

The prolonged influence of subsequent context on spoken word recognition

Christine M. Szostak · Mark A. Pitt

© Psychonomic Society, Inc. 2013

Abstract Connine, Blasko, and Hall (Journal of Memory and Language 30:234–250, 1991) suggested that within a 1-second temporal window, subsequent biasing information can influence the identification of a previously spoken word. Four experiments further explored this hypothesis. Our participants heard sentences in which an ambiguous target word was followed less than or more than a second later by a word biased in favor of either the target word or another word. Overall, the effects of the contextual biases on responding, measured using phonemic restoration and phoneme identification, were almost as large after 1 second as before 1 second. The implications of these results for defining the window of contextual effects are discussed.

Keywords Spoken word recognition · Temporal window · Subsequent context effects · Phoneme restoration · Phoneme identification · Psycholinguistics · Speech perception

When hearing running speech, a listener must rapidly transform the continuous acoustic signal into the words that the talker intended. This task is complicated by the fact that the quality of the speech signal can vary in a number of ways, such as when the articulation varies in clarity or environmental noise masks the signal. Despite these challenges, listeners typically have little difficulty identifying the intended words. How is this accomplished?

One means of achieving successful word recognition is to utilize information in the surrounding sentential context. For example, in the utterance *The feathers were on the = ing*, where the = represents noise, *feathers* can help aid identification of the masked phoneme, since only if the masked

phoneme were to be identified as /w/ would the sentence make sense. Ample examples have shown preceding context being used in this way (e.g., Borsky, Tuller, & Shapiro, 1998; Cole, Jakimik, & Cooper, 1980; Gaskell & Marslen-Wilson, 2001; Marslen-Wilson & Welsh, 1978; Samuel, 1981; van Alphen & McQueen, 2001). For example, Samuel (1981) demonstrated that the presence of a semantically biasing word caused listeners to perceptually restore a phoneme that had been intermixed with or replaced by noise (e.g., /b/ in *ball*) to yield a semantically coherent sentence (e.g., *The boy swung his bat at the ball*). Likewise, Marslen-Wilson and Welsh (1978) demonstrated that a highly constraining preceding context, as in *He still wanted to smoke a...*, could cause listeners to misperceive a pseudoword such as *shigarette* as the word *cigarette* more often than when the pseudoword was preceded by a low-constraining prior context, as in *He noticed there was a shigarette*.

Far less is known about what influence biasing information has on the identification of a preceding word (e.g., *The = ing had feathers*). It is possible that in order to maximize the speed with which word recognition occurs, if no prior influential context is available, the processor might delay word identification for a short period of time. If no further biasing information is perceived within this window of time, the processor simply commits to a lexical representation (right or wrong) on the basis of any available information.

Findings from Connine, Blasko, and Hall (1991; see also Samuel, unpublished manuscript) support this hypothesis. Specifically, they suggested that during an approximately 1-s temporal window, the processor will delay word identification to wait for additional contextual information to become available. To test this hypothesis, they presented listeners with sentences in which a target word had an onset ranging from a clear /t/ to a clear /d/, followed by a subsequent word that was biased toward one endpoint of the continuum (e.g., *After the t/dent in the campgrounds collapsed, we went to a hotel*, where the underlined portions represent the target phoneme

C. M. Szostak (✉) · M. A. Pitt
Department of Psychology, Ohio State University,
1835 Neil Avenue, 281 Psychology Building,
Columbus, OH 43210-1222, USA
e-mail: szostak.1@osu.edu

and biasing word, respectively). They manipulated the distance between the target word and biasing word to determine whether there is a point in time after which subsequent biasing information no longer influences word recognition. In each sentence, the biasing word (e.g., *campground*) was presented three syllables (near condition) or six to eight syllables (far condition) after the target word (*_ent*). Participants were required to judge whether the word-initial phoneme was /d/ or /t/, and whether the sentence made sense. Connine et al. predicted that if the contextual influence on phoneme identification decreases as the distance between the target word and subsequent biasing word is increased, a reduced proportion of contextually congruent responses (e.g., responding “t” given *campground*) should be reported in the far as compared to the near condition.

Their findings supported this prediction, with the percentage of congruent responses decreasing in the far condition (approximately 57 % of responses) as compared with the near condition (approximately 61 % of responses). Because in the far condition the reaction times (RTs) fell slightly beyond 1 s, Connine et al. (1991) hypothesized that within a temporal window of approximately 1 s, the processor will wait before word identification is completed if subsequent biasing information has not become available.

In their study, approximately 84 % of the responses were given prior to the onset of the subsequent biasing word (e.g., *campground*) in the far condition. Given that nearly all participants responded early, strong evidence supported their conclusion. Their findings lead to another interesting question: Is it ever possible for the processor to delay resolution for more than a second? In other words, it is possible that even when the subsequent biasing information occurs more than 1 s after the word being identified has been perceived, such information may still have an influence on word identification. To directly test this hypothesis, it would be necessary to measure performance in the far condition when the subsequent biasing word has been perceived.

The purpose of the present study was to revisit the length of the temporal window and to test whether biasing information beyond 1 s could influence word identification. As in Connine et al. (1991), we explored this question by presenting sentences wherein a biasing word became available within or more than 1 s after target word onset. If the 1-s temporal window hypothesis were to be maintained even when listeners attempt to wait for an extended period, the biasing word should have no influence on performance when it occurs outside of the window. In contrast, if subsequent biasing information can influence identification of the target word beyond this time point, the biasing word should influence performance similarly, regardless of whether it occurs within or outside of the temporal window. This outcome would suggest that even after multiple intervening words, which together span more than 1 s, a word’s representation is maintained in memory and is still pliable, being subject to semantic influences.

Experiment 1

We evaluated whether the temporal window can extend beyond 1 s by presenting participants with sentences containing a target word (e.g., *wing*) followed by a biasing word (e.g., *feathers*) that occurred within or more than 1 s after target word onset. By requiring participants to wait to respond to the target word until after biasing word offset, we ensured that subsequent biasing information would have been heard before responses were made.

If listeners must wait for biasing information to become available, and if the processor can not delay commitment beyond 1 s, then the biasing word should not influence responding when it occurs after 1 s, even when the biasing word becomes available before responses are made. In contrast, if the processor is not bound by a 1-s window and delays commitment to a single lexical item beyond this time period, the biasing word should bias responding regardless of whether it occurs before or after the 1-s window.

We designed our conditions to be similar to those of Connine et al. (1991). Sentences were constructed in which a semantically biasing word was presented either one syllable (near condition) or six to eight syllables (far condition) subsequent to a monosyllabic target word containing a noise-disrupted onset phoneme (e.g., *The = ing had feathers* vs. *The = ing had an exquisite set of feathers*). In line with Connine et al., monosyllables were used as the target words. The biasing word was either congruently related to the target word (congruent condition: *The wing had feathers*) or incongruently related to the target word but congruently related to a competitor word sharing all but the onset of the target word (incongruent condition: *The wing had diamonds*, where *ring* is the congruent competitor).

We chose to use the phoneme restoration paradigm (Samuel, 1981; Warren, 1970) as it lends itself to exploring the interaction between acoustic input and subsequent biasing information on word identification, and it allows for flexibility in stimulus selection. Further, it has been used in other studies exploring subsequent biasing context (e.g., Sherman, 1971; Warren & Sherman, 1974; see also Samuel, unpublished manuscript). Unlike phoneme identification, phoneme restoration is not restricted to the use of minimal pairs (e.g., *dent* vs. *tent*). In phonemic restoration, listeners hear stimuli in which a phoneme (e.g., /w/ in *wing*) has noise added to it or is replaced by noise (e.g., ~~w~~ing vs. =ing), and are required to make an “added”/“replaced” response. Because the added stimulus version (~~w~~ing) provides some acoustic evidence of the phoneme (e.g., formant transitions), use of this paradigm allows for an examination of the influence of subsequent context when signal-level information is also present, a situation similar to that of the phoneme identification paradigm used by Connine et al. (1991). The added stimulus versions make it possible to

examine what influence the subsequent biasing word has on word identification when the lexical hypothesis is generated. The replaced stimuli provide a case in which the biasing word’s influence on target word identification can be examined when there is no evidence of the onset phoneme (i.e., the lexical identity is ambiguous). If no acoustic-phonetic information is available that can provide cues to phoneme identification, then any “added” responses are due solely to contextual restoration of the missing phoneme.

Combining the two variables (distance between target and biasing word, congruency between target and biasing word) with added and replaced versions of the target word yielded eight conditions. They are listed in Table 1, along with example sentences. Predictions across the conditions differ as a function of whether sentences contained the added or replaced target phoneme. According to Warren (1970; see also Samuel, 1981), if the biasing word influences word identification, the likelihood of perceiving the phoneme as “added” should increase as compared to a case in which the biasing word has no influence.

We focus first on the added-stimulus versions. If there is a 1-s temporal window after which the system will commit to a lexical representation, the following should be observed. In the near condition (top two rows of the table), listeners would show a stronger bias toward responding “added” when the biasing word is congruent with the target word than when the two words (target and biasing) are incongruent with one another. This would happen because the processor will initially activate a lexical representation based on the available acoustic-phonetic information in the signal (/w/). When the subsequent biasing information is congruently associated with the identified word, this should confirm that the proper target word has been identified and trigger an “added” response. In contrast, when the target and biasing words are incongruent, a mismatch will be encountered, which will cause the processor to reject the word that was identified and thus trigger a “replaced” response. In the far condition (upper two middle rows of the table), no difference across the two congruency conditions should be observed. In contrast, if the temporal window can be extended beyond 1 s, we

would expect that the data pattern in the far condition would be nearly identical to the data pattern in the near condition.

For the replaced stimuli, we predicted a different outcome. If the 1-s window can not be extended, in the near condition (lower two middle rows in the table) a high proportion of “added” responses should be observed regardless of congruency condition. This is because when the phoneme is completely replaced by noise, the rhyme (e.g., *ing*) has a viable onset that would make the word congruent with *feathers* and *diamonds* (e.g., /w/ and /r/, respectively). In contrast, in the far condition (i.e., beyond 1 s), participants should be less biased toward responding *added* owing to the fact that the processor should have committed to a particular lexical representation. On the other hand, if the temporal window can be extended beyond 1 s, performance in the near and far conditions should be similar to one another.

Method

Participants A total of 30 undergraduates received course credit in an introductory psychology class for their participation. They reported no hearing disorders or impairments, and all were native speakers of American English.

Stimuli The stimuli included 24 monosyllabic target words with either a consonant–vowel–consonant (CVC) or CVCC structure. Each word had at least two additional rhyme competitors (e.g., *ring* and *king*; mean number of rhyme competitors = 7.8, range = 3–12). None of the target words contained a rhyme that by itself was a word (e.g., *at* in *cat*).

Samuel (1981) found that when presented in isolation, words having plosive or fricative onsets (e.g., *sing* or *ding*) showed poor discriminability, whereas vowels, liquids, glides, and nasals showed better discriminability. Target words therefore were selected only if they had a liquid, glide, or nasal onset. Rhyme competitors of the target words however were not subjected to this constraint owing to the limited number of such words in English. Because of this limitation, participants

Table 1 Conditions and sample sentences used in Experiments 1–3

Noise	Distance	Congruency	Example
added	near	congruent	<i>The wing had feathers</i>
added	near	incongruent	<i>The wing had diamonds</i>
added	far	congruent	<i>The wing had an exquisite set of feathers</i>
added	far	incongruent	<i>The wing had an exquisite set of diamonds</i>
replaced	near	congruent	<i>The =ing had feathers</i>
replaced	near	incongruent	<i>The =ing had diamonds</i>
replaced	far	congruent	<i>The =ing had an exquisite set of feathers</i>
replaced	far	incongruent	<i>The =ing had an exquisite set of diamonds</i>

added = added stimulus version, in which noise was added to the target phoneme; replaced = replaced stimulus version, in which noise replaced the target phoneme.

were presented with only the target word; the competitor word was never presented. The above constraints made it impossible to hold constant a number of properties of the target words, such as the number of rhyme competitors and frequency of occurrence of the target word (e.g., *wing*).¹

The target word always occurred as the second word of the sentence. This decision was made on the basis of findings from Warren and Sherman (1974), who found that in phoneme restoration, listeners often displaced the target phoneme by several phonemes when identifying it in a transcribed sentence. Because we wanted to ensure that participants perceived the intended phoneme, we placed this phoneme in an easily identified location. Furthermore, this allowed us to direct participants' attention to the target word. This differed from Connine et al. (1991), who presented their target word as the fifth word in the sentence.

Each sentence contained a biasing word that was semantically congruent with the target word (congruent condition: *The wing had feathers*) or congruent with a rhyme competitor of the target word (incongruent condition: *The wing had diamonds*). The semantically biasing word was always sentence final. Two versions of each of the two sentence types (congruent and incongruent) were constructed, yielding a total of 96 sentences. The near condition contained sentences in which the biasing word occurred one syllable (mean = 217 ms) after the target word (e.g., *The wing had feathers*). The far condition contained sentences in which the biasing word was presented six to eight syllables (mean = 1,113 ms) after target word offset (i.e., *The wing had an exquisite set of feathers*). All of the sentences began with "the" followed by the target word and ended with the biasing word. This sentence construction method was chosen to keep sentences short while at the same time remaining sufficiently flexible to accommodate the manipulation of distance. This sentence structure differed from that of Connine et al. (1991; see also Samuel, unpublished manuscript), who used embedded clauses. The reason that we chose to use this more simplistic syntactic structure, was that we wanted to provide a listening environment that would be more likely to be conducive to late commitment. In other words, it is possible that the more complex syntactic structure used by Connine et al. required more processing resources and thus early commitment occurred as a way to help reduce the demand placed on processing. Because we wanted to encourage delayed commitment if such commitment could be achieved, use of a more complex syntactic structure was less desirable in the present set of experiments.

For each sentence quad (near congruent, near incongruent, far congruent, far incongruent), the biasing word in the congruent

and incongruent conditions always shared the same part of speech. A rating experiment demonstrated that the biasing words (e.g., *feathers*, *diamonds*) biased listeners to expect only the related target word, not its competitor (e.g., *feathers* biased expectations of *wing*, not *ring*).² Appendix 1 contains the full set of stimuli used in this and the third and fourth experiments.

For the main experiment, the sentences were spoken in a slow, clear voice in a sound-dampened room by C.M.S., a native female speaker of American English with an upper central Ohio accent. In line with the stimuli reported by Samuel (unpublished manuscript) they were recorded with reduced prosody. The stimuli were recorded at 48000 Hz with 16-bit resolution and were then down sampled to 22050 Hz and saved as individual wave files. One member of each of the 24 near-condition sentence pairs had its biasing word digitally spliced off of the precursor using both visual and auditory cues. A second copy of the precursor was made, and the congruent (*feathers*) and incongruent (*diamonds*) biasing words from other recordings were spliced into the two precursors. This same cross-splicing procedure was performed to create the far-condition stimuli. The splice technique was used to ensure that information in the sentence fragment, including the target word, was identical across congruency conditions.

Added and replaced versions of the sentences (yielding a total of 192 stimuli) were constructed in the same manner as was performed by Samuel (1987). The onset and offset of the target phoneme (e.g., /w/ in *wing*) in each cross-spliced stimulus was identified, and then added and replaced versions were created of each sentence (see Samuel, 1987, for details).

Procedure As is typical in restoration experiments, the experimental design was completely within subjects. To ensure the largest possible interval between repetitions of a given target word and its biasing word, stimuli were separated into eight blocks of 24 trials, with each target word occurring in the same pseudorandomized position in each block. The participants were told nothing about the blocks used in the experiment. Because participants would hear two repetitions of each sentence, (e.g., an added and replaced version), Blocks 5–

² In a pencil-and-paper task, 30 participants read sentences with the biasing word present or absent. Each sentence had a blank line in place of the target word (e.g., present condition = *The _____ had feathers.*, absent condition = *The _____ had.*). The target and competitor words were presented below the sentence. Using a Likert-type 5-point scale, from 5 = *very good fit* to 1 = *very poor fit*, participants judged the goodness-of-fit of both words. Each participant saw only one of the six members of the sentence set [near congruent (e.g., *The _____ had feathers*), near incongruent (e.g., *The _____ had diamonds*), near absent (e.g., *The _____ had*) condition, or one of their three far counterparts]. Planned comparisons on the averaged rating data showed that across distances, ratings were reliably higher in the congruent (near = 4.7, far = 4.8) than in the absent (near = 3.7, far = 3.3) condition [near, $t(59) = 11.5, p < .001$; far, $t(59) = 15.2, p < .001$], whereas ratings were reliably lower in the incongruent (near = 0.7, far = 0.6) than in the absent [near, $t(59) = 63.6, p < .001$; far, $t(59) = 53.9, p < .001$] condition.

¹ The congruency hit and false alarm rate effects (congruent minus incongruent) and distance effects (near minus far) were calculated for each participant. Analyses indicated that lexical frequency, log frequency, and number of rhyme competitors did not reliably influence the sizes of the distance or congruency effects for hits or for false alarms.

8 had the identical ordering to that of Blocks 1–4, but contained the alternate member of the pair. For example, if the added version of *The *ing had feathers* was presented in the third position of Block 3, the replaced version of the same sentence was presented in the third position in Block 7.

Participants were tested in groups of one to four in sound attenuated rooms. They were told that they would hear sensible and nonsensical sentences over headphones and once the sentence was finished being presented the participants would make a judgment of whether the initial phoneme of the second word was present or absent by pressing one of two corresponding buttons on a response box labeled *added* and *replaced*. Participants were told that any responses made prior to sentence offset would not be recorded.

The stimuli were all presented diotically. Because added stimulus versions were expected to provide some information about the identity of the target phoneme, participants might notice that some of the sentences were incongruent (e.g., *The wing had diamonds*). The congruency manipulation was therefore pointed out to participants to avoid surprise when incongruent sentences were encountered.

At the start of each trial, a sentence was presented binaurally over headphones. Participants were given 2,500 ms after sentence offset to make an “added”/“replaced” response. There was then a 1,500 ms pause before the computer moved on to the next trial.³ A short break was offered after Trial 96. There were 24 practice trials and the experiment lasted approximately 50 min.

Results and discussion

Although the data from the phoneme restoration paradigm lend themselves nicely to analysis using signal detection theory (cf., Macmillan & Creelman, 2005; Samuel, 1981), the raw measures of hit rate (proportion of “added” responses out of all added stimulus items, hits plus misses) and false

³ In addition to the buttonpress responses in Experiment 1, participants performed a second task in which they were given an additional 4 s to write on an answer sheet the letter corresponding to the phoneme that was covered up or replaced by noise. Our goal with this task was to learn what phoneme had been perceived, if any, and how acoustic and semantic information influenced phoneme reports. Participants proved heavily biased toward reporting the spoken word, regardless of the condition (e.g., *wing*). This suggests that listeners began learning what phoneme had been covered or replaced. For the replaced condition, this likely provides little interpretable information, as no differences occurred as a function of congruency (likely participants simply were showing a baseline criterion for “added” responses). In the added condition, this finding is noteworthy in that even though listeners reported “wing,” they reported the phoneme as being added less often in the incongruent condition, which suggests that they were realizing that a mismatch in the identified word and biasing information had occurred, and thus assumed that they had misheard the word. In this case, it would make sense that they would reject the lexical representation, and thus make a “replaced” response.

alarm rates (proportion of “added” responses out of all replaced stimulus items, false alarms plus correct rejections) provide more insight into the questions of interest than do the derived measures of discriminability (d') and bias (beta). In particular, analysis of hit and false alarm rates allow for an exploration of differences in word processing when acoustic information conveying the identity of the phoneme is present in the signal (added version) or when it is absent (replaced version). For the interested reader, d' and beta results can be found in Table 2.

Figure 1 provides mean hit rates (left set of bars) and false alarm rates (right set of bars) across the distance and congruency conditions. An examination of the mean hit rates shows a clear effect of congruency (congruent minus incongruent condition) in the near condition (.32). Importantly, in the far condition, a similar congruency effect emerged (effect = .28). A two-factor analysis of variance (ANOVA) with Distance and Congruency as the two factors yielded reliable main effects of congruency, $F(1, 29) = 38.2, p < .001$, and distance, $F(1, 29) = 4.3, p < .05$. The interaction was not reliable, indicating that increasing the distance lowered hit-rate as a whole rather than shrinking the size of the congruency effect in the far condition. A distance by congruency interaction, with a large congruency effect emerging in the near condition and no congruency effect emerging in the far condition, should have been found if early commitment to a lexical representation occurred during a 1-s temporal window.

Inspection of the false alarm rates across the four conditions in Fig. 1 shows a very different outcome. In addition to the false alarm rates falling much lower than the hit rates across all conditions, the false alarm rates differed little across conditions (near congruency effect = .03, far congruency effect = .05), making the congruency effects minuscule, regardless of whether the biasing word was near or far from the target word. A two-factor ANOVA yielded both a marginally reliable distance main effect, $F(1, 29) = 2.7, p < .12$, and a marginally reliable congruency main effect, $F(1, 29) = 2.9, p < .06$; the interaction was not reliable. The marginal congruency effect may be the result of listeners building expectations about the identity of the target word in later blocks given the fact that the competitor (e.g., *ring*) was never presented. The effect could also be due to cues such as target phoneme duration or amplitude shape influencing phoneme

Table 2 Experiment 1: d' and beta values

Distance	Congruency	d'	beta
near	congruent	1.93	0.65
near	incongruent	0.90	1.13
far	congruent	1.60	0.81
far	incongruent	0.80	0.94

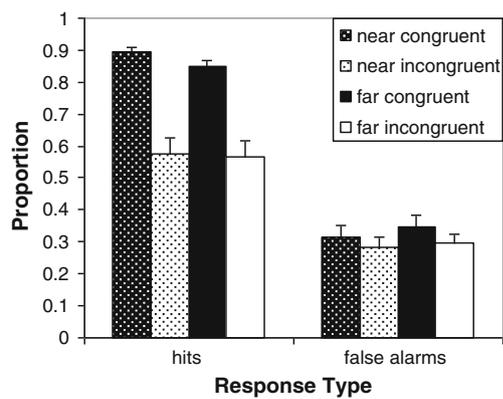


Fig. 1 Mean hit rates (“added” responses to noise-added stimuli) and false alarm rates (“added” responses to noise-replaced stimuli) as a function of congruency and distance in Experiment 1. Error bars represent one standard error of the mean estimate

identification given that the shape of the amplitude was maintained in both the added and replaced stimulus versions (cf., Bashford, Warren, & Brown, 1996; Sherman, 1971). To the extent this occurred, the magnitude of these differences suggests that such cues are not very discriminating for listeners.

The results of Experiment 1 suggest that when listeners are required to delay responding until the biasing word has become available, the temporal window may be extended. At a distance beyond 1 s, a subsequent biasing word can maintain a strong influence on responding, similar to that observed within 1 s. The small but reliable distance effect suggests that some decay of the trace of the target word took place. Over time, as the memory trace of the word decays, “added” responses will drop overall. The idea here is that because the trace is decaying, the memory for any perceived acoustic features is fading and thus, listeners become less confident that they heard anything along with the noise but this should have no specific impact on the size of the congruency effect. This is exactly the pattern that was found in the data (a main effect of distance without any interaction between this variable and congruency). The fact that the false alarms show the reverse distance effect suggests that listeners are showing some degree of regression toward the mean.

Experiment 2

Although the findings from Experiment 1 provide support for the idea that the processor may delay word identification when the biasing word is perceived at a far distance, we used phoneme restoration, whereas Connine et al. (1991) used phoneme identification, a paradigm in which listeners identify which of two phonemes served as the onset to the target word. The memory trace of an ambiguous phoneme might be

far less malleable than a noise-covered phoneme. If this is true, the window of time during which following context could have influenced responding may have been much shorter in phoneme identification.

Alternatively, it could also be that the context effects observed in Experiment 1 differ from those reported by Connine et al. (1991) because listeners adopted a postperceptual response strategy that is not linguistic in nature. For example, participants might have consciously revised their added/replaced judgments about stimulus quality after hearing the biasing context rather than subsequent biasing information actually influencing target word processing online. In other words, it may be that the nonlinguistic responses made in Experiment 1 evoked some type of response strategy that was not linguistic in nature and thus caused our data patterns to differ from those reported by Connine et al.

The purpose of Experiment 2 was to address the above concerns. We reran Experiment 1 using phoneme identification. Participants were presented with sentences having the same structure as those of Experiment 1. The second word in each sentence fell along a continuum ranging from a clear *sip* to a clear *ship*. The biasing word was congruently associated with one endpoint (e.g., *The sip was sunk*). Participants judged whether the target phoneme was /s/ or /ʃ/. As with Experiment 1, participants were required to wait until sentence offset before responding. If task differences caused the discrepancy, a congruency effect should be found only in the near and not in the far condition.

Method

Participants A group of 20 new individuals from the same pool took part in Experiment 2, meeting the same criteria as those of Experiment 1.

Stimuli The stimuli were 16 sentence pairs. One member of each pair served in the near condition and the other served in the far condition. For eight of the sentence pairs, the biasing word was biased toward *sip* (/s/-biased condition: *The sip was drunk*) and for the other eight pairs it was biased toward *ship* (/ʃ/-biased condition: *The sip was sunk*). We chose to use a *sip-ship* continuum in place of a *dent-tent* continuum for two reasons: (1) Continua from /s/-/ʃ/ often result in large shifts in labeling (e.g., McQueen, 1991; Pitt & Samuel, 2006). Shifts in labeling found with /d/-/t/ continua tend to be smaller (Pitt & Samuel, 1993), which could make it difficult to detect contextual effects, let alone differences across the near and far biasing conditions. (2) Because the primary goal of this study was to determine whether there are ever times that the processor may delay resolution, it was important to provide an environment in which such findings would be most likely to emerge if they will ever in fact emerge. Appendix 2 contains a full list of the sentences used in this experiment.

It should be noted that this experiment might be more accurately thought of as a hybrid between Connine et al. (1991) and Experiment 1. Although the paradigm was identical, our sentences were syntactically simpler for the reasons noted earlier.

Recordings of the sentences were made using the procedure in Experiment 1. As with that experiment, the sentences were recorded in a clear, natural conversation style with reduced prosody. A clear token of the /s/ in *sip* was removed from one sentence and the /f/ in *ship* was removed from another sentence. The two phonemes were matched in length and amplitude.

A /s-/f/ continuum was constructed by digitally blending the tokens in varying proportions on a 21 step continuum from 0 % (clear /s/) to 100 % (clear /f/) such that each step increased or decreased the proportion /s/ by 10 %. Piloting on multiple steps of the continuum was performed to make a first-pass selection of ambiguous items. During the pilot, participants made forced-choice responses between the two fricatives within the near- and far-condition biasing sentences. Steps were selected as potential candidates if a clear biasing word effect (e.g., more /s/ responses given on a step in the *sip*-biased condition than in the *ship*-biased condition) was observed in the near condition. Following the pilot, the continuum was widened to 51 steps wherein steps were altered in 2 % increments. The same piloting procedure was repeated. This allowed for more fine-grained selections to be made. The same criteria were made to select the finalized steps that included Steps 14, 18, and 22, along with the two clear endpoints (Steps 1 and 51). Each step was then spliced into the 16 sentence pairs, for a total of 160 stimuli.

Procedure Except for a change in task, the procedure was the same as Experiment 1. Participants heard the sentences in a fixed pseudorandomized order across five blocks with each of the 16 sentences presented once per block. Endpoint steps were presented half as often to reduce the number of trials (i.e., sentence repetitions) in the experiment. The participants were instructed to decide whether the first “letter sound” in the second word was /s/ or /f/ and to respond by pressing the corresponding labeled button. They were told that for some sentences the /s-/f/ response would be relatively easy, whereas in others it would be difficult. They were told that they should use any information in the sentence to help make their decision (cf., Connine et al., 1991 for similar instructions). They were told, however, as in Experiment 1, that some of the sentences would make sense, and that others would not.

As in Experiment 1, participants were not permitted to make a response (/s/ or /f/ buttonpress) until sentence offset. Participants were given 2,500 ms to respond, after which there was a 1,500-ms pause before the computer moved on to the next trial. The experiment began with 16 practice trials, and the testing session lasted approximately 30 min.

Results and discussion

The proportions of /s/ responses were calculated across steps and distance conditions for each participant. These values were then averaged across participants and are plotted in Fig. 2. The influence of biasing context was measured by quantifying the degree to which each pair of functions, near and far, split apart in the ambiguous region of the continuum. This value was obtained by averaging over the three ambiguous steps and then calculating the difference across the two biasing conditions. If the discrepancy between Experiment 1 and the results of Connine et al. (1991) were due to task differences, the labeling functions should split apart in the ambiguous region of the continuum in the near condition, but not in the far condition.

Shifts in labeling are present in the near and far conditions, demonstrating that a biasing context beyond 1 s influenced phoneme categorization. The labeling shift in the far condition, however, was half the size of the one in the near condition (.29 vs. .14, respectively), indicating that the influence of the biasing word diminished with time. An ANOVA with Distance and Bias as the factors yielded a reliable main effect only of bias, $F(1, 19) = 47.4, p < .001$, as well as a reliable Distance \times Bias interaction, $F(1, 19) = 18.5, p < .001$. Importantly, planned comparisons showed that the simple effects of bias in the two distance conditions were reliable [near condition, $t(19) = 7.0, p < .001$; far condition, $t(19) = 5.0, p < .001$].

The shapes of the functions in Fig. 2 suggest that listeners did not employ a strategy in which they responded using context alone while ignoring the fricative. Had they done so, biasing effects as large as those found across the ambiguous steps should have been seen at the endpoints. Instead, the functions converge to identical or very similar values, which indicates that they were sensitive to the clarity of the fricative.

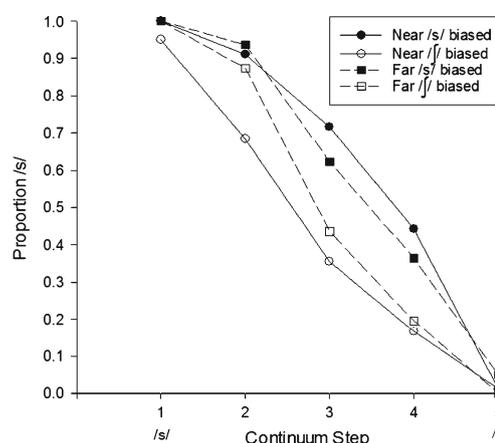


Fig. 2 Mean proportions of /s/ responses across the five steps as a function of biasing condition and distance

The findings of Experiment 2 suggest that the results from Experiment 1 were not a result of a change in paradigm. They also further supported the notion that beyond 1 s, if biasing information is perceived, it can help to resolve a lexical ambiguity. One reason that we may have obtained our effect in the far condition is that the memory trace of fricatives may be longer lasting than the memory trace for plosives. Thus, had we used plosives like Connine et al. (1991), the trace may have decayed so completely by onset of the target word in the far condition that no biasing effects could have been found. In other words, our data patterns may have been identical to their data patterns had we used a /d/-/t/ continuum.

At first glance, the data of Experiment 2 may appear somewhat at odds with those of Experiment 1. In Experiment 1, hits decreased across biasing contexts in a manner that suggests the incongruent context inhibited target word processing, whereas in the present experiment, the biasing context facilitated responding in a congruent fashion. The different outcomes are expected, however, if one considers the nature of the target-word stimuli. Recall that in Experiment 1, the added version of the target word (e.g., *ring*) contained acoustic information that resulted in the biasing word (e.g., *diamonds*) being incongruent with the target word. In contrast, in Experiment 2, when the word-initial fricative was unclear, it was also ambiguous. This made it possible to resolve the ambiguity so that it was always congruent with the biasing context (e.g., *?ip* could be biased to be heard as *sip* in the sentence *The ?ip was drunk*, or as *ship* in the sentence *The ?ip was sunk*). The phoneme identification data demonstrate a processing bias toward integration to achieve semantic coherence. The restoration results, in particular, those in the incongruent condition, identify a limiting condition of this process.

One additional question that needs to be addressed is why the effect of distance (far vs. near) was greater in Experiment 2 than in Experiment 1. In Experiment 1, the effect of congruency in the near versus the far condition was rather small (.32 vs. .28), whereas in Experiment 2, the near-condition effect (.29) was more than twice the size of the far-condition effect (.14). Statistical comparisons across the experiments, in which the size of the congruency effect was calculated for each participant (near minus far condition) was used as the dependent measure were performed. An ANOVA with Experiment (1 vs. 2) and Distance (near vs. far) as the factors yielded a main effect of distance, $F(1, 48) = 20.0, p < .001$, and an interaction between distance and experiment, $F(1, 48) = 11.0, p < .005$. The main effect of experiment was not reliable, $F < 1.9$. Planned comparison analyses indicated that the near condition was reliably different from the far condition in both experiments [Exp. 1, $t(29) = 1.7, p < .05$; Exp. 2, $t(19) = 4.3, p < .001$]. Importantly, the two experiments also differed reliably in the far condition, $t(48) = 2.2, p < .05$. There was no difference between the near

conditions across experiments, $t < 0.5$. One likely reason for the difference across experiments, is that as discussed earlier, those in Experiment 1 merely had to maintain some semblance of a signal, whereas those in Experiment 2 had to maintain a specific phoneme in memory over time. It is possible therefore that the memory trace for the specific features (e.g., the ability to distinguish differences in spectral energy of two different fricatives) has a somewhat faster decay rate than is the case for the perception of the presence of something more basic such as simply perceiving the presence of some signal in addition to noise.

Experiment 3

Although the results of Experiment 2 reinforce those of Experiment 1 in demonstrating that biasing influences can extend backward in time more than 1 s, one major difference between our work and that of Connine et al. (1991) is that we required our participants to respond after sentence offset, whereas in the previous study participants were allowed to respond at any time. It may be that our participants were engaged in a strategy wherein they were relying on the biasing word without really attending to the target word. In other words, to reduce their memory load, they may have been simply depending upon the biasing word while allowing the memory trace to fade from memory as a way to reduce the memory load. One piece of evidence disfavoring this idea is that a congruency effect was observed in the far condition of Experiment 1 (congruent minus incongruent). Recall that when listeners heard sentences such as *The ring had an exquisite set of diamonds*, they were less likely to report the target phoneme as being present. The explanation for this finding was that listeners were experiencing a mismatch in the acoustic and biasing information that led them to report the phoneme as absent. If they were ignoring the target word, both conditions should have shown relatively high or low hit rates. In other words, there should have been little to no congruency effect in this experiment.

The purpose of Experiment 3 was to provide a further test of the question above. To do so, we replicated Experiment 1 but this time, in line with Connine et al.'s (1991) work, we allowed our participants to respond at any time. If listeners were engaged in a strategy of relying on the biasing word, if listeners were permitted to respond as soon as they heard the target phoneme, such a strategy would likely be counterproductive as accurate performance would decrease and word recognition would suffer. Thus, we would predict that if a strategy was being employed in Experiment 1 to reduce the demands on processing, our data patterns in the present experiment should look similar to those of Connine et al. (1991; e.g., no congruency effects in the far condition).

Method

Participants A group of 20 new individuals from the same pool, who met the same criteria as in Experiment 1, participated for course credit.

Stimuli The stimuli were the 192 sentences used in Experiment 1.

Procedure The procedure was identical to Experiment 1 with one exception. During the instructions, in addition to what they had been told in Experiment 1, participants were permitted to respond at any time after sentence onset.

Results and discussion

Because minimal differences were obtained across conditions in the false alarm data of Experiment 1, and no differences were observed in the false alarms in this and the subsequent experiment, only the hit rate data will be discussed. For completeness, the false alarm data are included in the figures.

Mean hit rates and false alarm rates are shown in Fig. 3a. A congruency effect in the hit rates can be seen in both the near (.23) and far (.14) conditions. ANOVAs with Congruency and Distance as factors yielded a main effect of congruency, $F(1, 19) = 14.4, p = .001$, and a Distance \times Congruency interaction, $F(1, 19) = 6.0, p < .05$. The main effect of distance was not reliable. Paired t tests showed that the congruency effect in the near and far conditions were reliable [near condition, $t(19) = 4.1, p < .001$; far condition, $t(19) = 2.96, p < .005$].

The change in instruction, allowing participants to respond at any time rather than wait until sentence offset, had the effect of reducing the size of the congruency effect in both the near and far conditions relative to Experiment 1. In the near condition, the congruency effect shrank by .09 (.32 vs. .23). In the far condition, the difference was .14, (.28 vs. .14). A comparison of the congruency conditions from Experiments 1 and 3 was performed to demonstrate that these differences were reliable. A three-factor ANOVA based on experiment (1 vs. 3), distance, and congruency indicated that both the main effect of congruency, $F(3, 48) = 52.5, p < .001$, and the Congruency \times Distance interaction, $F(1, 48) = 9.0, p < .005$, were reliable. The variable of experiment showed a marginally reliable main effect, $F(1, 48) = 3.1, p < .09$. Its interactions with distance, $F(1, 48) = 3.0, p < .09$, and congruency, $F(1, 48) = 2.8, p < .11$, were marginally reliable. No other effects were reliable. These results indicate that when participants were not required to wait for a biasing word before responding, the size of the congruency effect shrank, but more so in the far than in the near condition.

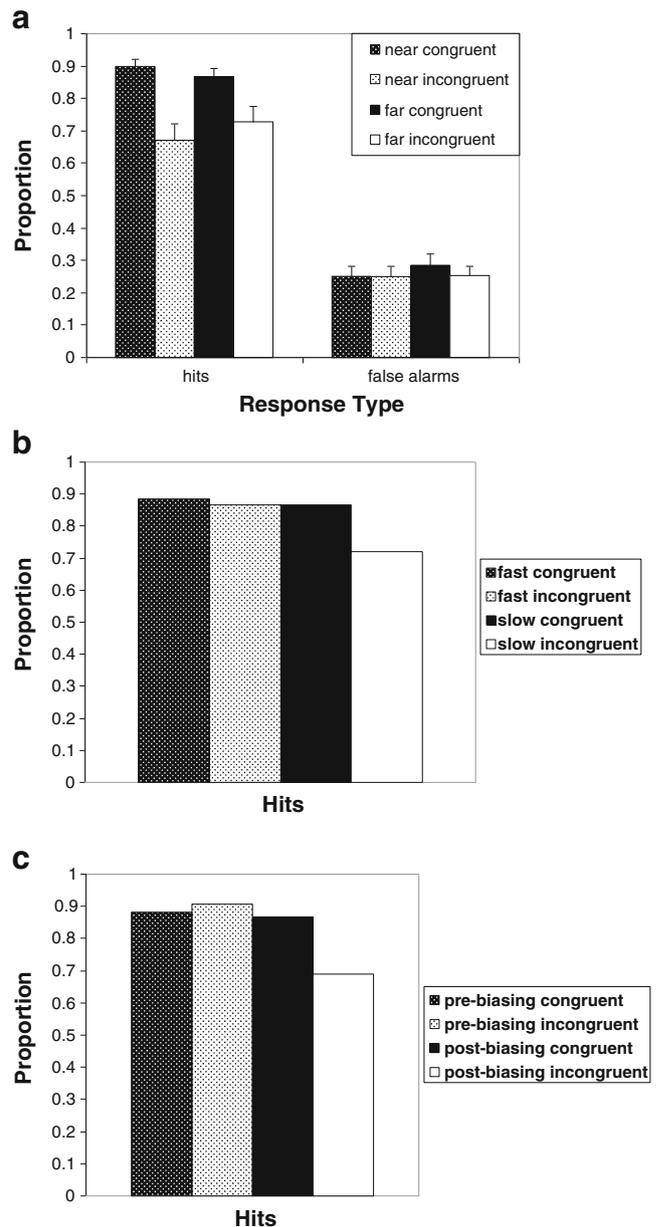


Fig. 3 Results of Experiment 3. **a** Mean hit rates and false alarm rates across the four conditions. **b** Cumulative hit rates across the fast and slow (1-s boundary) response bins. **c** Cumulative hit rates across the prebiasing word and postbiasing word response bins

To more closely explore the effect of responding at any time, we performed two sub analyses on the far-condition data. The first was a direct test of the 1-s temporal window and examined the magnitude of congruency effect as a function of whether participants responded before or after 1 s. Responses with latencies less than 1 s were placed into the fast-response bin, and those slower than 1 s were placed into the slow-response bin. Because only a small number of data points fell within the fast bin, the total proportion of hits was calculated for each condition using all responses rather than obtaining an average over participants. The low N also made

it unwise to perform inferential analyses, as they could be misleading. The hit rates in the fast and slow partitions are shown in Fig. 3b. We found no congruency effect in the fast bin, with the proportions of hits in the congruent condition differing from those in the incongruent condition by only .02. In contrast, a .15 congruency effect emerged in the slow bin. This analysis shows that the congruency effect in the far condition was confined to trials in which responses exceeded 1 s. This makes sense, as the onset of the biasing word occurred no earlier than 1.276 s after target word onset.

A weakness of the preceding analysis is that the data were partitioned without reference to the biasing word itself. To remedy this situation, we repeated the analysis, but instead split the data as a function of whether the response occurred before or after the onset of the biasing word. Responses occurring before this word were placed into the prebiasing word bin (congruent = 14.2 % of the data, incongruent = 18.1 % of the data). Those made after its onset were placed in the postbiasing word bin (congruent = 85.9 % of the data, incongruent = 81.9 % of the data). As in the first subanalysis, owing to the small number of responses before the biasing word, the proportions in each condition are cumulative across participants, and inferential statistics were not performed on the data. The data can be found in Fig. 3c. In the prebiasing word bin, a reverse congruency effect (−.03) was obtained, while in the postbiasing word bin, a .18 congruency effect was obtained, which was similar to that obtained when the data were not partitioned. Together, the data from these two subanalyses point to a difference in instruction, specifically when responses can be made, as the cause of the differences across studies. Slow responding, whether it is defined by time or the onset of the biasing word, yields a healthy congruency effect.

One additional question that needs to be addressed is why the RT rates in this experiment were much slower than those of Connine et al. (1991). They reported RT rates in the far condition of around 1,090 ms, whereas the mean in the present experiment was around 2,400 ms. It may be easier and therefore increase response rates to make a linguistic-based judgment than making a signal detection judgment.

Experiment 4

The purpose of the final experiment was to test an assumption implicit in the preceding experiments, that the biasing word, and not the words in the intervening region (e.g., *The = ing had an exquisite set of feathers*), was at least in part the cause of the observed changes in performance across congruency conditions. We addressed this possibility by presenting participants with the far-condition sentences from Experiment 1 with the biasing word present (present biasing word condition: e.g., *The = ing had an exquisite set of*

feathers) or absent (absent biasing-word condition: e.g., *The = ing had an exquisite set of*).

As with Experiment 1, participants were required to wait until after sentence offset before responding. If only the biasing word influenced responding, there should be no congruency effect in the absent biasing-word condition, whereas a large congruency effect should be obtained in the present biasing-word condition. If, instead, other properties of the sentences, such as the intervening words, contributed to the congruency effect, then both conditions should yield congruency effects.

Method

Participants A group of 22 new participants from the same pool meeting the same criteria as those in Experiment 1 participated.

Stimuli The 96 far-condition sentences from Experiment 1 (congruent = *The wing had an exquisite set of feathers*, incongruent = *The wing had an exquisite set of diamonds*), in their added and replaced versions, served as the stimuli in the present biasing-word condition. To construct the items for the absent biasing-word condition, each sentence had its biasing word spliced out (e.g., *The wing had an exquisite set of*), yielding a total of 192 items.

Procedure The procedure was identical to Experiment 1 with one exception. To prevent listeners from inferring or restoring the biasing words in the absent biasing-word condition sentences after hearing them in the present biasing-word sentences, the stimuli were split across two lists such that no list had both the present and absent biasing-word versions of the same sentence (e.g., if a listener heard *The = ing had an exquisite set of feathers*, this individual would not hear *The = ing had an exquisite set of*). This constraint required creating two stimulus lists, each with 96 sentences. Each list contained an equal number of stimuli from the four conditions.

Results and discussion

Figure 4 contains the hit and false alarm data. When the biasing word was present, a strong congruency effect was found (.23), with the congruent condition showing an increased proportion of hits relative to the incongruent condition. In contrast, the congruency effect disappeared when the biasing word was absent (−.01). An ANOVA indicated that only the congruency main effect, $F(1, 21) = 18.2, p < .001$, and a Biasing Word Presence \times Congruency interaction, $F(1, 21) = 25.3, p < .001$, were reliable. The main effect of biasing word presence was not reliable. Planned comparisons indicated that the congruency effect was reliable only in

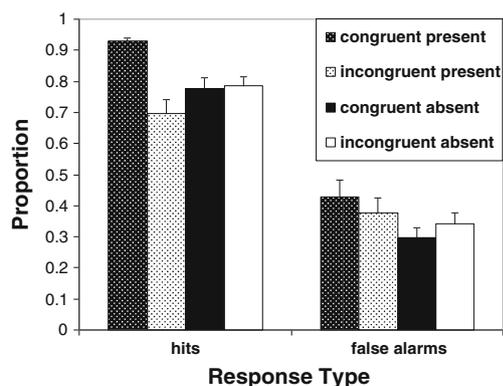


Fig. 4 Mean hit rates and false alarm rates in the present and absent biasing-word conditions in Experiment 4

the present biasing-word condition [$t(21) = 5.16, p < .001$], whereas the effect in the absent biasing-word condition proved unreliable.

The findings of Experiment 4 demonstrate that the results of Experiments 1 and 3 were not caused by the intervening words (e.g., *The = ing had an exquisite set of feathers*), but rather were the result of the sentence-final biasing word influencing identification of the target word.

General discussion

In four experiments, we examined how biasing subsequent context alters the perception of prior lexical ambiguities, in particular revisiting the idea that ambiguities are resolved within 1 s (Connine et al., 1991). Experiment 1 demonstrated that when participants were required to wait until after biasing word offset to respond, even when that word occurred more than a second after the target word, the biasing word influenced performance. Experiment 2 showed that the effect was not paradigm-specific, by generalizing the results to phoneme identification. Experiment 3 demonstrated that when participants were permitted to respond at any time, the biasing word maintained a strong influence on performance as long as the biasing word was heard before a response was made. Experiment 4 revealed that the large congruency effects observed in the first three experiments were caused by the biasing word, and not by the words present between the target and biasing word. Combined, the findings suggest that biasing information more than a second after the offset of an ambiguous word can maintain a strong influence on word identification, at least when memory load (e.g., simplistic syntax) is held at a minimum.

These results should not be interpreted as implying that context effects can always extend backward more than a second in time. As we mentioned earlier, there are likely stimulus and processing constraints on the size of the window, and a comparison of the present results with those of

Connine et al. suggests two important possibilities. One possible reason for the difference across studies is that in the experiments reported by Connine et al., the sentence structures were relatively complex, as they included embedded clauses, whereas in the present set of experiments the syntactic structure of the stimuli was kept relatively simple. It is therefore possible that in their experiment, having to maintain an ambiguity in memory while continuing to process a relatively complex sentence structure may have increased memory load. To reduce this load, listeners may have responded early as a way to enhance processing while reducing the demands placed on memory. Another possibility relates to the way in which the sentences were spoken. Connine et al. reported that their sentences were spoken with prosodic cues present. In the present experiments, the prosodic markers were reduced. It is therefore possible that one factor influencing early commitment was the presence or absence of prosodic cues. Given that Samuel (unpublished manuscript) reported producing sentences with reduced prosody while showing effects similar to those of Connine et al., the latter argument seems less likely to explain these differences.

A somewhat consistent finding across the restoration experiments is that listeners showed a decrease in their “added” responses in the incongruent as compared with the congruent condition when the biasing context was present. Does the presence of a mismatch somehow cause a rapid decay of the signal? Although we have no direct test of this question, we hypothesized that the “replaced” response in the incongruent condition was based on conscious decision-making. In other words, it is entirely possible that listeners perceived /w/, and because a mismatch was presented, they second-guessed their perception of the acoustic-phonetic features that they had perceived, and therefore reported the phoneme as being replaced by the noise.

One question that arises from the present data is whether the long-lasting biasing word influence is perceptual or postlexical in nature. In other words, does the biasing word directly influence word recognition, or does it exert influence at a postlexical stage, at which the ambiguous target word is integrated into the sentence? Responses to the replaced stimuli are a key piece of the data used to address this issue, because they provide the strongest test of perceptual restoration, as the target phoneme has been completely replaced by noise. If the biasing word truly affected lexical activation, we would have expected to see a much higher rate of restoration to the replaced stimuli. That is, top-down influences should have been so strong that listeners would have responded “added” on the majority of trials. Although listeners did produce “added” responses, their low rate (~25% on average) suggests that biasing influences, although present, were decidedly weak, which is more suggestive of a baseline level of response uncertainty across conditions than of perceptual restoration.

In the context of these false alarm data, the huge effect of congruency when the stimuli were added seems most plausibly a response bias to achieve semantic coherence, one that surprisingly was minimally sensitive to time with restoration, being approximately the same size in the near and far conditions. The phoneme identification data demonstrated greater sensitivity to temporal distance. Nevertheless, the long-lasting backward context effects are a desirable property of a postlexical biasing mechanism, because semantic integration can be required over long and short stretches of subsequent speech. In addition, that such large effects were found in the near condition demonstrates how quickly a following context can influence processing.

How far backward in time does the contextual-integration window extend? Our experiments did not address this question, so we will not speculate on possible values. Instead, we propose that the end of the sentence or phrase itself might serve as an important boundary at which ambiguity resolution could naturally occur, even if it is unsuccessful because of the absence of sufficient biasing information; this prediction would fit nicely with the work of Connine et al. (1991). This window, therefore, is defined by considerations of language use, notably meaning and/or syntactic complexity, not by time, and possibly is shaped by experience with the language. This idea would be easily testable by extending the present setup to include conditions involving multiple sentences.

This same line of reasoning can be extended to suggest that a corresponding structural bound might extend to temporally closer contextual biases, specifically those within words. The phoneme restoration paradigm (and, to a lesser degree, phoneme identification) has yielded two qualitatively different data patterns, depending on the type of context, which may reflect different processes. When the context is a sentence, changes in criterion dominate, with minimal changes in d' —a result that is suggestive of postlexical decision processes (Samuel, 1981). However, when the context is an isolated word (e.g., = *avern*), changes in d' are the main finding, which may suggest a different origin of the context effect (Norris, McQueen, & Cutler, 2000; Samuel, 1981). Perhaps the time course of this intraword context effect, which is also found with phoneme identification, is also defined structurally by the boundary of a word. In the present experiments, the later responses may have been reflective of a criterion change. This interpretation predicts that changes in d' should not be found when part of the context crosses a word boundary. Word boundaries are thus closely linked to the type of process responsible for contextual influences.

The preceding discussion leads to the question of how the ambiguous target word is resolved. It may be that when a word that is somewhat ambiguous is encountered, the system initially activates and selects the best-fitting representation (Marslen-Wilson, 1991; McClelland & Elman, 1986). If the

signal is too unclear for a perfect match to be made, the identified word is then moved into some type of short-term memory buffer, in which it is actively maintained until later incoming information becomes available and can provide the necessary resolution. For example, mechanistically, if some fit threshold is not exceeded in the lexicon, the utterance is held in this buffer to wait for further evaluation to take place. According to this account, the identification of ambiguous words might take place in what can be thought of as two stages, the latter of which includes a repair mechanism that is engaged when a mismatch is encountered between the lexical representation being maintained and the subsequent biasing information (Marslen-Wilson, 1991). The goal of the repair process is to generate the most plausible interpretation of what was said, given the immediate and possibly more distant context. Computationally, this is a very difficult problem to solve, because of the uncertainty in identifying the proper interpretation of the sentence that the talker intended, especially online (Norris & McQueen, 2008).

This conceptualization of semantic integration is in line with work suggesting that word recognition is not always sequential (Bard, Shillcock, & Altmann, 1988; Dahan, 2010; McMurray, Tanenhaus, & Aslin, 2009), but exhibits characteristics of a parallel process that is made possible by long-lasting memory traces of words. Biasing information subsequent to the word being identified can be a useful source of information for achieving accurate word identification and message comprehension. In this regard, it makes sense for the processing system to have a means of delaying the final commitment to a specific lexical entry.

In sum, the findings of this study provide additional insight into the time course of the influence of subsequent biasing context on word identification. They suggest that the processor can delay final commitment to a lexical representation for longer than 1 s in order to influence word identification, at least when memory load is kept low. Because talkers most often speak to convey a message, this capability may be one means to ensure that comprehension is robust.

Author note We thank Cindy Connine, Dennis Norris, and several anonymous reviewers for comments on earlier drafts of this work. We also thank Hartman Brawley, Victoria Cook, Emily Jamison, Pricilla Ju, Emily Kurtz, Paige Lisko, Karin Luk, Kristen Peters, Nicholas Satariano, Sara Stanutz, and Vanessa Waller for assistance with stimulus preparation and data collection. Portions of this work were presented at the 50th Annual Meeting of the Psychonomic Society in Boston, Massachusetts. Experiments 1 and 4 served as part of the first author's master's thesis.

Appendix 1: Stimuli used in Experiments 1, 3, and 4

In the sentences below, the word in italics is the competitor, which is congruent with the incongruent context. The word

prior to the parenthetical information is the intervening syllable in the near biasing-word condition, whereas the words within parentheses made up the additional intervening syllables in the far, present, and absent biasing-word conditions. The final two words in each row are the congruent and incongruent biasing words, respectively, and were removed in Experiment 4.

The knock/*lock* was (not quite able to be) heard/snapped.
 The leak/*beak* was (unexpectedly) fixed/sharp.
 The lip/*rip* was (very noticeably) chapped/stitched.
 The list/*wrist* was (found to be thoroughly) numbered/injured.
 The log/*dog* was (very actively) rotting/barking.
 The look/*book* was (not going to be) modeled/written.
 The match/*latch* was (readily able to be) struck/closed.
 The math/*bath* had (included a lot of) fractions/bubbles.
 The mole/*role* was (not adequately) digging/assigned.
 The mug/*rug* was (going to be completely) filled/swept.
 The nap/*wrap* was (unfortunately not) restful/tasty.
 The nest/*guest* had (some very colorful) sparrows/luggage.
 The niece/*lease* was (finally going to be) married/cosigned.
 The night/*light* was (not actually) starless/shining.
 The roast/*ghost* was (considered to be very) salty/spooky.
 The rum/*gum* was (not going to be) poured/chewed.
 The well/*bell* was (unexpectedly) dry/rung.
 The wick/*nick* was (insufficiently) burning/repared.
 The wife/*knife* was (most definitely) pregnant/pointed.
 The wig/*rig* was (not going to be) shampooed/driven.
 The wing/*ring* had (an exquisite set of) feathers/diamonds.
 The wish/*dish* was (unexpectedly) granted/shattered.
 The wool/*bull* was (not going to be) knitted/angered.
 The yarn/*barn* was (in the first stage of being) woven/painted.

Appendix 2

The sentences used in Experiment 2 are provided below. The *__ip* noted below was produced as *sip*, as an ambiguous *?ip*, and as *ship*. The portions in parentheses were spoken only in the far condition.

sip-biased sentences

The *__ip* was (not able to be) tasted.
 The *__ip* was (quite surprisingly) slurped.
 The *__ip* was (not going to be) drunk.
 The *__ip* was (not able to be) pored.
 The *__ip* had (a large amount of) sweetener.
 The *__ip* was (almost ready to be) gulped.
 The *__ip* was (not going to be) refreshing.
 The *__ip* was (not going to be) diluted.

ship-biased sentences

The *__ip* was (unexpectedly) traveling.
 The *__ip* was (fairly noticeably) leaky.
 The *__ip* was (not expected to be) navigated.
 The *__ip* was (unlikely to be) captured.
 The *__ip* was (not able to be) wrecked.
 The *__ip* was (quite noticeably) damaged.
 The *__ip* had (to have a lot of) pirates.
 The *__ip* was (expected to be) steered.

References

- Bard, E. G., Shillcock, R. C., & Altmann, G. T. M. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context. *Perception & Psychophysics*, *44*, 395–408.
- Bashford, J. A., Warren, R. M., & Brown, C. A. (1996). Use of speech-modulated noise adds strong “bottom-up” cues for phonemic restoration. *Perception & Psychophysics*, *58*, 342–350.
- Borsky, S., Tuller, B., & Shapiro, L. P. (1998). “How to milk a coat”: The effects of semantic and acoustic information on phoneme categorization. *Journal of the Acoustical Society of America*, *103*, 2670–2676.
- Cole, R. A., Jakimik, J., & Cooper, W. E. (1980). Segmenting speech into words. *Journal of the Acoustical Society of America*, *67*, 1323–1332.
- Connine, C. M., Blasko, D., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*, *30*, 234–250.
- Dahan, D. (2010). The time course of interpretation in speech comprehension. *Current Directions in Psychological Science*, *19*, 121–126.
- Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, *44*, 325–349.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- Marslen-Wilson, W. D. (1991). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.
- Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86. doi:10.1016/0010-0285(86)90015-0
- McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from “lexical” garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language*, *60*, 65–91.
- McQueen, J. M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 433–443. doi:10.1037/0096-1523.17.2.433
- Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357–395. doi:10.1037/0033-295X.115.2.357

- Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *The Behavioral and Brain Sciences*, *23*, 299–370.
- Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 699–725. doi:10.1037/0096-1523.19.4.699
- Pitt, M. A., & Samuel, A. G. (2006). Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1120–1135. doi:10.1037/0096-1523.32.5.1120
- Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, *110*, 474–494.
- Samuel, A. G. (1987). Lexical uniqueness effects on phonemic restoration. *Journal of Memory and Language*, *26*, 36–56.
- Sherman, G. L. (1971). *Phonemic restoration: An insight into the mechanisms of speech perception (Unpublished master's thesis)*. Milwaukee, WI: University of Wisconsin.
- van Alphen, P., & McQueen, J. M. (2001). The time-limited influence of sentential context on function word identification. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1057–1071. doi:10.1037/0096-1523.27.5.1057
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 392–393.
- Warren, R. M., & Sherman, G. L. (1974). Phonemic restoration based on subsequent context. *Perception & Psychophysics*, *16*, 150–156.