such a graded effect would account for the model's underestimation of response times in the matched non-word condition relative to the other nonword conditions.

Fig. 2B shows our simulation results for Frauenfelder et al.'s (1990) Experiment 4 conditions: targets in NNW were recognized slower than in CNNW, but this difference was not statistically reliable ($t(9) = 1.795$, $p = .106$). To understand this lack of statistical reliability it is useful to examine the processing of two NNW-CNNW pairs. For the NNW /prɔgrʌʃ/ (from progress), the lexical item "progress" dominated activity in the lexical layer (with minor initial competition from "plot" and "plug"). This lexical activity provided feedback support for /s/, which competed with the acoustically present /ʃ/. For the paired CNNW /brɔgrʌʃ/ lexical activity

was dominated by /b/-initial items ("block" and some others) with partial activation of "progress." This lexical activity did not provide feedback to the phoneme layer that distinguished between /s/ and /ʃ/. Thus, this pair demonstrated a lexically mediated delay in phoneme recognition. In contrast, consider the NNW–CNNW pair /prɔduʃ/–/trɔduʃ/ (from "produce"). The NNW /prɔduʃ/ activated a somewhat higher density lexical neighborhood, thus the lexical item "produce" did not dominate lexical activation as clearly as "progress" did for /prɔgrʌʃ/. As a result, the /s/ phoneme did not receive as much feedback support and provided less competition to /ʃ/. The paired CNNW, /trɔduʃ/, activated a dense /trV/-initial neighborhood ("trot," "true," "tree," "treat," "treaty") with many fricative-initial words slightly active in response to the /ʃ/ (i.e., interpreting the input as the beginning of two words, such as "treaty sheet"). The TRACE lexicon happened to have more /s/-initial words than /ʃ/-initial words, so the /s/ phoneme received additional feedback support, causing a small reversal of the lexical delay effect (target detected slower in CNNW than in NNW). The key point of these examples is that, according to the TRACE model, lexically induced delay effects are very sensitive to the structure of the lexical neighborhoods of the items and lexical neighborhood effects can be stronger than lexical delay effects.

Discussion

TRACE simulation results were consistent with the previous experimental failures to demonstrate lexical

---

[2] For one of the TNW items (aplˆpl), no single phoneme response reached threshold, that is, the model produced no response. Thus, this trial was excluded from the analysis.

delay effects. In the TRACE model, activation of different elements increases gradually and a delay in reaching threshold can be caused by competition among multiple partially active units. For phoneme recognition to be delayed through lexical influence, lexical feedback must cause (or contribute to) competition among different units in the phoneme layer. However, initially phoneme unit activations are at rest and below the threshold they must reach to begin competing with other phoneme units. Lexical feedback alone is not enough to activate a phoneme unit above threshold. As a result, without additional support, the lexically consistent phoneme does not enter into competition with the acoustically present phoneme and does not delay its recognition. For example, when the model is presented with the inhibitory nonword /parʃʌt/, the phoneme unit for the lexically consistent phoneme /l/ does not become active and the feedback input from the lexical item ''partial'' is not enough to excite the /l/ unit above threshold. Since the phoneme unit for /l/ does not go above threshold, it does not compete with /t/, so it does not delay the recognition of /t/. This analysis shows that despite its use of feedback, the TRACE model does not ''hallucinate'' a lexically consistent phoneme in the face of strong bottom-up support for an alternative.

In consideration of the importance of acoustic similarity (and simulations reported by Peeters et al., 1989), Frauenfelder et al. (1990) tested delays in phoneme recognition due to lexical support of similar phonemes (their Experiment 4), but once again found no difference between response times to targets in NNW (e.g., ''*simplicidé*'') and CNNW (e.g., ''*fimplicidé*''). Our simulations indicate that this comparison is more appropriate for demonstrating lexically induced delays; however, it is still highly dependent on lexical neighborhood structure. (The report of the Peeters et al. simulation is very brief, and provides no measures of reliability of the reported effects. Thus it is possible that the effect they found in their reported simulation would not have been reliable across items.) Furthermore, as mentioned above, the INW–CNW (and NNW–CNNW) comparison is conflated with differences between the conditions in the word status of the stimulus prior to the occurrence of the target phoneme. That is, at the point when the target phoneme occurs, the INW (or NNW) is a word but the CNW (or CNNW) is already a nonword. An extra processing load that arises from processing the nonword could slow performance on CNW's (and CNNW's), potentially masking a real inhibitory effect that would otherwise have been apparent. Thus, the experimental comparison used by Frauenfelder et al. may not have been optimal for revealing a lexically induced delay effect.

This analysis provides two guidelines for empirically demonstrating lexically induced delays in phoneme recognition: (1) The lexically consistent phoneme must be partially active for lexical influence to cause a delay in recognition of the acoustically present, but lexically inconsistent, phoneme. One case that would satisfy this condition is acoustic similarity between the lexically consistent phoneme and the phoneme present in the input. Thus, the lexically consistent and acoustically present phonemes must be acoustically similar for delay effects to emerge. (2) The point of deviation from the lexical item must be matched between the comparison conditions to control for neighborhood effects and processing differences between words and nonwords. The next simulation tests lexically induced delays in phoneme recognition using a design informed by these concerns.

Simulation 2: New prediction

Methods

Stimuli were designed to test for a delay in phoneme detection when the target phoneme was lexically inconsistent but was acoustically similar to the lexically consistent phoneme (e.g., /s/ and /ʃ/ or /t/ and /k/). Under these conditions, the acoustic input partially matches multiple phonemes, which compete through lateral inhibition. Lexical feedback can contribute to and extend the duration of that competition, thus delaying the response. For example, when /s/ is presented as input, initially both /s/ and /ʃ/ phoneme units become active because the input is partially consistent with both. The units compete through lateral inhibition. Since the input is more consistent with /s/ than /ʃ/, the /s/ unit wins after a short period of competition. However, if a lexical item that is consistent with /ʃ/ becomes active, feedback will provide additional activation to /ʃ/ and extend the period of competition, thus delaying recognition of the presented phoneme. In contrast, the phoneme unit for /k/ is not active because it is a very poor match to the /s/ input. Thus, there is no competition between the /s/ and /k/ phoneme units and lexical feedback activation of /k/ cannot significantly delay recognition. It follows that the most appropriate comparison condition is one in which the lexical feedback specifies a very different phoneme, which would not have a significant effect on extending the competition at the phoneme layer. In addition, in this design, the point at which the target-bearing item deviates from the lexical item is matched between the two nonword conditions to avoid differential processing demands.

To test for lexically induced delays, we created ''distant nonwords'' (DNW), which were derived from words by changing one phoneme to a dissimilar phoneme (i.e., a phoneme that differs by at least two features, for example /sikrʌʃ/ from ''secret'' or /pragrʌk/

from "progress"). The number of cycles required for TRACE to recognize a phoneme in DNW was compared to the response time for near nonwords from Simulation 1 (which were derived from words by changing one phoneme to a similar phoneme, e.g., /sikrʌk/ from "secret" or /pragrʌʃ/ from "progress"). The prediction was that response time to detect the target phoneme in a NNW would be slower than in a DNW because lexical feedback will extend the competition between phonemes for NNW but not for DNW. Note that DNW and NNW items have identical lexical neighborhoods up to the point of the target and deviate from words at the same point. Thus, these stimuli are matched for both of the factors that may have caused previous failures to demonstrate lexically induced delays in phoneme recognition.

The targets always occurred after the uniqueness point of the word. The stimuli were counterbalanced so that individual phoneme differences would not influence the results. That is, for half of the stimuli /k/ was the lexically inconsistent similar phoneme and /ʃ/ was the lexically inconsistent dissimilar phoneme, but for the other half this was reversed. The full set of stimuli is available in Appendix A.

Results and discussion

Fig. 3 shows mean simulation reaction times for each of the three conditions. The crucial finding was that the model was slower to detect targets in NNW than in DNW (e.g., /k/ in /sikrʌ<u>k</u>/ compared to the /ʃ/ in /sikrʌʃ/ and the /ʃ/ in /pragrʌʃ/ compared to the /k/ in /pragrʌ<u>k</u>/;
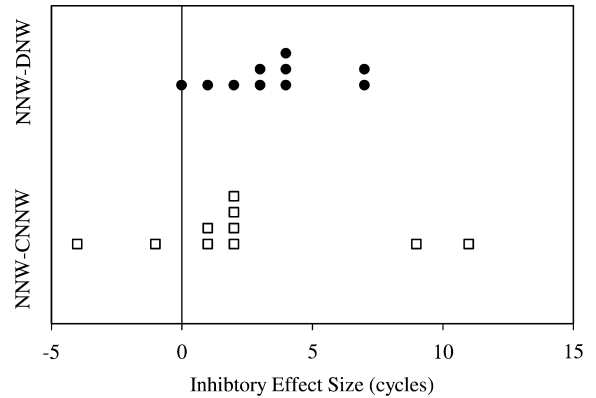
Fig. 4. Effect size histograms. Each point marks the inhibitory effect size for a single pair of items. The open squares correspond to NNW–CNNW pairs, which show a broad distribution with two points showing a reversal of the inhibitory effect. The filled circles correspond to NNW–DNW pairs, which show a moderately higher and a much narrower distribution.

$t(9) = 4.869$, $p = .001$). This difference arose because lexical feedback provided additional activation to a phoneme that was partially active due to partial matching to the input (e.g., /t/ when /k/ is presented) and thus extended the duration of competition at the phoneme layer, causing a delay in recognition. In contrast, lexical feedback activation of a dissimilar phoneme, which was not partially active (e.g., /s/ when /k/ is presented), did not extend competition. The overall NNW–DNW difference (3.5 cycles) was moderately larger than the NNW–CNNW difference (2.5 cycles) in Simulation 1 and much more reliable (standard error is two times larger for NNW–CNNW than NNW–DNW). To illustrate these aspects of the data, Fig. 4 shows the distribution of inhibitory effect sizes for the two comparisons. The NNW–DNW comparison yielded a narrower distribution that is more positive (i.e., greater inhibitory effect) than the NNW–CNNW comparison. In sum, the NNW–CNNW comparison showed a weaker inhibitory effect that was more susceptible to other influences, which showed up as noise in the inhibitory effect (see discussion in Simulation 1).

To provide a clear demonstration of the effect of the lexicon, this Simulation 2 was repeated with the lexicon turned off. The results (Luce choice probabilities) are shown in Fig. 5. For NNW Fig. 5A, the lexical influence (compare filled symbols to open symbols) caused an increase in response likelihood for the lexically consistent phoneme (filled squares are higher than open squares). Since phonemes compete through lateral inhibition (and due to the competitive Luce choice rule), this increase came at the cost of the acoustically present phoneme (filled circles are lower than open circles). The difference between filled circles and open
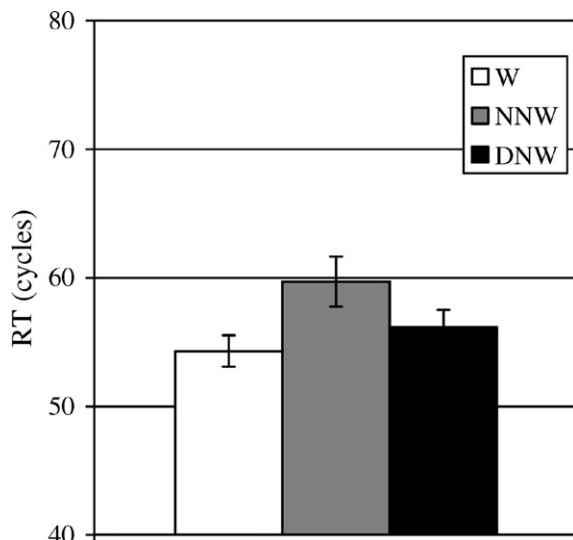

Fig. 3. Simulation RT for new prediction. Phoneme targets in NNW are detected more slowly than in DNW. Error bars show standard error.
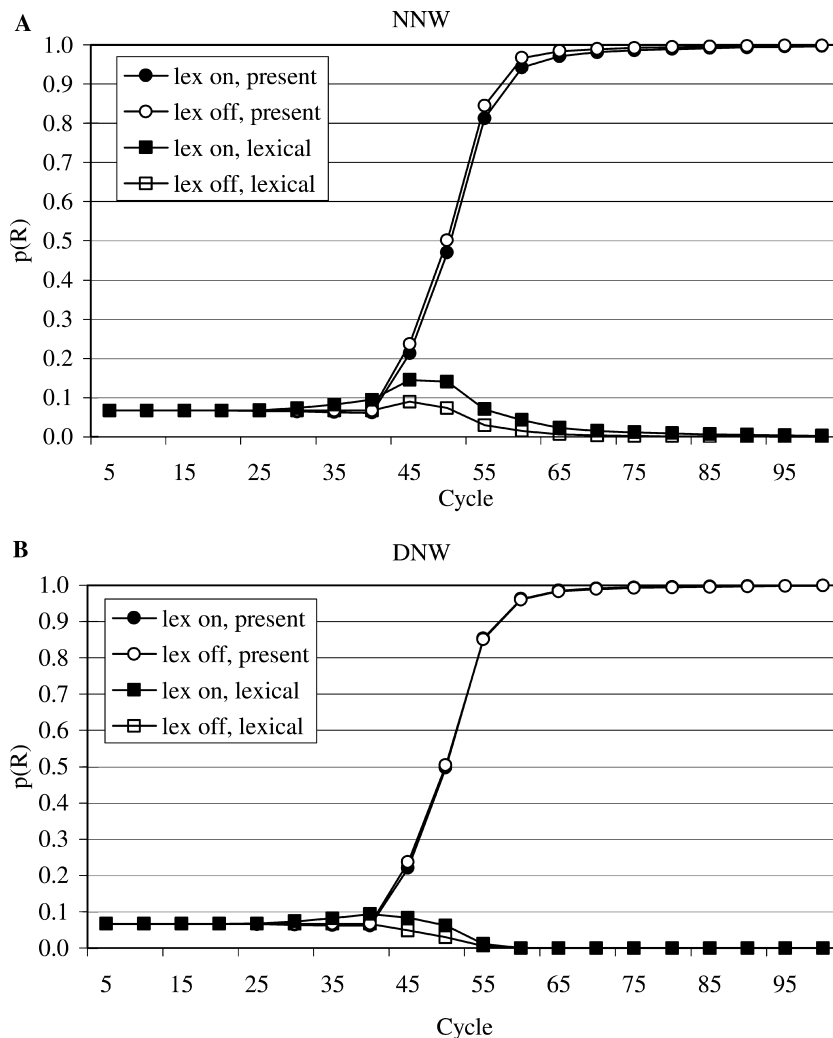
A

NNW

B

DNW

Fig. 5. Luce choice probabilities for acoustically present (circles) and lexically consistent (squares) phonemes in NNW (A) and DNW (B) with (filled symbols) and without (open symbols) influence of the lexicon.

circles is the indirect lexical inhibitory effect. For DNW Fig. 5B the lexical feedback also caused an increase in response likelihood for the lexically consistent phoneme (filled squares are higher than open squares), but since the activation remained below threshold (0), this phoneme did not enter into competition with the acoustically present phoneme and did not influence its activation or response probability (no difference between filled and open circles). In sum, according to the TRACE model, lexical feedback can delay phoneme recognition when the lexically consistent phoneme partially matches the acoustic input and competes with the presented phoneme target. Before testing this prediction in behavioral experiments, we examine the robustness of this prediction to variation of two critical parameters.

Simulation 3: Robustness to parameter changes

Methods

In the TRACE model, lexically induced delays in phoneme recognition can result when lexical feedback contributes to competition at the phoneme layer. This effect is mediated by two critical parameters: word-phoneme feedback excitation and phoneme–phoneme lateral inhibition. The main simulation reported in the previous section was repeated for a ±50% range of values for these two parameters. Specifically, three values for word-phoneme feedback excitation were tested (0.015, 0.03, 0.045) and five values for phoneme–phoneme inhibition were tested (0.02, 0.03, 0.04, 0.05, 0.06).

Results and discussion

Fig. 6 demonstrates that the pattern of simulation RT reported in Simulation 2 persisted across a range of word-phoneme excitation values. Not surprisingly, the lexically induced delay effect (NNW > DNW) grew as the feedback excitation grew; however, even at the comparatively low value of 0.015 the pattern was weak,
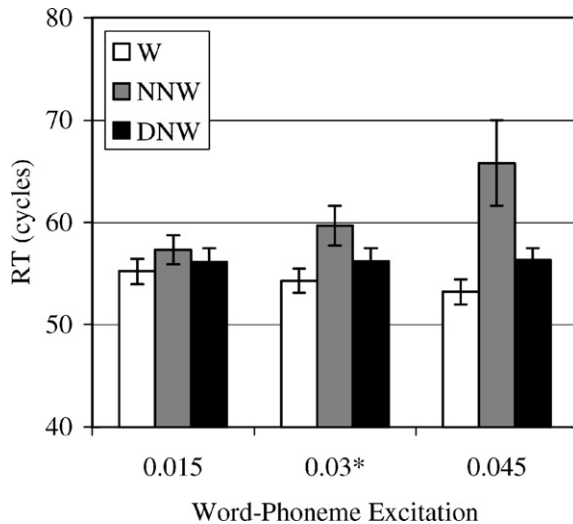

Fig. 6. Simulation RT for new prediction across three values of word–phoneme feedback excitation. The asterisk denotes the standard value of the parameter, which was used in Simulation 2. Error bars show standard error.

but still in the predicted direction. Variation of the phoneme–phoneme inhibitory strength produced similarly consistent results, as shown in Fig. 7. As phoneme–phoneme inhibition was decreased, the time required for the phoneme layer to settle to a single phoneme interpretation increased (because the phonemes competed less strongly), thus allowing more time for lexical feedback to influence phoneme activations. As a result, the lexically induced delay effect was bigger for smaller values of phoneme–phoneme inhibition. The relationship between within-layer inhibition strength and settling time (i.e., response time) is highly nonlinear; as a result, when phoneme–phoneme inhibition is decreased, RT to NNW items (which is most influenced by phoneme–phoneme competition) increases much more rapidly than RT to the other items. At the smallest value (0.02), the competition was so weak that for several NNW items multiple phoneme units remained partially active and no phoneme unit reached the strict 0.95 response threshold. Therefore, the results for the 0.02 condition in Fig. 7 are based on the somewhat lower response threshold of 0.85 (the RT's are faster and the lexical differences are smaller because this threshold is reached more quickly). The robustness of these results indicates that lexically induced delays in phoneme are predicted by the interactive mechanism of the TRACE model, but as shown in Simulation 1, these effects require carefully controlled stimulus materials. The following sections describe experiments testing whether these conditions produce the predicted lexical delay effect in behavioral experiments with human listeners.
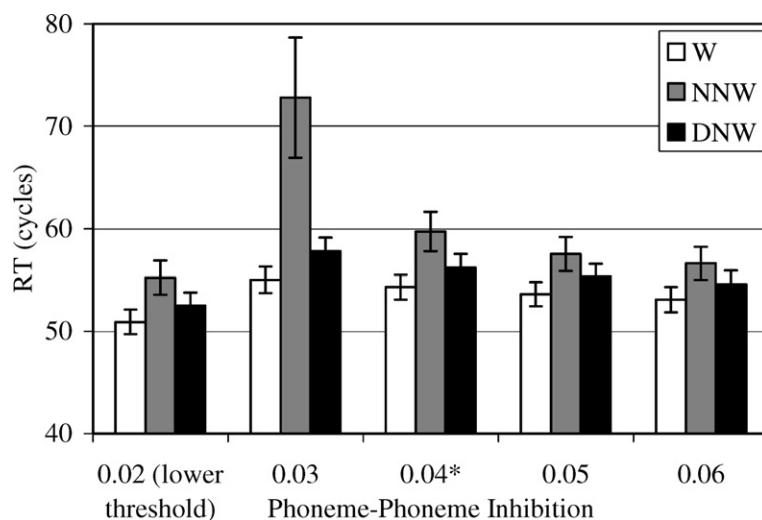

Fig. 7. Simulation RT for new prediction across five values of phoneme–phoneme lateral inhibition. The asterisk denotes the standard value of the parameter, which was used in Simulation 2. Error bars show standard error. Note that RT for the 0.02 condition was computed using a lower response threshold (0.85, 0.95 for other conditions). See text for explanation.

Experiment 1

The purpose of this experiment was to test the prediction of the TRACE model that lexically inconsistent phoneme targets that are similar to the lexically consistent phoneme will be detected more slowly than phonemes that are not similar to the lexically consistent phoneme.

Methods

Procedure

Participants were seated in sound attenuating booths where they heard words and nonwords presented through headphones and made responses using an electronic button box. For each token, participants were asked to determine whether the spoken item contained the target phoneme (/s/ or /ʃ/) or not ("Did the real or fake word contain a 'ss' ['sh'] sound?").

Target phoneme was counterbalanced across participants (i.e., half of the participants monitored for /s/ and half for /ʃ/). Button label assignments ("yes" and "no" to left and right buttons) were also counterbalanced across participants. Before beginning the main block of trials, participants completed a short practice session (10 trials) in which they performed the detection task with feedback. No feedback was provided after the initial practice session. The experiment was completed in a single session lasting approximately 30 min.

Critical Stimuli

The critical stimuli fell into three conditions: real words (W), near nonwords (NNW), and distant nonwords (DNW). Near nonwords were created by changing the word-final fricative of a word ending in /s/ or /ʃ/ to the other fricative (/ʃ/ to /s/ or /s/ to /ʃ/).

Distant nonwords were created by changing the word-final stop consonant of a word ending in /k/ to a fricative (/k/ to /s/ or /ʃ/). Each set of critical items was divided in half and the assignment of each half of the items to target condition (/s/ or /ʃ/) was counterbalanced (i.e., half of the participants heard "abol*ish*" and the other half heard "abol*iss*"; likewise, half heard "academ*ish*" and half heard "academ*iss*").

Critical items contained no /s/ or /ʃ/ phonemes other than the targets and the three sets of words (/ɪs/-, /ɪʃ/-, and /ɪk/-final) were matched as closely as possible for word frequency (Kucera & Francis, 1967), word length in phonemes and syllables, and distance between the uniqueness point and the target position. The full set of word bases for critical items is available in Appendix C.[3]

Filler stimuli

The purpose of the filler stimuli was to prevent participants attending to the particulars of the NNW and DNW stimuli. Filler items with targets in initial and medial positions were added to prevent listeners from focusing attention on the ends of tokens (critical stimuli were all target-final). The majority of filler items were words so that during the experiment, lexical information was generally helpful to task performance, thus promoting its use. Some of the filler items were "true nonwords" (as defined by Wurm & Samuel, 1997) to make the single phoneme manipulation less apparent. In addition, target-initial near nonword filler items were added to reduce the idiosyncrasy of the target-final near nonword critical stimuli. Finally, 120 non-target-bearing words were added to promote use of lexical information and make fricatives somewhat less conspicuous. Table 1 shows examples of each type of token and the number of trials with each type of token. Stimulus lists were compiled with the help of the MRC Psycholinguistic Database (Wilson, 1988)[4] and the CMU Pronouncing Dictionary.[5] Overall, approximately one-third of items contained one phoneme target (121/362), the same number of items (121/362) contained the other phoneme target, and the last third (120/362) were words with no targets, which were added to encourage participants to attend to the lexical level.

Stimulus construction

All stimulus materials were spoken by a phonetically trained male native speaker of unaccented American English in the context of the sentence "The next word is [item]" and recorded at a 11025 Hz sampling rate. Critical items were created by splicing a standard /ɪs/ or /ɪʃ/ ending (taken from recordings of the words "Venice" and "vanish" spoken by the same speaker) onto a word stem to create three different types of items: words (e.g., "abol*ish*"), NNW (e.g., "not*ish*," from "notice"), and DNW (e.g., "frant*ish*," from "frantic"). For the filler items, tokens were spoken in the form to be presented during the experiment and no splicing was performed. Stimuli were low-pass filtered at 5512 Hz to remove high frequency noise.

Participants

Participants were 51 undergraduates from Carnegie Mellon University who received course credit for participation. All participants reported normal hearing and English as their native language.

---

[3] Sound file examples of critical stimuli can be found at http://www.psy.cmu.edu/~lholt/php/gallery.php and the complete set is available from DM (dmirman@andrew.cmu.edu).

[4] http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm.
[5] http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

Table 1
Stimuli for Experiment 1: Examples of each type of token and number of each type of trial (in parentheses)

	Word	NNW	DNW	Total
Critical items				
/ʃ/	Abolish (7)	Apprentish (7)	Academish (7)	21
/s/	Actress (7)	Diminis (7)	Automatis (7)	21
	Word	NNW (INW)	TNW	
Filler items				
/ʃ/ early	Shampoo (20)	Shalary (10)	Shembep (10)	40
/s/ early	Sample (20)	Selter (10)	Sugkuk (10)	40
/ʃ/ mid	Bishop (20)		Deshin (10)	30
/s/ mid	Absorb (20)		Bismep (10)	30
/ʃ/ late	Cherish (20)		Daitish (10)	30
/s/ late	Analysis (20)		Baitis (10)	30
No targets	Apology (120)			120
Total	254	34	74	362

Results and discussion

Ten participants were excluded from analyses because their overall target detection accuracy was below 80% (7 participants; low accuracy may indicate difficulty perceiving the phonemes or lack of motivation) or their overall mean reaction time (measured from target offset) was more than 2 standard deviations above the mean (6 participants; 3 participants fit both exclusion criteria). Analyses including these participants showed the same overall pattern of results, but the additional noise made the differences less reliable. Mean response times (Fig. 8A) and error rates (Fig. 8B) for the remaining 41 participants are shown in Fig. 8. Response times were mea-

sured from target offset and only target-present trials on which the participant provided the correct response were included in the RT analyses. Recall that TRACE predicts slower phoneme recognition for NNW due to extended competition among similar phonemes as a result of lexical feedback. This prediction was borne out in the pattern of response time results: detection of targets in NNW (e.g., "noti<u>sh</u>") was slower than in DNW (e.g., "franti<u>sh</u>"). The main effect of item condition (word, NNW, or DNW) was significant ($F(2, 38) = 10.041$, $p < .001$), pair-wise comparisons indicated that this was because responses to the word condition were faster than NNW ($t(40) = 4.482$, $p < .001$) and DNW ($t(40) = 3.308$, $p = .002$), but the
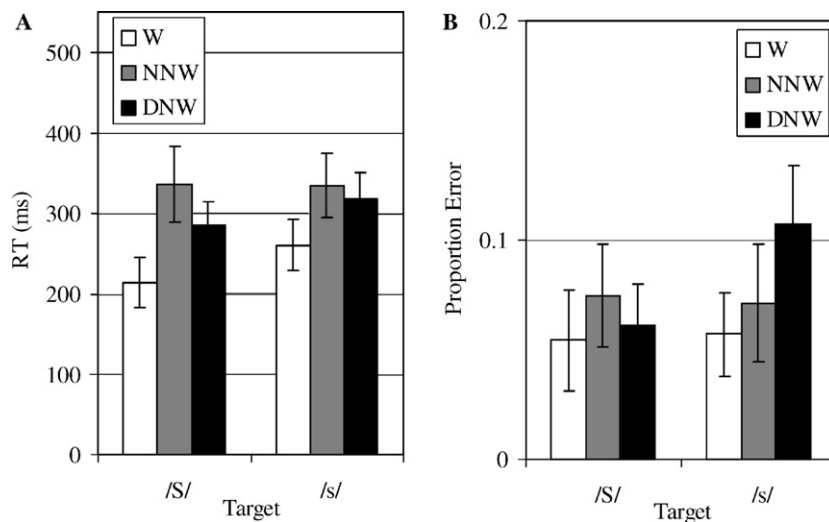

Fig. 8. Experiment 1 results. (A) Listeners were slower to detect phoneme targets in NNW than in DNW for both targets. This difference was not reliable overall. (B) Error rates for /ʃ/-monitoring also showed the predicted lexical inhibition pattern, but not the /s/-monitoring error rates. Differences were not reliable. Error bars show standard error.

Table 2
Stimuli for Experiment 2: Examples of each type of token and number of each type of trial (in parentheses)

	Word	NNW	DNW	Totals
Critical items				
/t/	Deficit (7)	Arsenit (7)	Abolit (7)	21
/k/	Dynamic (7)	Deposik (7)	Blemik (7)	21
	Word	NNW (INW)	TNW	
Filler items				
/t/ early	Tailor (20)	Talendar (10)	Tembep (10)	40
/k/ early	Cabbage (20)	Kenure (10)	Kuspun (10)	40
/t/ mid	Abstain (20)		Detin (10)	30
/k/ mid	Alcohol (20)		Mikep (10)	30
/t/ late	Helmet (20)		Daipit (10)	30
/k/ late	Pacific (20)		Bainik (10)	30
No targets	Apology (120)			120
Total	254	34	74	362

difference between nonword conditions was not reliable ($t(40) = 1.379$, $p = .176$). There was neither a main effect of target (/s/ or /ʃ/, $F < 1$) nor a condition × target interaction ($F < 1$). The same analyses for error rates showed no reliable differences. Inspection of Fig. 8 suggests that there may have been some (albeit statistically unreliable) differences in performance for the two phoneme targets. The /ʃ/-monitoring group ($N = 21$) showed the predicted pattern in both RT and error rates, in fact, /ʃ/ detection in NNW was marginally slower than in DNW ($t(20) = 2.033$, $p = .056$; a complete set of results for target-based analyses is in Appendix D). The /s/-monitoring group ($N = 20$) showed the predicted overall pattern, but the critical difference was not reliable ($t(19) = 0.371$, $p = .715$).

A number of factors may have contributed to the slight asymmetry between the /ʃ/ and /s/ results. The recording and filtering procedure removed all acoustic information above 5512 Hz, which may have removed important /s/-specifying information. In fact, a comparison of overall accuracy (including filler items, since this is a matter of general recording procedure) indicated that /ʃ/ recognition was marginally more accurate than /s/ identification (91.9% for /s/, 95.6% for /ʃ/; $t(39) = 1.789$, $p = .081$). In sum, Experiment 1 provided suggestive evidence but not a clear demonstration of the lexically induced delay in phoneme recognition predicted by TRACE. Experiment 2 used the same basic design but incorporated adjustments reflecting the concerns described above and other minor methodological issues.

Experiment 2

Methods

Although the results of Experiment 1 were consistent with the TRACE predictions, acoustic influences may have obscured the effects. Four changes were introduced in Experiment 2. (1) To avoid possible acoustic effects, Experiment 2 was based on a new stimulus set that used /t/–/k/ as the near phoneme distinction and /t/–/ʃ/ and /k/–/ʃ/ as the distant phoneme distinctions[6] (see Table 2; the full list of word bases for critical items is available in Appendix C). The base tokens were produced by the same speaker as in Experiment 1 and the stimulus construction procedure was the same. The standard /Ik/ and /It/ endings were taken from the words "eunuch" and "unit." (2) The critical items (including "eunuch" and "unit" for the endings) were recorded in the context of the sentence "Say [item] again." According to the TRACE model, for the delay effect to emerge, the bottom-up input must provide some support for the lexically defined phoneme. Embedding the critical item in the middle of the sentence rather than the end (as in Experiment 1) was one way to make the final consonant less distinct. (3) The first 50 trials were constrained to be filler trials with feedback presented on the first 20 trials. This was done to make sure participants learned the task before beginning the critical trials (subsequent time course analysis showed no overall RT decrease during the critical trial phase). (4) To alert participants to the orthographic-phonological correspondence (specifically, that the phoneme /k/ occurs as the letter 'c'), the on-screen instructions read "Did the real or fake word contain a 'k' ['t'] sound as in 'car' ['top']?" Aside from

---

[6] The design of this experiment requires a moderately large set of multisyllabic words that end with each phoneme target. The phoneme preceding the target phoneme must be matched to control for coarticulatory effects (e.g., in the present experiments all target phonemes were preceded by the vowel /I/) and the critical item sets must be matched for word frequency, word length, and distance between target and uniqueness point. The phonemes /t/ and /k/ were the best solution to these numerous constraints.

these differences, the design and procedure of Experiment 2 was identical to Experiment 1.

Participants

Participants were 58 undergraduates from Carnegie Mellon University who had not participated in Experiment 1. Participants received course credit for participation. All participants reported normal hearing and English as their native language.

Results and discussion

Eight participants were excluded from analyses because their overall target detection accuracy was below 80% (6 participants) or their overall mean RT was more than 2 standard deviations above the mean (2 participants). Analyses including these participants showed the same overall pattern of results, but the additional noise made the differences less reliable. Mean response times (Fig. 9B) and error rates (Fig. 9A) for the remaining 50 participants are shown in Fig. 9. Response times were measured from target offset and only correct target-present trials were included in the RT analyses. For reaction times, a main effect of item type was found ($F(2,47) = 23.787$, $p < .001$) and a main effect of target (/k/ was detected faster than /t/; $F(1,48) = 4.891$, $p = .032$), but no item type × target interaction ($F < 1$). Pair-wise comparisons showed that phoneme targets in words were recognized faster than in NNW ($t(49) = 6.042$, $p < .001$) or DNW ($t(49) = 5.379$, $p < .001$). Critically, targets in NNW were recognized more slowly than in DNW ($t(49) = 2.407$, $p = .02$), as predicted by the TRACE model. Furthermore, analysis

of error rates showed a similar pattern: a main effect of item type ($F(2,47) = 3.445$, $p = .04$), no main effect of target ($F < 1$) nor item type × target interaction ($F(2,47) = 2.252$, $p = .111$). Pair-wise comparisons showed that there were fewer errors to words than NNW ($t(49) = 2.597$, $p = .012$; the word-DNW difference was not reliable: $t(49) = 1.323$, $p = .192$) and marginally more errors to NNW and DNW ($t(49) = 1.865$, $p = .068$).

There is a possible confound in these results: DNW word bases may have less mismatch with the endings than NNW word bases. Specifically, the /It/ and /Ik/ tokens came from /n/ contexts ("unit" and "eunuch") and more DNW (5/14 or 10/28 since there is a /It/ and a /Ik/ version of each DNW) have /n/ preceding the ending than NNW (3/28). The incompatibility of non-/n/ context may have unequally slowed NNW responses and thus would account for the NNW > DNW difference. To test this possibility we compared DNW response times to targets in /n/ context and in non-/n/ context. If /n/ context mismatch influenced our results, targets in non-/n/ context DNW items should be detected more slowly than targets in /n/ context DNW items. Phoneme targets in /n/ context DNW were detected 6 ms faster than targets in non-/n/ DNW. This difference was not reliable either when all 28 DNW are considered ($t(26) = 0.182$, $p = .857$), nor when the two DNW types (/It/-final and /Ik/-final) are considered separately ($t(12) < 0.5$, $p > .6$ for each). Considering that the NNW-DNW effect size was nearly 10 times bigger (55 ms) and that this /n/-context effect did not even approach statistical significance, we conclude that phonological context mismatch did not influence our results.
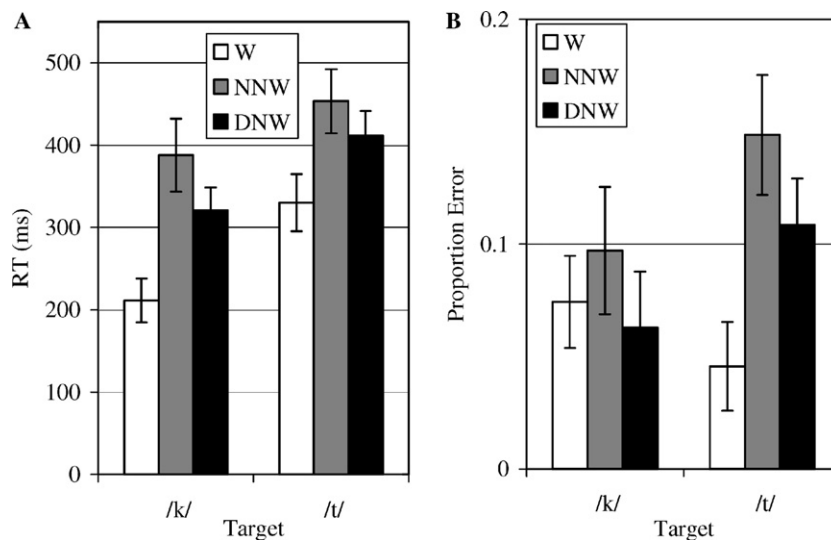

Fig. 9. Experiment 2 results. (A) Listeners are slower to detect phoneme targets in NNW than in DNW. (B) Listeners were marginally more likely to miss a phoneme target when it occurred in NNW than DNW. Error bars show standard error.

The results of this experiment provide behavioral evidence in support of the TRACE model prediction of lexically induced delays in phoneme recognition. This lexical delay was demonstrated when the acoustically present phoneme was not lexically consistent, but was acoustically similar to the lexically consistent phoneme. Under these conditions, the lexical feedback was able to extend competition at the phoneme layer and thus cause a delay in phoneme recognition. By comparison, when the lexically consistent phoneme was very different from the acoustically presented phoneme, lexical feedback was not strong enough to cause competition and there was no lexical delay effect. In addition, by matching the point at which the item becomes a nonword between the two critical conditions, the design used in these experiments controlled for the possible additional costs of processing nonwords. The lexically induced delay in phoneme recognition was demonstrated most clearly in Experiment 2, but the same pattern of response time data was found for each of the four phoneme targets across both experiments and the pattern was mirrored in error rates for three of the four phoneme targets.

General discussion

The interactive view of speech perception holds that phoneme perception is based on both bottom-up acoustic processing and top-down feedback from lexical processing. Under this view, lexical feedback facilitates recognition of phonemes that are consistent with the lexical context. This robust lexical facilitation is demonstrated, for example, by faster recognition of phonemes when they are in words than nonwords (e.g., Connine et al., 1997; Rubin et al., 1976). Such direct influence on phoneme processing might also delay recognition of phonemes that are not consistent with the lexical context. Two attempts to demonstrate such lexical delay effects in behavioral experiments have failed (Frauenfelder et al., 1990; Wurm & Samuel, 1997), and thus cast doubt on the interactive view (cf. Norris et al., 2000).

In this report, an empirical investigation of TRACE with careful attention to details such as lexical neighborhoods has revealed that the interactive model is consistent with the previous findings. In TRACE, between-layer connections are solely excitatory (as in other interactive activation models, see Usher & McClelland, 2001 for discussion). Thus, feedback connections from the lexical layer to the phoneme layer can cause direct facilitation effects but inhibition effects must be indirect. Such indirect inhibition can result from the competitive dynamics of phoneme processing. If lexical feedback contributes to the activation of a competing phoneme, it will take longer for the acoustically present phoneme to win the competition, thus delaying the response time. The simulations suggested that lexically induced delay effects did not emerge in previous experiments because the lexically consistent phonemes were not active enough for lexical feedback to be able to extend the competition that would delay phoneme recognition. In other words, the acoustic information regarding phoneme identity was too strong to allow lexical influences. One previous experiment was designed to test acoustically similar phonemes, but this experiment also failed to show inhibitory lexical effects (Experiment 4 in Frauenfelder et al., 1990). Our simulations indicate that indirect lexical effects are very sensitive to lexical neighborhood structure, but in previous experiments the critical target-bearing items had differing onsets and thus, differing lexical neighborhoods. In addition, previous studies have conflated lexical status of the target-bearing item with the experimental manipulation, which may obscure lexical delay effects if phoneme targets embedded in nonwords are harder to detect than phoneme targets in words (beyond a simple lack of lexical facilitation). Thus, the TRACE model was found to be consistent with previous behavioral results and suggested that a more balanced control condition is required.

A simulation informed by these results predicted that a lexically induced delay in phoneme recognition could be observed in the comparison of time to detect two different kinds of lexically inconsistent word-final phonemes. Lexically inconsistent phonemes that were acoustically similar to the lexically consistent phoneme they replaced were recognized more slowly by TRACE than phonemes that were very different from the lexically consistent phoneme. In the similar condition, the lexically consistent phoneme was partially active because the acoustic input was partly consistent with it. Lexical feedback increased the activation of the lexically consistent phoneme, thus making it a stronger competitor and extending the number of processing cycles required for the acoustically present phoneme to win the competition and reach the response threshold. In the dissimilar condition, the lexically consistent phoneme was not active because it was not consistent with the acoustic input. As in the similar condition, lexical feedback provided excitatory input to the lexically consistent phoneme, but lack of other support and competition from the acoustically presented phoneme kept the activation of the lexically consistent phoneme below threshold, so that it did not inhibit the acoustically presented phoneme. Thus, the lexical feedback did not delay phoneme recognition.

Two behavioral experiments were carried out to test this prediction with human participants. The critical

comparison was between response times to detect phoneme targets when they were similar to the lexically consistent phoneme (e.g., "apprenti<u>sh</u>") or very different from the lexically consistent phoneme (e.g., "academi<u>sh</u>"). Like the TRACE model, human participants were slower to detect phonemes when they were similar to the lexically consistent phoneme than when they were dissimilar. Although the strength of this effect varied across experiments, the basic pattern of results was found for all four phoneme targets tested in response times and for three of four targets in error rates. The variability in effect size could be due to loss of acoustic information due to recording and filtering procedures.

It is important to note that the autonomous model Merge (Norris et al., 2000) may also be consistent with the delay effect observed in these experiments. However, as with Merge's account of other lexical effects, the autonomous account differs in that the effect comes from information integration and competition at the decision layer. There are no inhibitory lexical-to-decision connections and Merge has a bottom-up priority rule that stipulates that a phoneme must have bottom-up support before entering into competition at the decision layer. Partial support provided by a similar phoneme may be enough to satisfy this requirement; thus, lexically consistent phonemes similar to the presented phoneme would compete with the presented phoneme and produce the same pattern of results as TRACE. Therefore, the data in this report do not provide a direct challenge to the Merge model. Rather, the results presented here answer one of the challenges that have led some researchers to reject the interactive view of speech perception. Nonetheless, we prefer the interactive view because we believe that it provides a more parsimonious account of the results. The TRACE model does not rely on an arbitrary bottom-up priority rule, but rather on the general principle that resting activation of units is below their interactive threshold. This generality means that any raising of phoneme units above their threshold (e.g., by priming) can open the door to lexically mediated delay effects.

In a recent critique of the interactive view, Norris et al. (2000) provided four reasons to prefer the autonomous view. (1) Norris et al. questioned evidence of predicted consequent effects of lexical feedback, perhaps the strongest evidence for interactive processing in speech. An example of this type of effect is the influence of a phoneme whose identity is affected by lexical constraints on identification of an adjacent phoneme in another word (lexically mediated compensation for coarticulation, Elman & McClelland, 1988). Norris et al. questioned the existence of this effect based on a possible confound with transitional probability and a failure to replicate the lexically mediated

compensation for coarticulation effect (Pitt & McQueen, 1998). (2) A second criticism was based on results of experiments testing lexical influences on sub-categorical mismatch effects (Marslen-Wilson & Warren, 1994; McQueen, Norris, & Cutler, 1999). In particular, Marslen-Wilson and Warren found that TRACE simulations did not match their behavioral results. Norris et al. also argued (3) that the lack of demonstrated inhibitory lexical effects contradicts the interactive view of speech perception and (4) that the TRACE model cannot address attentional modulation of lexical effects.

Three out of four of Norris et al.'s (2000) criticisms have been addressed. First, in their original report, Elman and McClelland (1988) showed that the lexically mediated compensation for coarticulation effect was found even when local phonetic context (and therefore transitional probability) was held constant (their Experiment 3). Furthermore, the key results have been replicated with controls for transitional probability (Magnuson et al., 2003a; Samuel & Pitt, 2003) and other such indirect effects have been demonstrated (Samuel, 1997, 2001). Note also that Samuel and Pitt discussed the complex interplay of stimulus factors that may undermine lexically mediated compensation for coarticulation, suggesting reasons why Pitt and McQueen's (1998) failure to replicate this effect may not be so surprising. It is important to note that in TRACE indirect effects are small; an acoustically unambiguous fricative will cause more robust phoneme unit activation than a lexically biased ambiguous fricative, a more active phoneme will cause stronger effects on its neighbors. As a result, it is quite reasonable that in some circumstances an unambiguous fricative will cause compensation for coarticulation, but a lexically defined ambiguous fricative will not (i.e., the results found by Pitt and McQueen). The critical finding is that there are conditions under which lexical feedback does cause pre-lexical effects (such as compensation for coarticulation), thus providing strong evidence for direct lexical feedback. McQueen (2003) has suggested that there may be longer-range transitional probability differences not fully controlled for in any of the experiments. Magnuson, McMurray, Tanenhaus, and Aslin (2003b) have done further analyses that do not support McQueen's criticism, but it is possible that some, as yet undetected, non-lexical effect is responsible for the results. In the absence of clear evidence of such an effect, the TRACE model appears to be consistent with the existing data.

Second, recent studies of lexical influences on sub-categorical mismatch effects combining eye-tracking and behavioral experiments and TRACE simulations have shown that TRACE is consistent with findings in this area (Dahan, Magnuson, Tanenhaus, & Hogan,

2001). The reasons for the discrepancy between the simulation results reported by Marslen-Wilson and Warren (1994) and Dahan et al. remain unclear. However, Dahan et al. found that their pattern of results persisted across a wide range of parameters and stimuli in two different implementations of the TRACE model (including the MacTRACE version used by Marslen-Wilson and Warren). Details regarding these attempts to replicate Marslen-Wilson and Warren's results are provided in Magnuson, Dahan, and Tanenhaus (2001). In consideration of the robustness of Dahan et al.'s findings, it seems safe to conclude that interactive processing is consistent with results of sub-categorical mismatch experiments. It may also be worth noting that the lexical enhancement of mismatch found by Marslen-Wilson and Warren, McQueen et al. (1999), and Dahan et al. can be interpreted as a type of lexical inhibition that arises under conditions of acoustic ambiguity. That is, when acoustic cues to phoneme identity conflict and the preceding context is consistent with a word the mismatch is greater than when the preceding context is a nonword.

Third, in the present simulations and experiments, we have shown that an interactive model of speech perception is consistent with previous studies of lexical inhibition, which were thought to be incompatible with direct lexical feedback. Analysis of the dynamics of interactions in the TRACE model explained why previous studies have failed to demonstrate lexically induced delays in phoneme recognition and further simulations predicted conditions that should produce such effects. Additionally, this pattern persists across a range of parameter settings, indicating that this is a robust prediction of interactive processing embodied in TRACE. Finally, our experiments tested these conditions and provided behavioral evidence that is consistent with the lexical delay effect predicted by TRACE.

Fourth, attentional modulation of lexical effects is an issue that remains to be addressed within the interactive-activation framework; this is a matter for future investigations to consider. In summary, three of four of the criticisms of TRACE that were raised by Norris et al. (2000) have been addressed. This is not to say that TRACE is proven correct, but only to say that it is consistent with the current evidence related to these criticisms.

On the helpfulness of lexical feedback

Critics of the interactive view have argued that lexical feedback cannot improve phoneme recognition and, in fact, can cause inaccuracies in phoneme identification (Norris et al., 2000). That is, Norris et al. argued that lexical feedback could cause "hallucination" of lexically consistent phonemes. Simulation 1 results showed that the interactive TRACE model does not hallucinate phonemes based solely on lexical feedback, because lexical feedback alone is not enough to activate a phoneme unit inconsistent with the acoustic input. However, as predicted by Simulation 2 and confirmed by the behavioral experiments, lexical feedback can delay recognition of lexically inconsistent phonemes due to competition with phonemes that are partially consistent with the acoustic input as well as the lexical constraints. Why would the cognitive system include a mechanism that can delay phoneme recognition? To answer this question it is helpful to consider interactive processing in terms of a rational Bayesian decision maker. It has been shown that stochastic versions of the TRACE model (McClelland, 1991) actually implement optimal Bayesian inference (Movellan & McClelland, 2001.) In an optimal Bayesian decision maker, the optimal policy is to combine information from different sources to assign posterior probabilities to possible interpretations of an input and then choose the alternative with the highest posterior probability. In a natural situation where real words are much more likely than nonwords, a policy of using lexical context as well as acoustic information can lead to a higher overall probability of correct phoneme recognition. If such a policy has become ingrained in the processing machinery, as it is in TRACE, then it may continue to be applied even in laboratory situations where an experimenter has chosen to create stimuli that are unlikely to occur in natural environments.

An important point in this discussion is the fact that lexical context is only one factor in the Bayesian decision; acoustic information is another. When the acoustic information overwhelmingly favors one alternative over another (i.e., the probability of the acoustic information given one of the alternatives is very small compared to the probability of the acoustic information given another alternative), then lexical information will have little effect. That is, if there is no chance that the acoustic information could have arisen from the pronunciation of a phoneme, then it will not be chosen even in the face of lexical information that favors it. It is only when the acoustic information is partially consistent with multiple interpretations, one of which is lexically appropriate, that a lexical influence will become apparent.

This analysis is consistent with the pattern seen in our TRACE simulations and behavioral experiments, where detection of a lexically inconsistent phoneme is slowed when it is similar to a lexically consistent phoneme, but not when it is completely inconsistent with it. Other phoneme perception findings are also consistent with this approach; for example, cases in which listeners fail to detect mispronunciations (Bond & Small, 1983; Cole, 1973; Cole, Jakimik, & Cooper, 1978; Marslen-Wilson & Welsh, 1978; Small & Bond, 1986). Of particular rele-

vance is the finding that multi-feature mispronunciations are detected at greater rates than single feature mispronunciations (Cole, 1973; Marslen-Wilson & Welsh, 1978). That is, listeners perceive the lexically consistent phoneme in place of the acoustically present one more frequently when the two phonemes are acoustically similar.

Similar illusory phoneme perceptions have been demonstrated in the phonemic restoration effect (Samuel, 1981a, 1996, 1997; Warren, 1970). As with our account of lexically induced delays in phoneme recognition, similarity to the expected phoneme plays a key role in the phonemic restoration effect. When part of an utterance is deleted, listeners perceive a silent gap. However, when the gap is filled with noise, listeners report that the utterance is intact, thus perceiving an illusory phoneme. This kind of phonemic restoration is even stronger when the noise filling the gap is acoustically similar to the absent phoneme (Samuel, 1981b). In Bayesian terms, listeners are most likely to restore a missing phoneme when there is some acoustic evidence for the presence of the phoneme because both the conditional probability of the acoustic information and of the context affect the inference.

TRACE differs from autonomous models (including Massaro, 1998 Fuzzy Logical Model of Perception, as well as Merge) in feeding back lexical information to the phoneme level rather than combining lexical and phonological information at a later decision stage. Both approaches can lead to the same degree of adherence to optimal policies of Bayesian inference (Movellan & McClelland, 2001). However, proponents of the autonomous view have claimed that a "theoretical argument against perceptual feedback is that it cannot benefit word recognition, and therefore has no function to serve in normal listening" (McQueen, Norris, & Cutler, 2003, p. 267; see also Norris et al., 2000). This statement appears at first glance to be consistent with Movellan and McClelland's finding that both feed-forward and feedback approaches can lead to optimal identification policies. However, the statement neglects the fact that decisions about the identity of one phoneme have implications for the identification of other phonemes and hence, other words. For example, the acoustic realization of a stop consonant (e.g., /k/ or /t/) is affected by the presence of a prior fricative (e.g., /s/ or /ʃ/; Mann & Repp, 1981). In this case, using lexical information to determine whether an ambiguous fricative is a /s/ or a /ʃ/ could influence identification of a subsequent /k/ or /t/. This is exactly what feedback accomplishes in the TRACE model and exactly what is found in experiments; as discussed above, lexical information can mediate compensation for coarticulation (Elman & McClelland, 1988; Magnuson et al., 2003a, 2003b; Samuel & Pitt, 2003)

and similar indirect lexical effects have also been demonstrated by Samuel (1997, 2001). For example, consider the case when the auditory input is "fooli[X] [?]apes" where [X] is an ambiguous fricative (/s/ or /ʃ/) and [?] is an ambiguous stop (/t/ or /k/). Norris and colleagues are quite correct in saying that feedback has no particular advantage in identifying the word "foolish" since this can be accomplished with equal fidelity to Bayesian principles with or without feedback. However, "tapes" and "capes" are equally good candidates for the second word given that /t/ and /k/ are equally consistent with the input. Preceding phonetic context can indicate whether /t/ or /k/ is more likely to be correct, but only if the lexical information is used to disambiguate it. The tendency for the perception of the ambiguous stop to be perceived as /t/ in the example above is evidence that the lexically influenced identification of /ʃ/ is not simply a decision made at the output of perceptual processing but feeds back into the phonemic processing. While such influence of one decision on another can be arranged in other architectures, the interactive architecture of TRACE intrinsically predicts such indirect effects.

Based on the above, the evidence appears to be consistent with our view that speech perception is an interactive process. Indeed, we would suggest that interactive processing as embodied in TRACE is a domain-general principle of perception and cognition. As such, it fits well with the perception of illusory contours in the visual domain (e.g., Kanizsa, 1979). That is, the Kanizsa triangle and similar visual illusions are an interpretation of the visual input that may be consistent with previous experience, but is "incorrect" in a laboratory setting. Electrophysiological studies in monkeys have provided evidence that illusory contours arise from feedback interactions (Lee & Nguyen, 2001), consistent with our view that interactive processing occurs in other domains.

In sum, interactive processing is a domain-general mechanism that brings multiple sources of knowledge together to recognize complex stimuli in a way that is consistent with the perceiver's experience. Although this mechanism can cause delays and errors, these detrimental effects are the inevitable consequences of optimal inference when applied to stimuli that conflict with prior knowledge.

Appendix A. Stimuli for simulations

The list of input patterns used in TRACE model simulations. For ease of replication, the inputs are presented in standard TRACE notation. IPA equivalents of TRACE phoneme symbols are presented in the key below.

A.1. Simulation 1: Past experiments

Words: b^b^l, k^p^l, parS^l, pas^bli, sut^b^l, ^br^pt, pak^t, sikr^t, targ^t, tr^sti
Matched Nonwords: p^b^l, g^p^l, barS^l, kas^bli, Sut^b^l, abr^pt, bak^t, Sikr^t, darg^t, pr^sti
Inhibitory Nonwords: b^b^t, k^p^t, sut^b^t, parS^t, pas^bti, ^br^pl, pak^l, targ^l, tr^sli, sikr^l
Control Nonwords: p^b^t, g^p^t, Sut^b^t, barS^t, kas^bti, abr^pl, bak^l, darg^l, pr^sli, Sikr^l
True Nonwords: pug^t, gib^t, Sakub^t, buls^t, kuSibti, apl^pl, big^l, dilk^l, plisli, Supr^l
Near Nonwords: pak^k, tr^ski, ^br^pk, sikr^k, targ^k, art^St, pal^Si, dikriS, praduS, pragr^S
Control Near Nonwords: tak^k, br^ski, ibr^pk, Sikr^k, parg^k, urt^St, tal^Si, pikriS, traduS, bragr^S

A.2. Simulation 2: New prediction

Words: pak^t, tr^sti, ^br^pt, sikr^t, targ^t, art^st, pal^si, dikris, pradus, pragr^s
Near Nonwords: pak^k, tr^ski, ^br^pk, sikr^k, targ^k, art^St, pal^Si, dikriS, praduS, pragr^S
Distant Nonwords: pak^S, tr^sSi, ^br^pS, sikr^S, targ^S, art^kt, pal^ki, dikrik, praduk, pragr^k

A.3. Key

p=/p/, b=/b/, t=/t/, d=/d/, k=/k/, g=/g/, s=/s/, S=/ʃ/, r=/r/, l=/l/, a=/a/ or /ɑ/, i=/i/, u=/u/, ^=/ʌ/ or /ə/

Appendix B. TRACE simulation parameters

Maximum unit activation:	1.0
Minimum unit activation:	−0.3
Maximum word–word inhibition:	3.0
k for Luce choice equation:	15
Continua feature weights:	1.0 for all 7 continua
Rest activation:	−0.1 for all 3 levels
Feature level decay:	0.01
Phoneme level decay:	0.03
Word level decay:	0.05
Input–feature excitation:	1.0
Feature–phoneme excitation:	0.2
Phoneme–word excitation:	0.5
Word–phoneme excitation:	0.03
Feature–feature inhibition:	0.04
Phoneme–phoneme inhibition:	0.04
Word–word inhibition:	0.03
All other interaction parameters:	0
Print (output) frequency:	1 cycle

Appendix C. Stimuli for experiments

C.1. Experiment 1

The following is a list of base words for critical stimuli in Experiment 1. For each word, -/Is/ and -/Iʃ/ versions were created using standard endings (see text for details).

/ʃ/ words: abolish, anguish, blemish, british, danish, diminish, english, extinguish, foolish, goldfish, Jewish, punish, turkish, varnish
/s/words: actress, apprentice, crevice, endless, fortress, goddess, harness, necklace, notice, novice, pelvis, reckless, waitress, witness
/k/words: academic, automatic, clinic, fabric, frantic, garlic, gothic, heretic, lyric, panic, terrific, topic, traffic, tragic

Words were matched on four key variables: word frequency (/ʃ/: 32.1, /s/: 11.9, /k/: 20.2, $F = 1.072$, $p = .352$) and log-transformed word frequency (/ʃ/: 0.89, /s/: 0.80, /k/: 1.05, $F = 0.583$, $p = .563$), number of phonemes (/ʃ/: 6.1, /s/: 6.0, /k/: 6.2, $F = 0.137$, $p = 0.872$), number of syllables (/ʃ/: 2.2, /s/: 2.1, /k/: 2.4, $F = 1.647$, $p = .206$), and distance between target and uniqueness point (/ʃ/: 1.6,/s/: 1.9, /k/: 1.8, $F = 0.619$, $p = .544$).

C.2. Experiment 2

The following is a list of base words for critical stimuli in Experiment 2. For each word, -/Ik/ and -/It/ versions were created using standard endings (see text for details).

/t/ words: bracelet, deficit, deposit, elicit, faucet, forfeit, gadget, habit, implicit, limit, omelet, orbit, spirit, summit
/k/ words: arsenic, dynamic, epidemic, ethnic, fabric, frolic, garlic, gothic, lyric, mosaic, music, panic, prolific, specific
/ʃ/ words: abolish, admonish, anguish, blemish, danish, diminish, english, foolish, gibberish, goldfish, Jewish, lavish, punish, Spanish

Words were matched on four key variables: word frequency (/t/: 23.6, /k/: 31.6, /ʃ/: 25.6, $F = 0.084$, $p = 0.920$) and log-transformed word frequency (/t/: 0.912, /k/: 0.991, /ʃ/: 0.812, $F = 0.255$, $p = 0.776$), number of phonemes (/t/: 5.86, /k/: 6.29, /ʃ/: 5.86, $F = 0.791$, $p = 0.461$), number of syllables (/t/: 2.29, /k/: 2.50, /ʃ/: 2.29, $F = 0.745$, $p = 0.481$), and distance between target and uniqueness point (/t/: 1.79, /k/: 1.64, /ʃ/: 1.64, $F = 0.291$, $p = 0.749$).

Appendix D. Target-based analyses

Effect sizes and statistical tests by target (significant effects are in bold, marginal effects are marked by *)

			W-NNW	W-DNW	NNW-DNW
RT	E1	/s/	74 ms; **t(19) = 2.979, p = .008**	58 ms; t(19) = 1.778, p = .091*	16 ms; t(19) = .371, p = .715
		/ʃ/	121 ms; **t(20) = 3.412, p = .003**	71 ms; **t(20) = 3.108, p = .006**	50 ms; t(20) = 2.033, p = .056*
	E2	/k/	177 ms; **t(24) = 4.826, p < .001**	109 ms; **t(24) = 5.739, p < .001**	68 ms; t(24) = 1.916, p = .067*
		/t/	123 ms; **t(24) = 3.689, p = .001**	81 ms; **t(24) = 2.701, p = .012**	42 ms; t(24) = 1.434, p = .164
Error	E1	/s/	1.4%; t(19) = .567, p = .577	5.0%; t(19) = 1.926, p = .069*	3.6%; t(19) = 1.045, p = .309
		/ʃ/	2.0%; t(20) = 1.142, p = .267	0.7%; t(20) = .252, p = .803	1.3%; t(20) = .491, p = .629
	E2	/k/	2.3%; t(24) = .625, p = .538	1.1%; t(24) = .401, p = .692	3.4%; t(24) = 1.238, p = .228
		/t/	10%; **t(24) = 3.392, p = .002**	6.0%; **t(24) = 2.529, p = .018**	4.0%; t(24) = 1.371, p = .183

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Bond, Z. S., & Small, L. H. (1983). Voicing, vowel, and stress mispronunciations in continuous speech. *Perception & Psychophysics, 34*, 470–474.

Cole, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics, 1*, 153–156.

Cole, R. A., Jakimik, J., & Cooper, W. E. (1978). Perceptibility of phonetic features in fluent speech. *Journal of the Acoustical Society of America, 64*, 44–56.

Connine, C. M., Titone, D., Deelman, T., & Blasko, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language, 37*, 463–480.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16*(5/6), 507–534.

Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America, 27*, 761–764.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language, 27*, 143–165.

Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory?. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 77–91.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 110–125.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kucera, H., & Francis, W. N. (1967). *Computational Analysis of Present-Day American English*. Providence: Brown University Press.

Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in early visual cortex. *Proceedings of the National Academy of Sciences, 98*, 1907–1977.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358–368.

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes, 16*, 565–581.

Luce, R. D. (1959). *Individual choice behavior*. Oxford, England: Wiley.

Magnuson, J. S., Dahan, D., & Tanenhaus, M. K. (2001). On the interpretation of computational models: The case of TRACE. In J.S. Magnuson & K.M. Crosswhite (Eds.), *University of Rochester Working Papers in the Language Sciences, 2,* pp. 71–91.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003a). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science, 27*, 285–298.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003b). Lexical effects on compensation for coarticulation: A tale of two systems?. *Cognitive Science, 27*, 801–805.

Mann, V. A., & Repp, B. H. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America, 69*, 546–558.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29–63.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*(4), 653–675.

Massaro, D. W., & Cohen, M. M. (1983). Phonological context in speech perception. *Perception & Psychophysics, 34*, 338–348.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. Cambridge, MA, USA: The MIT Press.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–86.

McClelland, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology, 23*(1), 1–44.

McQueen, J. M. (2003). The ghost of Christmas future: didn't Scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus, and Aslin (2003). *Cognitive Science, 27*, 795–799.

McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision making: Evidence from sub-categorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25*(5), 1363–1389.

McQueen, J. M., Norris, D., & Cutler, A. (2003). Flow of information in the spoken word recognition system. *Speech Communication, 41*, 257–270.

Movellan, J. R., & McClelland, J. L. (2001). The Morton-Massaro law of information integration: Implications for models of perception. *Psychological Review, 108*(1), 113–148.

Newman, R. S., Sawusch, J. R., & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 873–889.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*, 299–370.

Peeters, G., Frauenfelder, U. H., & Wittenburg, P. (1989). Psychological constraints upon connectionist models of word recognition: Exploring TRACE and alternatives. In R. Pfeifer, Z. Schreter, F. Fogelman-Soulie, & L. Steels (Eds.), *Connectionism in perspective* (pp. 395–402). North-Holland: Elsevier.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon?. *Journal of Memory and Language, 39*(3), 347–370.

Rubin, P., Turvey, M. T., & Van Gelder, P. (1976). Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception & Psychophysics, 19*, 394–398.

Samuel, A. G. (1981a). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110*, 474–494.

Samuel, A. G. (1981b). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1124–1131.

Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes?. *Journal of Experimental Psychology: General, 125*, 28–51.

Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology, 32*, 97–127.

Samuel, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science, 12*, 348–351.

Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language, 48*, 416–434.

Small, L. H., & Bond, Z. S. (1986). Distortions and deletions: Word-initial consonant specificity in fluent speech. *Perception & Psychophysics, 40*, 20–26.

Tyler, L. K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics, 34*, 409–420.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*, 550–592.

Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sublexical and Lexical Representations in Speech Production: Effects of Phonotactic Probability and Onset Density. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 514–529.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*, 392–393.

Wilson, M. D. (1988). MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments, and Computers, 20*, 6–11.

Wurm, L. H., & Samuel, A. G. (1997). Lexical inhibition and attentional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory & Language, 36*, 165–187.