



Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2008 September ; 34(5): 1207–1223. doi:10.1037/a0012796.

Context and Spoken Word Recognition in a Novel Lexicon

Kathleen Pirog Revill,

Department of Brain and Cognitive Sciences, University of Rochester

Michael K. Tanenhaus, and

Department of Brain and Cognitive Sciences, University of Rochester

Richard N. Aslin

Department of Brain and Cognitive Sciences, University of Rochester

Abstract

Three eye movement studies with novel lexicons investigated the role of semantic context in spoken word recognition, contrasting three models: restrictive access, access-selection and continuous integration. Actions directed at novel shapes caused changes in motion (e.g., looming, spinning, etc.) or state (color, texture, etc.). Across the experiments, novel names for the actions and the shapes varied in frequency, cohort density, and whether the cohorts referred to actions (Experiment 1) or shapes with action-congruent or incongruent affordances (Experiments 2 and 3). Experiment 1 demonstrated effects of frequency and cohort competition from both displayed and non-displayed competitors. In Experiment 2 a biasing context induced an increase in anticipatory eye movements to congruent referents and reduced the probability of looks to incongruent cohorts, without the delay predicted by access-selection models. In Experiment 3, context did not reduce competition from non-displayed incompatible neighbors as predicted by restrictive access models. We conclude that the results are most consistent with continuous integration models.

A long-standing issue in spoken word recognition is how syntactic and semantic information from the prior context constrains the activation of lexical candidates and the eventual decision that a portion of fluent speech is a particular word. The effects of context could be pre-emptive, immediate, or delayed relative to the time course of spoken word recognition, and all three hypotheses have found support over the last several decades.

One hypothesis, often called *selective* or *restrictive access*, proposes that a sufficiently constraining context could restrict access to contextually congruent meanings (for a review see Simpson, 1984; Lucas, 1999). In the strongest version of a restrictive access model, a sufficiently constraining context can both bias recognition of contextually consistent lexical candidates and inhibit inconsistent lexical candidates, therefore reducing sensitivity to the bottom-up input. Thus, a suitably strong context could in principle pre-select a consistent subset of the lexicon by priming consistent items and blocking the activation of inconsistent items, thereby limiting the alternative lexical candidates that are activated with the arrival of the bottom-up speech input.

Correspondence concerning this article should be addressed to Kathleen Pirog Revill, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332. E-mail: kate.pirog@psych.gatech.edu.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at <http://www.apa.org/journals/xlm/>

A second hypothesis, called *exhaustive* or *multiple access*, proposes that all phonologically consistent lexical competitors are initially activated during a context-independent stage in which lexical activation is solely determined by the input. Context is later used to select the congruent word or meaning. One influential formalization suggests that spoken word recognition can be divided into two partially overlapping sub-processes: *access* and *selection* (Marslen-Wilson 1987; 1989). The access-selection framework assumes an early temporal window during which the activation of syntactic and semantic components of a lexical candidate is solely determined by the match to the auditory input without influence from, or integration with, available contextual constraints. Context is then used in a later time window to help guide selection. Once a candidate is selected and integrated with context, recognition is complete.

A third hypothesis, called *continuous integration*, has emerged more recently. According to this hypothesis, mapping from the input to meaning is a continuous process in which integration of contextual constraints is immediate -- there is no fixed set of lexical candidates, no discrete point in time at which lexical selection is completed, and no end to the influence of multiple sources of information, even after the speech signal for a word has ended (e.g., Allopenna, Magnuson & Tanenhaus, 1998; Gaskell & Marslen-Wilson, 1997; 2002). Contextual constraints provided by syntax and semantics are used as soon as they become available, but the word recognition system remains sensitive to information coming from the bottom-up input even after integration has begun. Crucially, the continuous integration framework, like the restrictive access model, predicts that context will exert immediate and continuous effects on spoken word recognition, whereas the access-selection framework predicts delayed effects of context. Unlike the strong form of the restrictive access model, however, the continuous integration framework predicts that context will not reduce the sensitivity of the lexicon to the input, and therefore should not eliminate effects of inconsistent lexical candidates.

The goal of the current studies is to distinguish among these three models by using the visual world paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995) in conjunction with an artificial language (Magnuson, Tanenhaus, Aslin & Dahan, 2003). As participants follow spoken instructions to manipulate pictures in a display, their eye movements provide a sensitive measure of the time course of spoken word recognition (e.g., Creel, Aslin & Tanenhaus, 2006; Creel, Tanenhaus & Aslin, 2006; Magnuson, Tanenhaus & Aslin, in press; Shatzman & McQueen, 2006; Wonnacott, Newport & Tanenhaus, in press). An emerging body of evidence demonstrates that eye movements to pictures are sensitive to the earliest moments of lexical processing. For example, even five ms within-category differences in voice onset time at the beginning of a word will modulate the timing of fixations to a referent and its competitors (McMurray, Tanenhaus & Aslin, 2002). Furthermore, eye movements are sensitive enough to demonstrate that asynchronous acoustic/phonetic cues that combine to determine a phonetic percept (e.g., whether a sound is voiced or voiceless) affect lexical processing as soon as each cue becomes available (e.g., McMurray, Clayards, Tanenhaus & Aslin, submitted). Artificial languages enable complete experimenter control over the distributional properties of the stimulus set, allowing for targeted tests of hypotheses that would be difficult to test with real words. In particular, we use an artificial language to create a well-defined contextual constraint and then examine the effects of that constraint on (a) looks to a phonologically similar but contextually inconsistent competitor and (b) competition from non-displayed contextually inconsistent competitors. These conditions allow us to distinguish between restrictive access models, which predict immediate use of contextual constraints and reduction of neighborhood effects, access-selection models, which predict delayed use of contextual constraints, and continuous integration models, which predict immediate use of context with continued effects of inconsistent neighbors.

The standard approach for examining the effects of context on spoken word recognition, which was introduced by Conrad (1974) and adapted in classic studies by Swinney (1979), probes for activation of associates of different meanings. Using cross-modal lexical priming, researchers focused on finding evidence for activation of contextually inconsistent alternatives at different temporal intervals from the offset of the ambiguous word to make inferences about the recognition process. We introduce a second, complementary approach, which is to examine how context affects the dynamics of lexical competition among words that have partial acoustic/phonetic overlap. Target words that have substantial overlap with many other words are recognized more slowly than words with no overlap. This dimension of acoustic/phonetic overlap, weighted by word frequency, is called frequency weighted neighborhood density (Magnuson, Dixon, Tanenhaus & Aslin, 2007; Luce & Pisoni, 1998; Vitevitch & Luce, 1999). This approach requires using an artificial lexicon because the sound and meaning of words in natural language are only very weakly correlated, making it extremely difficult, if not impossible, to select natural language stimuli in which a contextual constraint is consistent with some words in a phonological neighborhood but inconsistent with others.

Evidence for activation of contextually inconsistent alternatives

Throughout the 1970s and 1980s, the effects of context on spoken word recognition were addressed primarily by examining their influence on the processing of homophones, such as *bank* and *rose* (for reviews see Lucas, 1999; Simpson, 1984; Tanenhaus & Lucas, 1987). Initial data from researchers looking for evidence of activation of contextually inconsistent alternatives were generally interpreted as supporting an access-selection framework. Cross-modal lexical decision (Swinney, 1979) and cross-modal naming (Tanenhaus, Leiman & Seidenberg, 1979) were used to examine the time course of lexical ambiguity resolution in semantic and syntactic contexts, respectively. The results suggested that all meanings of an ambiguous word are initially accessed, with higher frequency meanings accessed more rapidly than lower frequency meanings (Simpson, 1981), and that a contextually congruent meaning is then selected from among the activated alternatives as rapidly as 200 ms after the offset of the ambiguous word (Tanenhaus et al., 1979; Seidenberg, Tanenhaus, Leiman & Bienkowski, 1982). However, more recent studies by Tabossi and colleagues using more restrictive contexts have found little or no delay between the onset of a spoken word and when the effects of context are first observed (e.g., Tabossi, 1988; Tabossi, Colombo & Job, 1987; Tabossi & Zardon, 1993).

Although many common words have distinct meanings and most words have multiple senses, full lexical ambiguity represents a special case in which the sensory input remains entirely consistent with more than one semantically unrelated lexical candidate (e.g., homophones). However, because spoken words unfold over time, even unambiguous words are temporarily consistent with multiple lexical candidates. Zwitserlood (1989) extended the cross-modal priming paradigm to temporarily ambiguous words by examining priming of target words and words that shared the same onset sounds (cohort competitors). Zwitserlood found that gated fragments primed words semantically related to cohorts, regardless of contextual bias. However, priming for the cohort competitor declined faster when the previous context was biasing than when it was neutral. Zwitserlood concluded that the input activated semantic features of both fragment continuations without restriction by the contextual information, but that the biasing context speeded the selection process, results that strongly support the access-selection model.

Dahan and Tanenhaus (2004) evaluated predictions of the access-selection and continuous integration perspectives by using the visual world paradigm. They used participants' eye movements to examine the processing of lexical competitors in sentential contexts with naturally occurring Dutch constructions in which the verb precedes its subject noun phrase.

Four objects were presented on a computer screen, and participants were instructed to use a mouse to move the object that they heard named in the sentence. The subject noun phrase followed either a main verb that placed strong semantic constraints on the upcoming subject (e.g., *Nog nooit krom een bok zo hoog* [Never before climbed a goat so high], where ‘climb’ requires an animate subject), or an auxiliary or modal verb that did not place such constraints on its subject (e.g., *Nog nooit is een bok zo hoog geklommen* [Never before has a goat climbed so high]). In sentences where the main verb was introduced after the target, they found clear evidence for activation of phonologically related cohort words like *bot* (bone) when the target word was *bok* (goat) in the form of fixations to both the bone and goat before the point of disambiguation. When the alternate sentence construction was heard, however, the contextual constraints generated at the main verb affected cohort competition for the upcoming subject. Fixation proportions did not differ for cohort and unrelated distracter objects when the cohort did not fit the thematic restrictions established by the main verb. This pattern of results is inconsistent with an access-selection model, since when contextual constraints were available, they had immediate effects on cohort competition before the target word was completely disambiguated. These results are consistent with the earlier studies by Tabossi and colleagues (Tabossi, 1988; Tabossi et al., 1987; Tabossi & Zardon, 1993) and with studies using event-related potentials (Connolly & Philips, 1994; van Berkum, Zwitserlood, Hagoort & Brown, 2003; van den Brink, Brown & Hagoort, 2001; Van Petten, Coulson, Rubin, Plante & Parks, 1999).

Evidence of contextual effects on neighborhood competition dynamics

In order to separate restrictive access models and continuous integration models, it is necessary to address the system's remaining sensitivity to bottom-up input in the face of a contradictory context. A second experiment by Dahan and Tanenhaus (2004) supported continuous integration models over restrictive access models by demonstrating that the system, while integrating sources of constraint as soon as they become available, remains sensitive to the bottom-up input. Misleading coarticulatory information about place of articulation was cross-spliced into the initial vowel of the target word, making the input temporarily more consistent with the contextually incongruent cohort. For example, the initial “bo” of *bot* was added to the “k” of *bok* to create a stimulus that initially had coarticulatory information favoring *bot*, the cohort competitor. Under these conditions, weak cohort effects emerged immediately (accounting for a saccade planning delay of approximately 200 ms, Matin, Shao & Boff, 1993) after the onset of the misleading coarticulation information, even when the cohort was contextually inconsistent with the main verb and there was no evidence of activation of the inconsistent cohort prior to the misleading coarticulatory information.

Although the results of Dahan and Tanenhaus (2004) are most consistent with continuous integration models, the scope of their conclusions is constrained by the limitations of the task and manipulations they used. One potential concern with the visual world paradigm (e.g., Corley, MacGregor & Donaldson, 2007) is that the use of a limited number of pictured referents results in closed-set effects that do not generalize to language processing in richer or more abstract environments. This seems unlikely, however, because several studies have shown that strategic effects do not seem to be the driving force behind eye movement patterns in visual world studies. Most crucially, the properties of *non*-displayed competitors do affect eye movements to a pictured referent (e.g., Dahan, Magnuson, Tanenhaus & Hogan, 2001; Magnuson et al., 2003; Magnuson et al., 2007; for a review, see Tanenhaus, 2007). However, Dahan and Tanenhaus (2004) did not examine effects of contextually or phonologically consistent lexical items outside of the visual display (non-displayed competitors). This is potentially problematic because the stimulus manipulation that Dahan and Tanenhaus used to demonstrate continuous sensitivity to bottom-up input is unlikely to occur outside the laboratory setting. Cross-spliced stimuli were created in which articulatory information

favoring a cohort competitor was embedded into the target word. This effectively created a ‘garden path’ situation where auditory information temporarily promoted a contextually inconsistent alternative. Dahan and Tanenhaus interpreted the resulting behavioral effects as indicating that bottom-up sensitivity remained in the face of contextual restrictions, which is consistent with continuous integration models but not with strong restrictive access. However, a more limited conclusion is that bottom-up information that is inconsistent with a contextually preferred interpretation but consistent with an additional displayed alternative can override probabilistic contextual biases. Although the results do indicate that the joint presence of a biasing context and a contextually consistent target picture do not eliminate sensitivity to low-level acoustic detail, a more natural and general demonstration of bottom-up sensitivity would be desirable to rule out strong restrictive access interpretations.

One way to attain a more natural measure of sensitivity to bottom-up phonological information in the face of top-down contextual restrictions is to examine how context interacts with lexical neighborhoods, an approach that has not been explored previously. Several paradigms have demonstrated that words in high-density neighborhoods (computed by taking a frequency-weighted measure of the number of lexical items sharing most but not all phonemes with the target item) are recognized more slowly than words in low-density neighborhoods (Luce & Pisoni, 1998; Vitevitch & Luce, 1998; Magnuson et al., 2007). Words with many neighbors or words with high frequency neighbors must overcome greater aggregate competition than words in sparse neighborhoods. If a constraining context can reduce or eliminate consideration of contextually inconsistent candidates, then contextually inconsistent neighbors would compete less effectively with the target, reducing the effects of neighborhood density and speeding target recognition. However, if context primes the activation levels of contextually consistent candidates, recognition of the target will be slowed by the need to overcome the additional competition from the primed competitors.

Use of an artificial lexicon to manipulate context

Examining the effects of context on lexical neighborhoods requires using constraining or biasing contexts that are clearly compatible with some of the lexical candidates in the lexicon but incompatible with others. Because of the probabilistic nature of contextual biases and the free relationship between word form and word meaning in natural languages, it is virtually impossible to select a set of phonological competitors in which some of the lexical items are highly consistent with a particular context while other items are highly inconsistent with it. The approach we adopt here is to use a miniature novel lexicon to control both the characteristics of lexical competitors and the nature of the constraint provided by context.

In order to provide a fine-grained measure of the time course of lexical processing, we tracked eye movements as participants performed manipulations in accordance with spoken instructions. In the present series of experiments, participants learned to associate new lexical items with two types of visually presented referents: a set of novel objects that could be further subdivided into two perceptually similar families based on edge contours (straight or curved), and a set of icon buttons that could be subdivided into two classes based on whether they caused changes in motion (i.e. spinning or looming) or surface appearance (i.e. texture or color) to occur.¹ Participants selected an object by clicking on it with a computer mouse and modified it by clicking on an icon button. We used the different subcategories of objects and icon buttons to create the strong contextual constraints necessary to test our hypotheses by controlling the pairings of objects and modifiers seen during training. Participants learned that certain

¹We chose these dimensions because they are known to activate different areas of the brain associated with relatively early visual processing, a design feature relevant to a follow-up experiment examining fMRI correlates of lexical activation. The motion modifications are known to strongly activate the MT/MTS complex in area V5 of visual cortex, whereas the color and texture modifications do not (Pirog Revill, Aslin, Tanenhaus, & Bavelier, submitted).

modifications could only apply to objects with similar perceptual 'affordances'. For example, a participant might see only modifiers that caused motion acting upon objects with curved edge contours, and only objects with straight edge contours undergoing modification by texture or color change icons. Thus, upon hearing the name of an icon with affordancy restrictions, participants had information about what the affordances of the upcoming object must be. Previous research has shown that the affordances of objects -- the perceptual features that define interactions with an object -- affect which objects participants deem to be plausible referents (Chambers, Tanenhaus, Eberhard, Filip & Carlson, 2002; Chambers, Tanenhaus & Magnuson, 2004). The artificial lexicon enabled us to control the acoustic/phonetic similarity of the set of lexical alternatives and allowed us to design competitor sets partially or fully consistent with the affordance-defined context. We also manipulated lexical frequency and the density of the lexical neighborhoods, as specified by number and frequency of cohorts within and between affordancy-defined categories.

There is, of course, no way to guarantee that results from artificial languages and lexicons will generalize to natural language. However, studies of spoken word recognition using artificial languages are emerging as important complements to studies with natural language stimuli (Magnuson et al., 2003; Creel, Aslin & Tanenhaus, 2006; 2008; Creel, Tanenhaus & Aslin, 2006; Shatzman & McQueen, 2006; Wonnacott et al., in press). Crucially, spoken word recognition in a novel lexicon shows the same signature effects as normal word recognition, including competition between displayed cohort and rhyme competitors, effects from non-displayed competitors, and frequency-dependent lexical activation. In addition, processing in a novel lexicon is relatively unaffected by lexical neighborhoods from English under certain circumstances (Magnuson et al., 2003). We conducted Experiment 1 to establish that our experimental environment shows the standard lexical effects seen in previous visual world studies of spoken word recognition, specifically those effects that would be crucial for diagnosing the loci of context effects. In particular it is crucial to demonstrate effects of: (a) displayed cohorts, with increased looks to phonologically related items relative to phonologically unrelated items; (b) target and cohort frequency, with earlier looks to high frequency targets compared to low frequency targets and more looks to high frequency than low frequency cohorts; and (c) non-displayed competitors, with higher frequency non-displayed competitors having stronger effects on target activation than lower frequency non-displayed competitors.

In Experiment 2, we evaluated evidence for activation of contextually inconsistent competitors. We measured the time course over which contextual constraints and the ongoing bottom-up input combine by examining the effect of constraining context on looks to contextually congruent and incongruent cohorts. In neutral contexts, all three models of contextual integration predict initial looks to both the eventual target and its cohort competitor. In biasing contexts, the access-selection framework predicts delayed effects of context; initial consideration of cohorts, as indexed by early looks, should not be influenced by context, as reported by Zwitserlood (1989). In contrast, both restrictive access and continuous integration models predict that context will immediately affect looks to the inconsistent competitor. In addition, a necessary precondition for evaluating whether context effects are delayed because of the architecture of the language processing system is to establish that the contextual constraint has been computed and is available prior to the onset of the temporarily ambiguous word. We used anticipatory eye movements (Altmann & Kamide, 1999) to confirm that participants had learned the contextual constraints, and most crucially to confirm that the constraint was available prior to the onset of the temporarily ambiguous target word.

Visual world eye tracking has previously been used as a measure of processing load to track the effects of different types of lexical neighbors (Magnuson et al., 2007). In Experiment 3, we examined the dynamics of lexical competition by observing the effects of non-displayed

cohort competitors that were either consistent or inconsistent with contextual affordancy restrictions on the probability of fixating the target. The interaction of context and non-displayed neighbors should be particularly informative about the loci of context effects. Continuous integration and restrictive access models each predict immediate effects of context. However, the two classes of models make different predictions about how context will affect the set of lexical candidates (lexical neighbors) that are consistent with the bottom-up input. Restrictive access models predict that a biasing context should eliminate, or at least sharply reduce, the effects of neighbors that are incompatible with the contextual constraint on recognition of the target word. In contrast, continuous integration models predict that context should not eliminate neighborhood density effects.

Experiment 1

In this experiment, we first verify that standard cohort and frequency effects can be found using an artificial lexicon and novel visual environment in the visual world paradigm. We then examine whether correlated perceptual features, in this case the spatial arrangement of manipulable items on a computer screen as well as the movement and surface properties of those items, can indeed serve as a biasing context to begin to explore how these factors may affect lexical access.

Method

Participants—Twenty-one native speakers of English who reported normal or corrected-to-normal vision and normal hearing were paid \$10/hr for their participation. Data from an additional four participants were excluded from the analysis because their word-referent performance on the final test was < 85% correct.

Materials—Participants learned a 16-word artificial lexicon. Eight CVCV lexical items referred to simple novel objects that appeared on a computer screen, and eight CVCVCV lexical items to modifiers of those objects, also represented onscreen by icon buttons. Each display contained four of the eight objects and four of the eight modifier icons (see Figure 1). Four different random mappings between the 16 words and their referents were created, and the mappings were counterbalanced across participants.

The eight modifiers were divided into two categories. Clicking on four of the modifier buttons caused the selected object to undergo changes in motion trajectory (horizontal oscillation, vertical oscillation, size oscillation, and clockwise rotation), while the other four buttons caused changes to the object's color or texture (grey to white, grey to black, grey to speckled, grey to marbled). The words referring to these modifiers were trisyllabic CVCVCV items, and each was identical to another modifier word up to the final vowel, creating four cohort pairs (see Table 1). Both items in two of the phonologically-related cohort pairs referred to icons within the same modifier category, for example, *kubera* and *kubero* respectively referring to horizontal oscillation and clockwise rotation. We refer to these as same-category cohorts. For the other two cohort pairs, one item in each pair referred to a motion icon and one to a surface change icon, for example, *goreba* and *gorebi* respectively referring to vertical oscillation and speckling. We refer to these as cross-category cohorts. In order to examine the effects of word frequency on activation in the artificial lexicon, one member of each cohort pair was presented as the target five times more often than the other member of the pair throughout training.

In Experiment 1, the spatial arrangement of the modifier icons was correlated with change category. During all training blocks, icons with related perceptual properties were consistently presented in one region of the display spatially distinct from the area assigned to modifiers from the other category. Icons from the first category were displayed in two slots (randomly assigned across trials) in a 'toolbar' running along the top of the screen, and two icons from

the other category were displayed in a toolbar along the bottom of the screen. Participants were not explicitly told of this arrangement, or that there was any sort of category structure (motion/surface appearance) among the modifiers.

The lexical items referring to objects were bisyllabic CVCV items with unique onset syllables. They were presented at equal frequencies throughout the experiment. The novel objects were not easily described with one-word English descriptions, were not highly confusable with each other, and were uniformly grey in color before any modification occurred. There were no affordancy restrictions in Experiment 1, so the objects were allowed to appear in any one of the four possible object positions in the center portion of the screen and were acted upon by each modifier an equal number of times throughout the course of the experiment.

The words were spoken as isolated tokens by a female native speaker of English and were recorded in a quiet room on a Kay Elemetrics Computerized Speech Lab (model 4300B) sampling at 11025 kHz, for use with the experimental control software PsyScope 1.2 (Cohen, MacWhinney, Flatt & Provost, 1993). The mean duration of the trisyllabic items was 769 ms (range 670-840 ms), and the mean duration of the bisyllabic items was 571 ms (range 528-610 ms).

Procedure—Each participant completed a 60 minute session on two consecutive days. Participants were told that they were going to use a computer paint program to create scenes on the screen in response to spoken instructions. A voice would give instructions in an unfamiliar new language, and their task was to try to follow the instructions and create the specified scene on the computer screen by using ‘tool’ icons to modify the objects in the workspace area. A click to a tool icon would ‘pick up’ its property, and clicking on an object would apply that property to the object and cause a change. To aid in their learning, they would be shown the correct scene after their own attempt. At the beginning of each training trial, a small fixation cross appeared in the center of the screen. Participants were instructed to fixate and click the cross to initiate each trial. The fixation cross disappeared and was immediately replaced on the screen by up to four objects and up to four icon buttons, depending on the stage of training. The spoken instruction began simultaneously, allowing no time for preview of the stimulus display. Immediately after participants had clicked on one icon button and one object, all incorrect objects and incorrect icon buttons disappeared. The correct object and/or icon button remained onscreen, with the object undergoing the specified change, while the spoken instruction was repeated. Five hundred milliseconds after the end of the spoken phrase (which varied in length from 1198 to 1450 ms in the full-complexity training trials), the correct icon and object disappeared and the fixation cross reappeared to signal the start of the next trial. Test trials only differed from training trials in that the trial ended immediately after the participant's final mouse click. No feedback was provided, and participants could immediately begin the next test trial by clicking on the central cross.

Training began with the presentation of simple displays and eventually became more complex. First, participants learned the names of the modifier icons by choosing between two icon buttons, one of each modifier type, and applying that modifier by clicking on a single circle. Throughout training, the icon representing the target's phonological cohort never appeared onscreen as a distracter in the same display as the target. In the second block, four modifier icons were available on the screen in two tool bars. After completing one 96-trial block of each type (2AFC and 4AFC), participants progressed from learning the modifier names to learning the names of the objects. They received 96 trials of a simplified version of the final task, in which they heard a two-word phrase consisting of a modifier word and the name of the object to be modified, while viewing a scene containing two objects and four modifier icons. Then they received a 96-trial block at full complexity, with four objects and four modifier icons on the screen. Finally, they were eye-tracked on a 64-trial test block without feedback. By the end

of the first day, participants had been exposed to each object name 24 times, each low frequency modifier 16 times, and each high frequency modifier 80 times.

On the second day, participants completed 96 trials of the simplified design (four tool icons, two objects) as a refresher before undertaking two more 4×4 training blocks. By the end of these three practice blocks on Day 2, participants had been exposed to each object name 60 times, each low frequency modifier 28 times, and each high frequency modifier 140 times across both days of training. After training, a final 64-trial testing phase occurred. All icons and objects were presented as targets with equal frequency, and each icon modified each object exactly once. On half of the trials, the target icon's phonological cohort also appeared onscreen as a distracter icon. No feedback was provided during the test. In addition to mouse click responses, eye gaze was monitored during the test with an EyeLink II head-mounted eye tracker, which provided a record of point-of-gaze in screen coordinates at a sampling rate of 250 Hz. The eyetracker's software automatically parsed the eye track into three categories: saccades, fixations, and blinks. A saccade was assigned to a referent if the fixation coordinates fell within a 250×250 pixel window centered around the object or icon buttons, which were sized at 150×150 pixels. A look to a pictured referent was defined from the onset of the saccade into that referent's window to the end of the fixation on that referent. The test block began with the calibration of the eye-tracker using a standard nine-point calibration procedure. At a viewing distance of 60 cm, the visual angle between each icon and the fixation cross was 14.5 degrees, and the visual angle between each shape and the fixation cross was 5.8 degrees.

Results

The primary objective of Experiment 1 was to examine lexical activation levels for the modifier items, where the dependent measure of activation is the proportion of looks to the modifier icons. Because of this, only data from the test block on day 2 involving the modifier icons will be discussed here. Data from the object identification portion of the trials will not be discussed because in Experiment 1 items referring to the object came second in the utterance, were phonologically distinct, and were equibaised. Overall, participants performed highly accurately ($M = 94.0\%$ correct, $SEM = 0.01$). However, in order not to overestimate competitor or distracter fixations, only trials in which the participant's first mouse click was on the correct modifier icon were included in the following analyses.

It is apparent that participants learned the lexicon well enough to exhibit standard cohort competitor effects. Figure 2 shows the proportion of fixations to the modifier icons beginning 200 ms after the onset of the CVCVCV word until 1500 ms post word-onset. Given what is known about saccadic planning and what has been seen in previous studies of the time-locking between speech comprehension and eye movements, signal driven fixations are expected to begin approximately 200 ms after the onset of the icon target word. Target and phonological competitor fixations rose beginning 200 ms after the onset of the target, and rapidly separated from the phonologically unrelated distracters. Approximately 200 ms after the point of disambiguation ($M = 510$ ms), looks to the cohort began to decline as the target's identity became unique. At a point 1500 ms after the onset of the target word, fixations to the target modifier icon peaked across all participants and conditions, and fixations to the second named object, here one of the objects in the central workspace, began to rise above the level of unrelated icon distracters (not shown). As in Magnuson et al. (2003), we chose this 200-1500 ms region, from signal-driven fixation onset to fixation peak, as the analysis window. The foregoing summary of Figure 2 was verified statistically. For trials in which the icon's phonological cohort competitor was present, there was a significant advantage for cohort icons relative to unrelated distracter icons, $t(1, 20) = 5.01$, $p < 0.001$, $d = 1.10$

The natural language spoken word recognition literature has demonstrated that high frequency words are recognized more quickly than low frequency words, and these results have also been

observed in visual world studies (Dahan, Magnuson & Tanenhaus, 2001; Magnuson et al., 2007). We see this same frequency effect in our data as well (Figure 3). Since half of the items had a trained frequency five times higher than the other items, it was possible to compare fixation proportions for high frequency and low frequency targets. A two-way repeated measures ANOVA (target frequency by cohort category) on target fixation proportions across the 200-1500 ms window for trials in which a cohort was present revealed a significant main effect of target frequency, $F(1, 20) = 5.42, p = 0.03$, generalized eta squared $\eta_G^2 = 0.05$.² There was no effect of cohort category ($F < 1$), but there was a marginal interaction $F(1, 20) = 3.54, p = 0.075, \eta_G^2 = 0.04$, as a result of target frequency having a stronger effect when the cohort and the target were present in opposite toolbars. Post-experiment debriefing revealed that participants were generally unaware of the categorical spatial grouping of the modifier icons. However, the marginal interaction suggests that participants implicitly learned the spatial properties inherent in the display organization. Strong initial competition from a high frequency cohort competitor not in the target's toolbar drives initial fixations away from the spatial location where the target will be found, increasing fixations to the cohort and decreasing fixations to the low frequency target relative to when both the target and competitor shared the same spatial location restrictions.

Finally, effects of absent competitors were also demonstrated in this artificial lexicon (Figure 4). Since the cohort referent was not onscreen in these trials, it was not possible to measure cohort activation directly, but it was possible to measure how the properties of an absent cohort affected fixations to the objects that were present on the screen. Given that an icon's position was fixed in one toolbar region but random within that toolbar, it was more accurate to examine fixations to the toolbar region as a whole, since participants could not initially know that only one of the referents currently consistent with the ambiguous input was displayed on that trial and might still direct eye movements toward all spatially permitted locations. In conditions where the absent cohort would appear in the opposite toolbar than the target, both toolbars are possible locations for the ambiguous word's referent. Therefore, the toolbar containing a target with an absent cross-class competitor should be fixated less than if the initial phonemes are consistent with items found in only one toolbar, as it would in the absent same-class cohort condition. A two-way ANOVA (target frequency by cohort category) on the ratio of fixations to the target's toolbar relative to the total toolbar fixations over the 200-1500 ms window showed a main effect of cohort category, $F(1, 20) = 18.52, p < 0.001, \eta_G^2 = 0.14$. There were proportionally fewer fixations to the target toolbar when the (absent) cohort icon could only appear in the opposite toolbar than when the absent cohort belonged to the same toolbar category as the target. There was no effect of frequency, $F(1, 20) = 1.87, p = 0.19$, but there was a marginal interaction, $F(1, 20) = 3.56, p = 0.074, \eta_G^2 = 0.04$. Planned comparisons showed that when the target was high frequency (and the absent competitor low frequency), the category of the cohort did not significantly affect target toolbar fixations, $t(20) = 1.36, p = 0.19$. But the presence of a high frequency, cross-category competitor, even when not displayed on the screen, significantly reduced the proportion of fixations to the target toolbar relative to same-category cohort conditions, $t(20) = 4.82, p < 0.001, d = 0.90$.

²For ANOVA effects, we report generalized eta-squared (η_G^2) as defined by Olejnik & Algina (2003) and as recommended for within-subjects designs by Bakeman (2005). This measure ranges from 0 to 1 and describes the variance accounted for by a factor in a similar way to partial eta-squared (η_p^2), but includes variability due to participants in the error term. This makes it numerically smaller than η_p^2 , but more generalizable across designs. Bakeman suggests using Cohen's η^2 guidelines for effect sizes, with η^2 of .02 defined as small, .13 as medium, and .26 as large. For t -tests, we report Cohen's d (Cohen, 1988), which is a standardized mean difference. We calculated these values from the means and standard deviations of each distribution, which includes individual participant variability to prevent overestimating effect size (see Dunlap, Cortina, Vaslow & Burke, 1996). Unlike η_G^2 , the absolute value of the d measure can vary from 0 to \pm infinity as the standard deviation approaches 0, but in practical terms Cohen (1988) provides reference values of .2, .5, and .8 as small, medium, and large effect sizes.

Discussion

Experiment 1 demonstrates that our paradigm allows participants to successfully learn both the lexical and semantic features of the novel 16-word referential environment we provide them. After only two hours of training, participants accurately selected target icons and objects when given their names. This process has become automatic enough to reveal standard lexical access effects, showing an advantage for phonological competitors over unrelated distracters and an advantage for high frequency items compared to low frequency items. Moreover, participants have implicitly learned the regularities of the physical structure of the environment, and these regularities, which we consider a form of semantic knowledge, affect the strength of competition. These effects are not solely driven by a bottom-up visual identification of the possible referents, since the semantics of the competitor, in this case spatial location restrictions, influences target activation even when the competitor is absent from the screen.

Although Experiment 1 demonstrates that contextual information (the spatial arrangement of the modifier icons) is affecting spoken word recognition, this information is insufficient to resolve the time course issues about context effects raised in the introduction, since in this experiment the contextual information was made available through the target word itself. In Experiment 2, we use a different contextual manipulation that establishes a biasing context before the onset of a target word, and then examine how this context affects target access and activation.

Experiment 2

Experiment 2 further explores the time course of integration of top-down contextual information with bottom-up phonological information during spoken word recognition by examining evidence for activation of contextually inconsistent competitors, replicating the results of Dahan and Tanenhaus (2004) with a semantic context and an artificial lexicon. As has been demonstrated in multiple investigations of spoken word recognition using the visual world paradigm (including Experiment 1), if the names of two objects are phonological cohorts, competition between these two alternatives is seen before the point of disambiguation is reached. However, both restrictive access and continuous integration perspectives suggest that suitably constraining contexts may immediately affect the consideration of contextually inconsistent alternatives, preventing or reducing activation. In Experiment 2, participants again used icon buttons to manipulate novel objects on a computer screen in response to spoken instructions. We turn the focus of our investigation from the first word in the phrase (the modifier) to the second (the object) so that we can introduce biasing contextual information through the modifier and observe its effects during the ambiguous portion of the object word. We allowed modifying icons to appear in either of the two toolbars but restricted each modifier category to acting on only one subset of perceptually similar novel objects, creating two families of objects with different “affordances.” For example, for some participants only straight-edged shapes could undergo motion modification, while only shapes with curved contours could change color or texture. Since each modifier could only be applied to the members of one object affordance family, hearing the name of the modifier reduced the set of target object alternatives to only those with contextually consistent affordance properties. For example, if a participant heard the name of a modifier causing motion and had learned that only straight-edged shapes could be moved, the set of potential object referents for the second word of the phrase could be restricted to only items referring to straight-edged shapes. If the contextual information is available for use without delay, it should have immediate effects on the activation levels of contextually inconsistent objects with names overlapping at onset with the target. This prediction is consistent with both restrictive access and continuous integration perspectives, but not with the access-selection framework.

Method

Participants—Sixteen native speakers of English who reported normal or corrected-to-normal vision and normal hearing were paid \$10/hr for their participation. None of the participants took part in Experiment 1. Data from an additional seven participants were excluded from the analysis due to poor performance during testing (< 85% correct).

Materials—Participants learned a 24-word artificial lexicon (Table 2) with one of eight random word-to-referent mappings. Eight trisyllabic (CVCVCV), phonologically unique items referred to the modifier icons. As in Experiment 1, four icons caused a selected object to experience a texture change and four caused a motion trajectory change. During training, icons were selected as targets with equal frequency and could appear in any of four positions around the perimeter of the screen.

Sixteen bisyllabic (CVCV) items referred to simple novel objects. Half of the objects were constructed using only straight lines and half were constructed using only curves. These two classes of objects served as perceptually similar “affordance families,” whose members could be acted upon by either motion or texture modifiers, but not both. For seven of the sixteen participants, only objects from the curved family could move while only objects from the straight-edged family could change texture. The affordance restrictions were reversed for the other nine participants. Four pairs of the CVCV items were cohorts that differed only by the final vowel. Two cohort pairs referred to items within the same affordance family, while in the other two pairs, one item referred to an object in the curved family and one to an object in the straight-edged family. Eight additional, phonologically unique items referred to the remaining objects, four in each affordance family.

The words were spoken as isolated tokens by a female native speaker of English and were recorded in a quiet room on a Kay Elemetrics Computerized Speech Lab (model 4300B) sampling at 11025 kHz, for use with ExBuilder, a custom software package for stimulus presentation and data collection. The mean duration of the trisyllabic items was 816 ms (range 752-862 ms), and the mean duration of the bisyllabic items was 621 ms (range 507-805 ms).

Procedure—The task used in training and testing phases was similar to that used in Experiment 1, except in the training blocks that featured both icons and objects. Although participants heard the name of the modification to be made followed by the name of the object to be modified (as in Experiment 1), they were instructed to click on the items in the reverse order, first highlighting the object to be modified and then selecting the modifier button to cause the change to occur (e.g., analogous to hearing “red square” and selecting the square followed by clicking on the red icon). This was done in order to encourage early looks to the objects and discourage early looks to the modifier icons, which might delay fixations to referent and competitor objects until after the point of disambiguation of the target object word. Again, only the first click on an object and the first click on an icon were accepted.

As in Experiment 1, training progressed in stages. In the first block on day 1, participants received a 64-trial block of 2AFC object training, followed by a second 64-trial block of 4AFC training. Participants then performed three blocks with two modifier buttons and two objects, and three blocks with four modifier buttons and four objects. On the second day, participants completed a single refresher block with two modifiers and two objects before progressing to four blocks of four-modifier, four-object training. All stimuli were presented with equal frequency, all pairings of modifier and object that were consistent with the affordance restrictions were seen, and cohort items were used as distracters on approximately 25% of four-modifier, four-object training trials. Feedback, in the form of repetition of the spoken instruction and a demonstration of the correct scene, was provided at the end of all trials during training. The first five participants saw the cohort items used as distracters on 50% of the trials,

but this had no noticeable effect on their performance during identical test blocks and so those data are combined here.

After training, all participants completed a final, 128-trial testing block. Each modifier was the target on 16 trials, each object was the target on eight trials, and affordancy restrictions were maintained throughout the test block. For each object with a phonological cohort, its cohort object appeared on the screen as a distracter on half of the trials in which it was the target. No feedback was provided during the test for the mouse click responses. During the test, eye movements were monitored using an EyeLink II eye tracker as in Experiment 1.

Results

In all of the analyses reported below, data were only used from trials in which participants' first two mouse clicks correctly identified both the modifier icon and the object ($M = 94\%$ of all trials, $SEM = 0.01$). In order to determine whether the affordancy restrictions had been learned, we examined anticipatory fixation proportions to the onscreen objects. During the anticipatory interval in which participants heard the name of the modifier but before information about the object name was expected to drive fixations (the interval from 200 to 1000 ms following the onset of the modifier word), participants reliably made more anticipatory looks to onscreen members of the family of objects that had the correct affordances for the modifier than to objects with the incorrect affordancy, $t(15) = 2.67, p = 0.017, d = 0.24$ (Figure 5), indicating that information from the modifier had been learned well enough to make predictions about the referents of the upcoming target word.

Only fixations beginning after the onset of the object word were used to examine target and competitor activation levels. As predicted, once object name information became available, cohort items appearing on the screen received significantly more consideration than unrelated distracter items. During the interval beginning 200 ms after the onset of the CVCV object target word and continuing through 1500 ms, the point at which looks to objects again peaked, we compared the average fixation proportion for cohort competitors with the average fixation proportions for the unrelated distracters. A two-way ANOVA contrasting phonological relatedness (cohort or distracter) with cohort consistency (consistent or inconsistent with affordancy expectations generated by the modifier) revealed a significant effect of phonological relatedness, $F(1, 15) = 7.49, p = 0.015, \eta_G^2 = 0.09$ (see Figure 6, collapsed across contextual consistency). Although there was no main effect of cohort consistency, $F(1, 15) = 1.25, p = 0.28$, there was a significant interaction, $F(1, 15) = 5.29, p = 0.036, \eta_G^2 = 0.08$, indicating that the affordancy restrictions generated by the modifier affected cohort competition (Figure 7). Compared to contextually consistent (afforded) cohort competitors, inconsistent (unafforded) competitors received fixations much more comparable with phonologically unrelated distracters. Planned comparisons revealed a significant difference between the cohort and the distracter when the cohort was consistent with the modifier affordancy restrictions, $t(15) = 3.38, p = 0.004, d = 1.16$, but not when the cohort had an inconsistent affordancy, $t(15) < 1$. Nevertheless, close inspection of Figure 7 shows a slight visual trend for more looks to inconsistent cohort competitors than unrelated distracters, suggesting that the consistency of the cohort competitor with the bottom-up phonological information is still affecting activation levels in spite of a contradictory context. In order to determine whether the relatively long analysis window used in the above analysis obscured a possible early interval where context did not affect the cohort advantage, we conducted additional analyses of shorter time windows by dividing the 1300ms analysis window into halves (200-850 ms, and 850-1500 ms) and quarters (200-525 ms, 525-850 ms, 850-1175 ms, and 1175-1500 ms). There were more fixations to affordance-consistent cohorts than affordance-inconsistent cohorts in all time windows, but we found no time window in which

the difference between contextually inconsistent cohort competitors and phonologically unrelated distracters was significant.

Since half of the object referents had cohorts and half did not, our lexicon contained two levels of neighborhood density, zero neighbors and one neighbor. There was a visible and numerical trend towards an effect of density on target fixation proportions (Figure 8). On trials in which no cohort was present on the screen, it appears that fixation proportions for the unique target items, the ones in low density neighborhoods, rose faster than fixation proportions for items which had cohorts and thus were in more dense neighborhoods. When fixation proportions to the target object were averaged across the 400-1200 ms time window (the interval between the point when target and distracter fixations first deviated across cohort-absent conditions and the point when target fixations peaked), however, the difference did not reach significance, $t(15) = 1.33, p = 0.20$.

Discussion

Affordancy-based anticipation is clearly seen in the looks to the objects during the modifier portion of the phrase. This confirms that biasing contextual information has been learned by the participants and is strong enough to affect behavior before the onset of the word itself. This is important since without a demonstration that contextual information is in play, it would be difficult to determine whether an absence of context effects during the processing of the object word was due to the time course of contextual integration or to the failure of our contextual manipulation to provide information that could be acquired and used effectively.

This experiment also demonstrates that semantically biasing information carried by the modifier affects cohort competition. When affordancy information was consistent with both members of an onscreen cohort pair, both items received initial consideration, with more looks to targets and cohorts than distracters. However, when the modifier restricted the possible target domain to only one member of the pair, looks to the cohort did not differ significantly from looks to the distracter, demonstrating that context reduced the behavioral manifestation of the cohort effect. It is apparent from Figure 7 that this effect begins early and persists across the entire analysis interval. This result is inconsistent with access-selection models of context integration, since there is no period where the effects of context are not seen.

Experiment 2 examined the effects of context on spoken word recognition by focusing on the activation of contextually inconsistent competitors. Although the data from Experiment 2 are inconsistent with an access-selection model, they are consistent with both the continuous integration and the restrictive access frameworks. In order to distinguish between these two models, it is necessary to examine how information about context affects competition dynamics. If context is able to prevent contextually inconsistent words from becoming active, it should reduce or eliminate the effects of contextually inconsistent neighbors on *target* activation, since lexical activation levels are affected by the number and frequency of concurrently active neighbors. In Experiment 2, we observed a trend toward neighborhood density differences, even though the difference in density was quite small. In Experiment 3, we extend the neighborhood density manipulation to interact with context and add an additional neighborhood containing both same and cross-category competitors to explore possible differential effects of contextually consistent and inconsistent neighbors.

Experiment 3

Although Experiment 2 demonstrated that contextual information present before the onset of the target word has immediate effects on competition between the target referent and a pictured cohort competitor, the exact locus for the contextual integration is still unclear. In Experiment 3, we approach this question from a different perspective. We examine the effects of context

on competition dynamics, specifically by exploring how contextually consistent and inconsistent non-displayed neighbors affect fixation proportions to target objects. This focus on target activation is necessitated by the absence of the cohort objects from the display. The three hypotheses about context integration during spoken word recognition make different predictions about the effects of a biasing context on neighborhood density effects. The access-selection framework again assumes a delay between bottom-up activation and the integration of top-down context, so context should not affect the influence of neighborhood density on target activation. The restrictive access framework suggests that context may act to pre-select lexical items consistent with the context. Thus neighborhood effects should be reduced or eliminated when the neighbors are inconsistent with the contextual constraint. The continuous integration framework suggests more moderate effects of context on neighborhood density based on the overall predictability of the contextual constraints. The addition of a neutral or non-predictive modifier in Experiment 3 allows us to directly investigate whether a contextually biasing modifier affects competition dynamics by priming and/or inhibiting competitors relative to competition in a neutral condition rather than simply searching for differences between contextually consistent and inconsistent items.

Method

Participants—Twenty-one speakers of English with self-reported normal or corrected-to-normal vision were paid \$10/session for their participation. None had participated in the previous experiments. Data from an additional two participants were excluded from analysis because of poor performance during the final test (< 80% correct responses).

Materials—Participants learned a 25-word artificial lexicon with one of four randomly assigned word-to-referent mappings. In order to manipulate the neighborhood density of the object names, we increased the number of syllables from two to three and reduced the number of syllables of the modifier names from three to two. Nine phonologically distinct CVCV items referred to modifier icons (see Table 3). As before, four of these caused motion changes and four caused texture changes. In addition, a ninth modifier icon caused the screen to be cleared. As in Experiment 2, the motion and texture modifiers were only applicable to one family of objects based on affordancy (curved vs. straight contours), counterbalanced across participants. In order to create a non-biasing condition in which the modifier did not promote or eliminate any of the objects from consideration, the screen-clearing icon could be applied to objects from either affordance family. The screen-clearing, non-biasing modifier was two times more frequent than the other modifiers during training and was paired an equal number of times with each object. The biasing modifiers were equibaised and were equally paired with all objects possessing the correct affordances.

Sixteen trisyllabic (CVCVCV) items referred to novel objects, perceptually divided into the two “affordance families”. Lexical neighborhoods of phonologically-related items were generated by creating sets of cohorts differing at the last vowel. There were eight items in relatively dense neighborhoods (two four-item sets), four items in moderately dense neighborhoods (two two-item sets), and four items in sparse neighborhoods (four unique items with no cohorts).

The structure of the moderately dense neighborhoods was designed to test the effects of bias on contextually *inconsistent* neighbors. In the moderately dense neighborhoods, one member of each cohort pair referred to a curved object and the other to a straight-edged object. This created a situation where biasing information available at the modifier rendered one of the competitors contextually inconsistent and could potentially cut the number of active items in the neighborhood in half. If context prevents inconsistent candidates from becoming active,

there will be less neighborhood competition in biasing contexts than in neutral contexts and target fixation proportions should rise.

In each four-item neighborhood, two objects were from the curved family and two from the straight-edged family. This relatively dense neighborhood was designed to examine the effects of biasing context on non-displayed *consistent* neighbors, since in addition to two neighborhood competitors with opposite affordancies, each item in a dense neighborhood also had a competitor with the same affordancy. If context promotes consistent neighbors, then looks to the target should be delayed for biasing contexts compared to neutral contexts due to the increased competition from the primed competitor. In sum, there are four possible effects of context on non-displayed neighbors: (a) no effects on neighbors; (b) increased competition from consistent neighbors and reduced competition from inconsistent neighbors; (c) increased competition from consistent neighbors but no effects on inconsistent neighbors; and (d) decreased competition from inconsistent neighbors but no effects on consistent neighbors.

Each object appeared an equal number of times during training, and a single cohort object appeared as one of the three distracters on 25% of the training trials for targets with cohort competitors. The words were spoken as isolated tokens by a female native speaker of English and the stimuli were recorded and presented to the participants in the same manner as Experiment 2. The mean duration of the trisyllabic items was 825 ms (range 790-863 ms), and the mean duration of the bisyllabic items was 643 ms (range 641-644 ms).

Procedure—Each participant completed a 75 minute session on two consecutive days. Training proceeded in a similar fashion to the previous experiments. Participants first completed two 64-trial 2 AFC blocks where only objects were present and two 64-trial 4AFC object blocks. They then progressed to two 80-trial blocks with two objects and two modifiers, and eventually to two 80-trial blocks with four modifiers and four objects to complete day 1. On day 2, participants performed one block of two object, two modifier training and then four blocks of four object, four modifier training. As in Experiment 2, participants heard the modifier name before the object name but were required to first click on the object to highlight it before selecting the appropriate modifier icon button. Feedback, in the form of repetition of the spoken instruction and a demonstration of the correct scene, was provided at the end of all trials during training. At the end of the training period, each object had been the target 61 times, each biasing modifier was the target 72 times, and the neutral modifier was the target 144 times.

At the end of the second day of training, participants completed an 80-trial final test. The neutral modifier was the icon target on half of the trials, and each of the other eight modifiers was a target on five of the remaining forty trials. Phonologically unique (low density) object referents were targets on 16 of the 80 trials, half with the neutral modifier and half with biasing modifiers. Items in the medium-density neighborhoods were targets on 24 trials. On twelve of these trials, a biasing modifier was used, and the neutral modifier was used on the other twelve trials. For each subset of twelve trials, on four trials (one per item), the target's phonological cohort was present on the screen as a distracter. Items from the densest neighborhoods were targets on forty trials, half with biasing modifiers and half with the neutral modifier. On each subset of twenty trials, a single cross-category cohort was present as a distracter on four trials and a single same-category cohort was present on eight trials. None of the phonologically related items served as distracters during the other eight trials. During the test, no feedback was provided. Mouse clicks were recorded and eye movements were sampled at 250Hz by the EyeLink II eyetracker as described previously.

Results

As in Experiment 2, the analyses were restricted to trials in which participants correctly identified both the modifier and the icon ($M = 92\%$ of all trials, $SEM = 0.02$) with the first two

mouse clicks. During the 200 to 1000 ms interval during which participants heard the names of the two-syllable modifier but *before* information about the three-syllable object name was expected to drive fixations, participants were significantly affected by biasing information. They made more anticipatory looks to onscreen members of the target family when the modifier they were hearing had restrictive affordances than they did when the neutral modifier was heard, $t(20) = 2.10$, $p = 0.048$, $d = 0.32$ (Figure 9), again indicating that the implicitly defined contextual restrictions had been learned and were available for use before the onset of the object word.

The general pattern of effects seen in cohort-present trials in Experiment 2 were replicated in Experiment 3. Once object name information became available, cohort items appearing on the screen received significantly more consideration than unrelated distracter items (Figure 10) when fixations beginning after the onset of the target word are considered. During the interval beginning 200 ms after the onset of the target object and continuing through 1500 ms, the point at which looks to objects again peaked, cohort competitors had a higher average fixation proportion than phonologically unrelated distracters. A three-way ANOVA (bias, phonological relatedness, and cohort affordancy) revealed a significant main effect of phonological relatedness, $F(1, 20) = 8.47$, $p = 0.009$, $\eta_G^2 = 0.09$.

The three-way interaction did not reach significance, $F(1, 20) = 2.25$, $p = 0.15$. However, because the trend was in the predicted direction, and it replicated the pattern seen in Experiment 2, we conducted separate ANOVAs for the same and different affordancy conditions. Cohorts in the same affordancy family as the target had a higher average fixation proportion than unrelated distracters, regardless of whether the preceding modifier was neutral or biasing (cohort, biasing modifier $M = 0.15$, $SEM = 0.02$; unrelated, biasing modifier $M = 0.07$, $SEM = 0.01$; cohort, neutral modifier $M = 0.15$, $SEM = 0.02$; unrelated, neutral modifier $M = 0.07$, $SEM = 0.01$). An ANOVA examining this relationship between phonological relatedness (cohort or distracter) and context (biasing or neutral) showed a main effect of phonological relatedness, $F(1, 20) = 26.64$, $p < 0.001$, $\eta_G^2 = 0.32$, but no effect of bias and no interaction (both $F < 1$), indicating normal cohort competition when both the target and cohort were equally consistent with context. On the other hand, the type of context provided by the modifier did affect fixations to the cohort when the cohort and target belonged to different affordancy families (Figure 11). For these different-affordancy cohorts, there was a marginal interaction between bias and phonological relatedness, $F(1, 20) = 4.20$, $p = 0.054$, $\eta_G^2 = 0.05$, but there were no main effects (both $F < 1$) of either bias or phonological relatedness. Planned comparisons showed that a biasing context significantly reduced cohort activation, $t(20) = 2.13$, $p = 0.046$, $d = 0.58$, but did not affect distracter activation, $t(20) = 1.07$, $p = 0.3$. As in Experiment 2, competition from contextually inconsistent cohorts was reduced in the presence of a biasing context, though the fixation curves again suggest small but non-significant cohort effects may remain.

The key question in this experiment was whether neighborhood density effects were modulated by biasing contexts, especially since such neighborhood effects are carried by non-displayed competitors. We asked whether context had a pre-selective effect by examining fixation proportions to the target object in the 600-1400 ms time window, which corresponds to the interval between the point when fixations to the target first deviated from fixations to the distracters across cohort-absent conditions and the point when fixations to the target peaked across cohort-absent conditions.

In order to determine whether context can reduce or eliminate competition from non-displayed, contextually inconsistent neighbors, we compared items from sparse (0 neighbors) and medium neighborhoods (1 cross-category neighbor), as seen in Figure 12. Average target fixation proportions were entered into a 2×2 repeated measures ANOVA (biasing/neutral context,

sparse/medium density). There was a significant effect of density, $F(1, 20) = 8.23, p = 0.002, \eta_G^2 = 0.05$, no effect of bias, $F(1, 20) = 1.64, p = 0.22$, and most importantly, no suggestion of an interaction with bias ($F < 1$). Planned comparisons revealed that bias did not significantly affect fixation proportions for targets from sparse or medium neighborhoods (both $t < 1.0$), indicating that the effects of contextually inconsistent competitors in the medium neighborhoods were not reduced. This result is consistent with continuous integration models, which predict immediate effects of contextual constraints, without reduced sensitivity to the bottom-up input, but is incompatible with restrictive access models.

In order to determine whether context can increase competition from consistent neighbors, we examined the effect of biasing and neutral contexts on looks to targets in dense neighborhoods where one neighbor had the same affordancy as the target and two neighbors had a different affordancy. If context primes consistent candidates, there should be more competition from the same-affordancy neighbor in biasing contexts than in neutral contexts, and therefore delayed looks to the target along with lower fixation proportions. Indeed, target objects had a mean fixation proportion of 0.27 ($SEM = 0.03$) when a biasing context was present and 0.35 ($SEM = 0.03$) in the neutral context. A paired t-test reveals that the biasing context significantly decreased the proportion of target fixations relative to the neutral context condition, $t(20) = 3.29, p = 0.004, d = 0.64$, which is consistent with the hypothesis that biasing contexts can prime related words.

One unexpected result was that, for neutral contexts, looks to targets in moderate and dense neighborhoods did not differ, $t(20) = 0.94, p = 0.36$. There are two plausible explanations for why dense neighborhoods did not generate measurable effects of greater competition, both of which need to be explored in future research. One possibility is that there is a tradeoff between phonotactic probability, which might facilitate recognition (Vitevitch & Luce, 1999), and neighborhood density, which interferes with recognition. A second possible explanation is that neighborhood density, like many other probabilistic effects, operates on a logarithmic scale. In this case, the difference between having 0 and 1 neighbor would be greater than the difference between having 1 and 3 neighbors. Future work will need to use a much larger lexicon in order to resolve this issue.³

Discussion

In Experiment 2, the affordancy restrictions carried by the modifier created strong contextual constraints on the identity of the upcoming word. Anticipatory effects were observed in the increased fixations to afforded objects relative to objects that had inconsistent affordance properties. As in Experiment 2, the results of Experiment 3 again demonstrate that the context set up by the modifier affordancy restrictions was strong enough to affect anticipatory eye movements to potential target objects before the onset of the target word. Furthermore, the neutral modifier acted as a baseline and provided a stronger test for the claim that the context is biasing, since we can now demonstrate that the eventual target objects receive more consideration when there is a biasing context than when the context is neutral, as opposed to simply arguing that consistent referents are more expected than contextually inconsistent ones. If context failed to modulate cohort or neighborhood activation, it was not because the context was insufficiently strong or constraining.

Again, the results of Experiment 3 replicated the standard cohort effect. Overall, items that competed phonologically with the target were more highly activated than phonologically inconsistent items. Biasing context, however, reduced this effect when the cohort belonged to a contextually inconsistent affordancy family. A cohort with a different affordancy than the

³We thank Florian Jaeger for suggesting this interpretation of the density effects we observed for medium and dense neighborhoods.

target received consideration when a neutral context did not act to support one family or the other, but when the modifier provided information about the affordancy of the upcoming object referent, the cohort received significantly fewer fixations. As in Experiment 2, a context that rendered one member of a cohort pair impermissible resulted in reduced cohort effects, with no suggestion that there was a temporal interval where context was not affecting fixations.

But the key question examined in Experiment 3 was whether a context that both manifests itself before the onset of the target word and is strong enough to reduce the competitiveness of inconsistent cohorts present in the set of displayed objects prevents bottom-up activation of competitors in the first place. The neighborhood density manipulation in Experiment 3 allowed us to assess the effects of contextual constraints on competition dynamics in the absence of displayed competitors. It is clear that the contextual information available to the participants before the onset of the target word did not prevent an inconsistent competitor from becoming active, since a biasing context did not act to reduce neighborhood competition for items with competitors with different affordancies. This is inconsistent with a strong restrictive access viewpoint.

There was significantly less target activation, seen though decreased fixation proportions to the target object, when a biasing context was present and the target object had both unafforded and afforded competitors. In competition models of spoken word recognition (e.g., TRACE, McClelland & Elman, 1986 and Shortlist, Norris, 1994), competition from a neighbor will increase as a function of its activation. Thus, it is plausible that two strongly activated neighbors may compete more strongly with each other than four weakly activated neighbors. If we assume that the afforded competitor and the target were primed by the context, then the primed neighbor would compete more strongly with the target and slow the rise of target activation relative to the neutral context. These results suggest that context can influence neighborhood effects by making consistent lexical candidates more accessible.

General Discussion

The goal of the current experiments was to use an artificial lexicon in conjunction with the visual world paradigm to explore the scope of contextual constraints on spoken word recognition. Using an artificial lexicon allowed precise control over the nature of the contextual constraints and the properties of the lexicon. Experiment 1 demonstrated cohort effects, with increased looks to phonologically related items relative to phonologically unrelated items; effects of target frequency, with earlier looks to high frequency targets compared to low frequency targets; and effects of non-displayed competitors, with higher frequency neighbors having stronger effects than lower frequency neighbors. Experiment 2 examined evidence for activation of contextually inconsistent alternatives. We first observed anticipatory looks to objects that were consistent with the affordancy constraints established by the preceding context, and further demonstrated that these contextual constraints reduced, but did not completely eliminate, looks to cohort referents with incompatible affordances. Experiment 3 replicated the effects reported in Experiment 2 while extending the approach to examine the effects of non-displayed neighbors on competition dynamics. Contextual constraints increased the degree of competition from contextually congruent non-displayed neighbors, but did not reduce competition from contextually incompatible neighbors.

In the access-selection framework, the delay imposed by the architecture of the system prevents contextual constraints from initially affecting the evidence for a lexical candidate. When the bottom-up evidence for two candidates is equal, there will be initial effects of both candidates, regardless of contextual constraints. Context should not modulate neighborhood effects, but it should also not affect cohort competition. Our results are inconsistent with the access-selection model because there is no evidence for delayed effects of context, especially with displayed

competitors. One possible argument is that the effects of context are indeed delayed relative to the effects of bottom-up neighborhood access, but that eye movements are only sensitive to a late stage in processing and thus obscure this difference. This is inconsistent with evidence that eye movements are sensitive to effects of non-displayed neighbors, as demonstrated in the current experiments, in previous experiments with artificial lexicons (e.g., Magnuson et al., 2003), and in studies manipulating neighborhood and cohort density with English stimuli (Magnuson et al., 2007).

In a strong restrictive access model, the top-down effects of context are initially weighted heavily because they are available before the onset of bottom-up input. This predicts strong early effects of context, including contextual modulation of neighborhood effects. Under these conditions, the effects of contextually inconsistent neighbors will not be seen. The results from the current studies are problematic for restrictive access accounts, because we do observe effects of inconsistent neighbors in biasing contexts.

Continuous integration models assume that evidence is weighted according to its strength and reliability and immediately integrated. Effects of context should therefore not be delayed. At the same time, competition from inconsistent neighbors should emerge even in predictive contexts. To see why, consider that a predictive congruent context provides evidence that is distributed across all of the candidates consistent with that context. In Experiment 3, each context is consistent with eight potential candidates. Thus the constraint provided by a predictive context is weaker than the constraint provided by the bottom-up input (which even in the densest neighborhoods has only four potential candidates). While immediate effects of context might modulate neighborhood effects, then, it should not eliminate those effects.

The pattern of results observed in these studies is also broadly consistent with findings that have emerged from the reading literature using ambiguous words (Kambe, Rayner & Duffy, 2001). The basic pattern of results is that a biasing context eliminates or reduces processing difficulty for the homographs with equally likely senses, but not for homographs where the context biases the less frequent sense. In the results reported here, we see context affecting the processing of equibased, temporarily ambiguous words. In addition, unpublished data from our laboratory show reduced effects of context and more initial looks to displayed inconsistent cohorts when the inconsistent cohorts have higher frequencies than the consistent targets.

Recent results suggest that sleep consolidation may be necessary for new lexical items that are near neighbors of words in a participant's native language to begin to show evidence of lexical competition (Dumay & Gaskell, 2007). Since all of the data reported here was from tests performed after the second day of training, it is possible that sleep consolidation would be necessary for the data patterns reported here to arise. However, the lexicon structure and tasks used here were more similar to those used by Magnuson et al. (2003), who did test participants after the end of training on the first day of the experiment. They did find lexical competition after only one day of training and no sleep, though the effects were smaller on day 1 than on day 2. Given the similarity of our study to that of Magnuson et al. (2003), we do not think that sleep was critical for the general pattern of results, though it may have affected the size of the effects. However, in future experiments we will perform a test after the initial training session to investigate this issue further.

In summary, the current experiments make two contributions to the literature on context and spoken word recognition. First, they have clear theoretical implications for models of how context affects spoken word recognition. The combination of results across these three experiments is inconsistent with the access-selection framework, which maintains that there is an initial stage of context-independent lexical access. If this were the case, then we would have seen some evidence of an early interval in which context had no effects on early looks to the

referents of contextually inconsistent cohorts, since these candidates were as consistent with the bottom-up input as the contextually consistent cohorts. The pattern of results is also inconsistent with restrictive access models. This class of model predicts that neighborhood effects would be reduced when the neighborhood contains non-displayed competitors that are inconsistent with the context. However, we found that the magnitude of neighborhood effects that contain semantically inconsistent neighbors was not modulated by context. These results are most consistent with continuous integration models.

Second, the empirical results reported here suggest that using miniature novel lexicons and languages can help shed light on the dynamics of how context affects lexical processing. In future research it will be important to further explore the effects of context on competition dynamics by manipulating the type of similarity, for example, by examining neighborhoods where not all candidates have similar onsets or the frequency of neighbors differs, and by exploring different classes of contextual constraint, including probabilistic rather than absolute constraints. These manipulations would create conditions sufficiently subtle and quantitative to evaluate and test more formal models. It will, of course, also be important to determine whether the pattern of results obtained with artificial languages will generalize to natural language. Selecting these stimuli from English, for example, will require analyses of how neighborhoods vary along both phonological and semantic dimensions. However, as a first step, the present results demonstrate the range of effects that very strong contexts can have on lexical competition during spoken word recognition.

Acknowledgements

This work was supported by National Institute on Deafness and Other Communication Disorders Grant DC-005071 and National Institute of Child Health and Human Development Grant HD-27206 to MKT and RNA, and a National Science Foundation Graduate Research Fellowship to KPR. A portion of the data from Experiment 1 was presented as a poster at the Seventeenth Annual CUNY Conference on Human Sentence Processing, College Park, MD (March 2004), and portions of the data from Experiments 2 and 3 were presented at the Twentieth Annual CUNY Conference on Human Sentence Processing, La Jolla, CA (March 2007). We thank Jim Magnuson for helpful comments on an earlier version of this manuscript and Dana Subik for recruiting and running study participants.

References

- Alloppenna PD, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 1998;38:419–439.
- Altmann GTM, Kamide Y. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 1999;73:247–264. [PubMed: 10585516]
- Bakeman R. Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, Instruments, and Computers* 2005;37:379–384.
- Chambers CG, Tanenhaus MK, Eberhard KM, Filip H, Carlson GN. Circumscribing referential domains in real-time sentence comprehension. *Journal of Memory and Language* 2002;47:30–49.
- Chambers CG, Tanenhaus MK, Magnuson JS. Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2004;30:687–696.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd. Hillsdale, NJ: Erlbaum; 1988.
- Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments and Computers* 1993;25:257–271.
- Connolly JF, Phillips NA. Event-related potential components reflect phonological and semantic processing of the terminal words of spoken sentences. *Journal of Cognitive Neuroscience* 1994;6:256–266.
- Conrad C. Context effects in sentence comprehension: A study of the subjective lexicon. *Memory and Cognition* 1974;2:130–138.

- Cooper RM. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology* 1974;6:84–107.
- Corley M, MacGregor LJ, Donaldson DI. It's the way that you, er, say it: Hesitations in speech affect language. *Cognition* 2007;105:658–668. [PubMed: 17173887]
- Creel SC, Aslin RN, Tanenhaus MK. Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language* 2006;54:1–19.
- Creel SC, Aslin RN, Tanenhaus MK. Heeding the voice of experience: The role of talker variation in lexical access. *Cognition* 2008;106:633–664. [PubMed: 17507006]
- Creel SC, Tanenhaus MK, Aslin RN. Consequences of lexical stress on learning an artificial lexicon. *Journal of Experimental Psychology: General* 2006;32:15–32.
- Dahan D, Magnuson JS, Tanenhaus MK. Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology* 2001;42:317–367. [PubMed: 11368527]
- Dahan D, Magnuson JS, Tanenhaus MK, Hogan EM. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes* 2001;16:507–534.
- Dahan D, Tanenhaus MK. Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2004;30:498–513.
- Dumay N, Gaskell MG. Sleep-associated changes in the mental representation of spoken words. *Psychological Science* 2007;18:35–59. [PubMed: 17362375]
- Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups of repeated measures designs. *Psychological Methods* 1996;1:170–177.
- Gaskell MG, Marslen-Wilson WD. Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes* 1997;12:613–656.
- Gaskell MG, Marslen-Wilson WD. Representation and competition in the perception of spoken words. *Cognitive Psychology* 2002;45:220–266. [PubMed: 12528902]
- Kambe G, Rayner K, Duffy SA. Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory and Cognition* 2001;29:363–372.
- Lucas M. Context effects in lexical access: A meta-analysis. *Memory and Cognition* 1999;27:385–398.
- Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 1998;19:1–36. [PubMed: 9504270]
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology* 1986;18:1–86. [PubMed: 3753912]
- McMurray B, Tanenhaus MK, Aslin RN. Gradient effects of within-category phonetic variation on lexical access. *Cognition* 2002;86:B33–B42. [PubMed: 12435537]
- McMurray B, Clayards MC, Tanenhaus MK, Aslin RN. Temporal uptake of continuous acoustic detail reveals the grain of sublexical processing. submitted
- Matin E, Shao K, Boff K. Saccadic overhead: information processing time with and without saccades. *Perceptual Psychophysics* 1993;53:372–380.
- Magnuson JS, Dixon JA, Tanenhaus MK, Aslin RN. The dynamics of lexical competition during spoken word recognition. *Cognitive Science* 2007;31:1–24.
- Magnuson JS, Tanenhaus MK, Aslin RN. Immediate effects of form-class constraints on spoken word recognition. *Cognition*. in press
- Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General* 2003;132:202–227. [PubMed: 12825637]
- Marslen-Wilson WD. Functional parallelism in spoken word recognition. *Cognition* 1987;25:71–102. [PubMed: 3581730]
- Marslen-Wilson, WD. Access and integration: Projecting sounds onto meaning. In: Marslen-Wilson, WD., editor. *Lexical representation and process*. Cambridge, MA: MIT Press; 1989.
- Norris D. Shortlist: A connectionist model of continuous speech recognition. *Cognition* 1994;52:189–234.

- Olejnik S, Algina J. Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods* 2003;8:434–447. [PubMed: 14664681]
- Pirog Revill K, Aslin RN, Tanenhaus MK, Bavelier D. Temporary lexical ambiguity activates human MT/V5 for motion-defined words. submitted
- Simpson GB. Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior* 1981;20:120–136.
- Simpson GB. Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin* 1984;96:316–340. [PubMed: 6385046]
- Seidenberg MS, Tanenhaus MK, Leiman JM, Bienkowski M. Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology* 1982;14:489–537.
- Shatzman KB, McQueen JM. Prosodic knowledge affects the recognition of newly acquired words. *Psychological Science* 2006;17:372–377. [PubMed: 16683922]
- Swinney DA. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 1979;18:645–659.
- Tabossi P. Effects of context on the immediate interpretation of unambiguous nouns. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1988;14:153–162.
- Tabossi P, Colombo L, Job R. Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research* 1987;49:161–167.
- Tabossi P, Zardon F. Processing ambiguous words in context. *Journal of Memory and Language* 1993;32:359–372.
- Tanenhaus, MK. Spoken language comprehension: insights from eye movements. In: Gaskell, G., editor. *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press; 2007.
- Tanenhaus MK, Leiman JM, Seidenberg MS. Evidence of multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior* 1979;18:427–440.
- Tanenhaus MK, Lucas MM. Context effects in lexical processing. *Cognition* 1987;25:213–234. [PubMed: 3581728]
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. *Science* 1995;268:1632–1634. [PubMed: 7777863]
- van Berkum JJA, Zwitserlood P, Hagoort P, Brown C. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research* 2003;17:701–718. [PubMed: 14561457]
- van den Brink D, Brown C, Hagoort P. Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience* 2001;13:967–985. [PubMed: 11595099]
- Van Petten C, Coulson S, Rubin S, Plante E, Parks M. Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1999;25:394–417.
- Vitevitch MS, Luce PA. When words compete: Levels of processing in spoken word perception. *Psychological Science* 1998;9:325–329.
- Vitevitch MS, Luce PA. Probabilistic phonotactics and spoken word recognition. *Journal of Memory and Language* 1999;40:374–408.
- Wonnacott E, Newport EL, Tanenhaus MK. Acquiring and processing verb argument structure: distributional learning in a miniature language. *Cognitive Psychology*. in press
- Zwitserlood P. The locus of the effects of sentential–semantic context in spoken-word processing. *Cognition* 1989;32:25–64. [PubMed: 2752705]

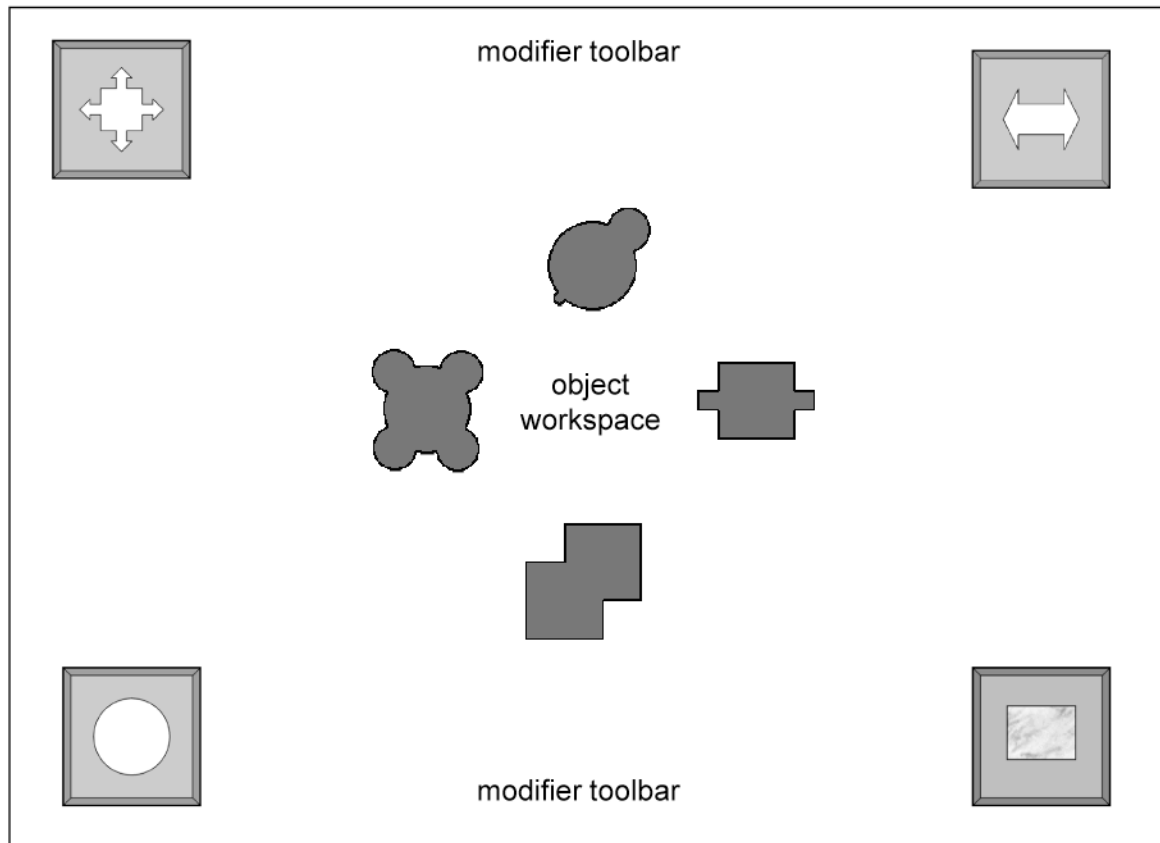


Figure 1.

Sample display used in all experiments. Modifier icons are located in toolbar regions along the top and bottom edges of the screen (modifier position is restricted in Experiment 1, and random in Experiments 2 and 3). Objects were positioned randomly in the central workspace for all experiments.

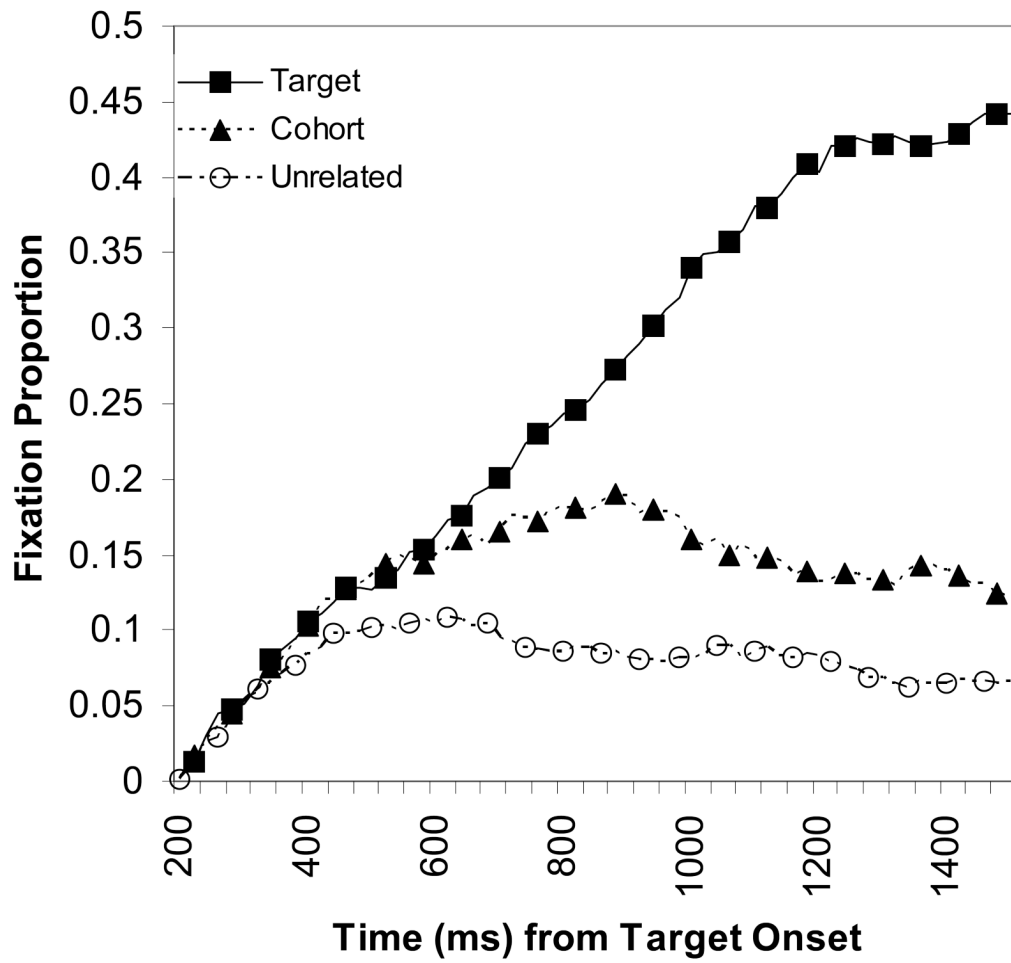


Figure 2. Fixation proportions over time to target, cohort (collapsed across toolbar), and unrelated distracter icons in Experiment 1. All fixations to modifier icons beginning 200ms after the onset of the target word are included.

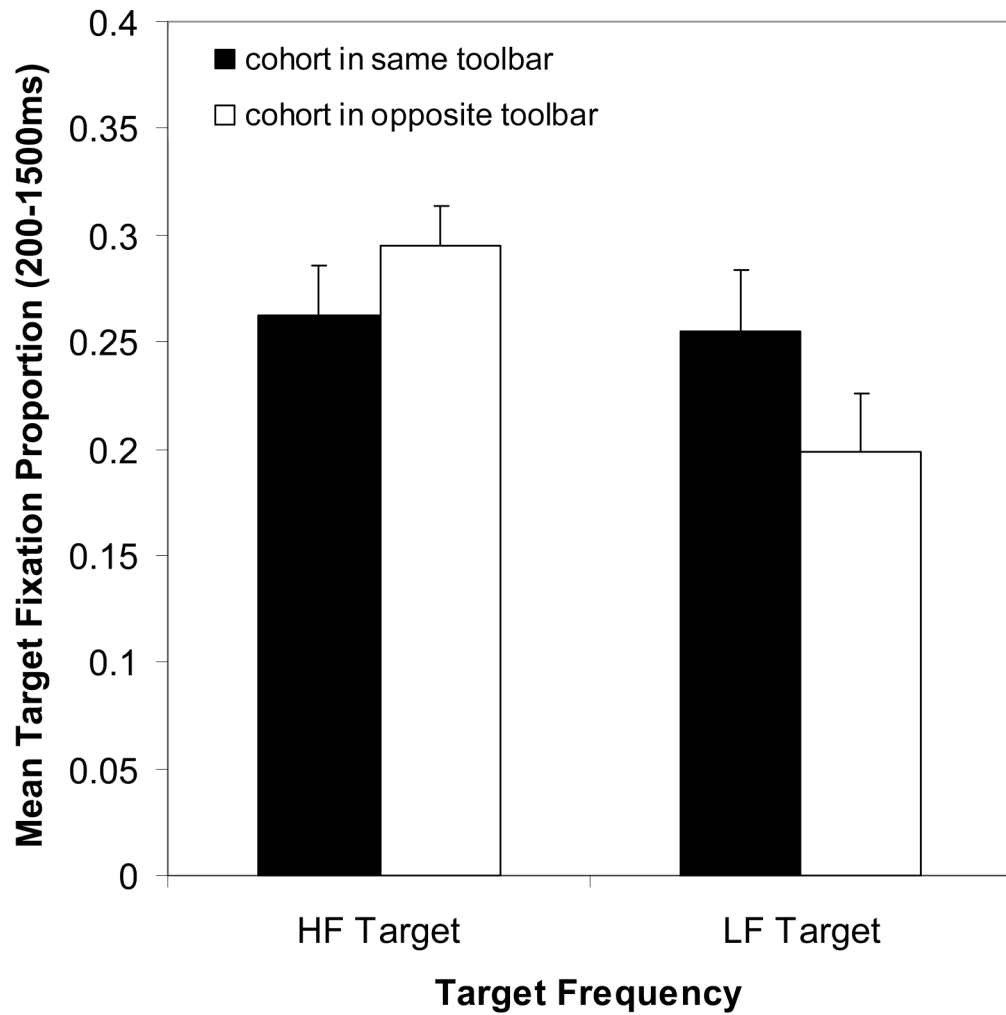


Figure 3. Mean target fixation proportions (200-1500 ms) for Experiment 1 cohort present trials. HF = high frequency. LF = low frequency. Error bars represent SEM.

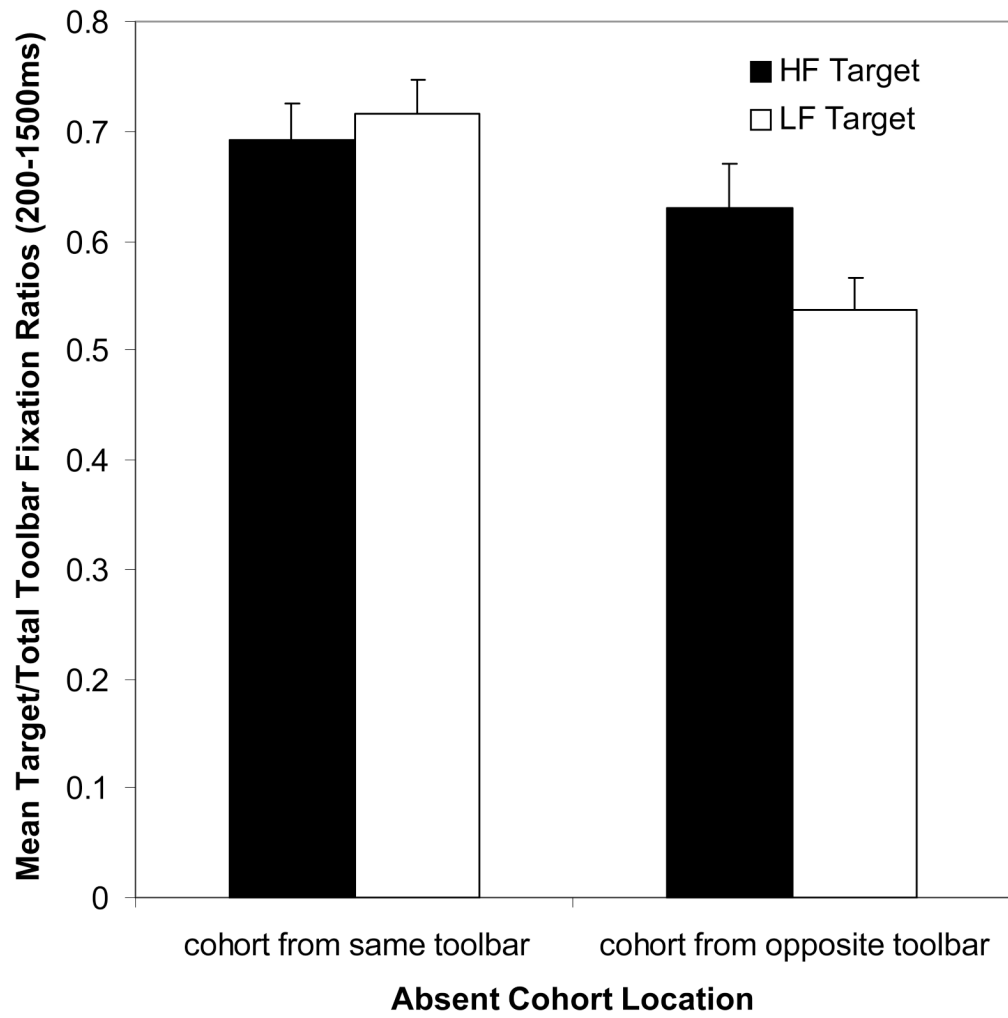


Figure 4. Mean ratio of fixations to target toolbar region to total toolbar fixations (200-1500 ms) for Experiment 1 cohort absent trials. HF = high frequency. LF = low frequency. Error bars represent SEM.

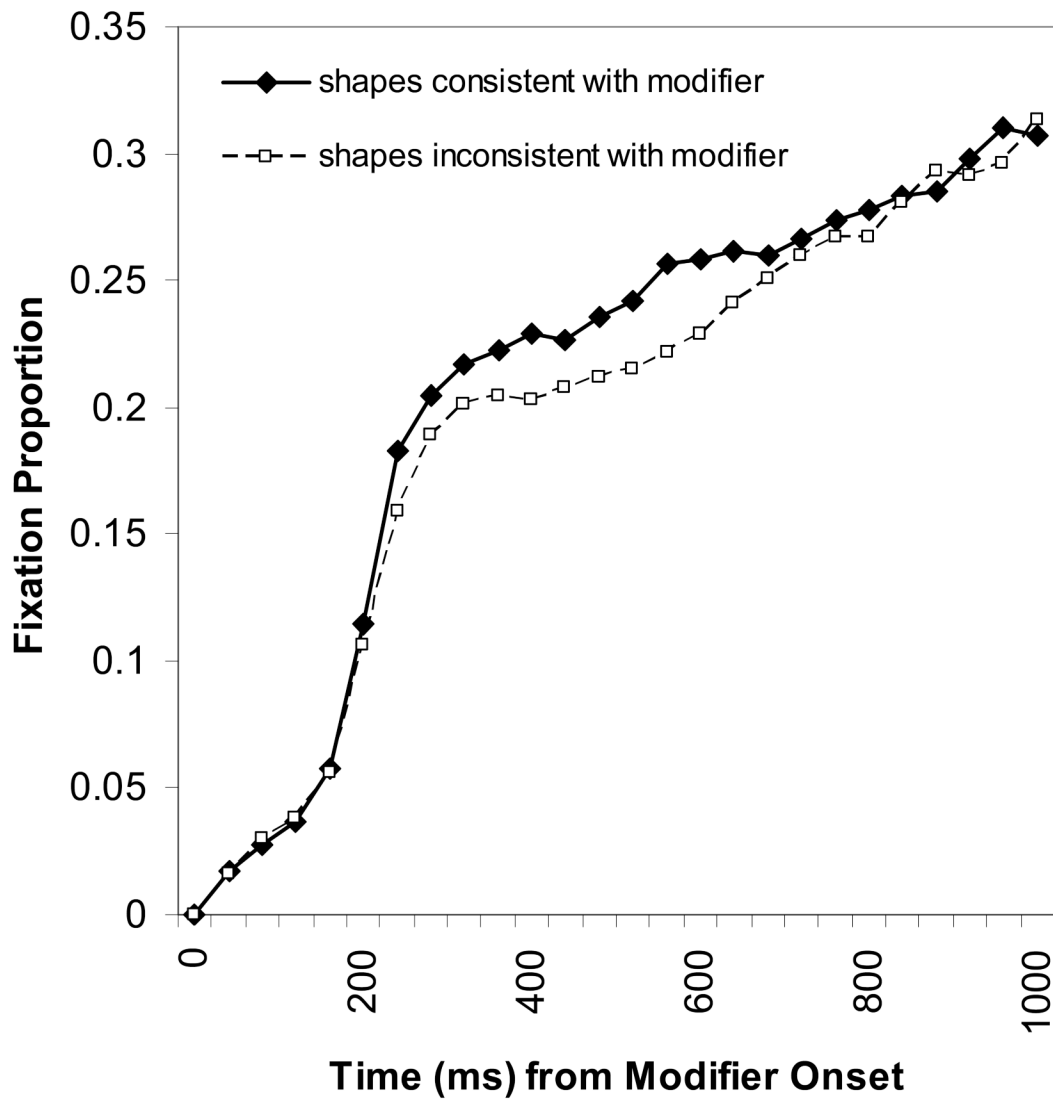


Figure 5. Anticipatory fixation proportions over time to contextually consistent and contextually inconsistent objects during the window in which only the biasing modifier is heard in Experiment 2. All fixations to objects begin 200ms after the onset of the modifier word.

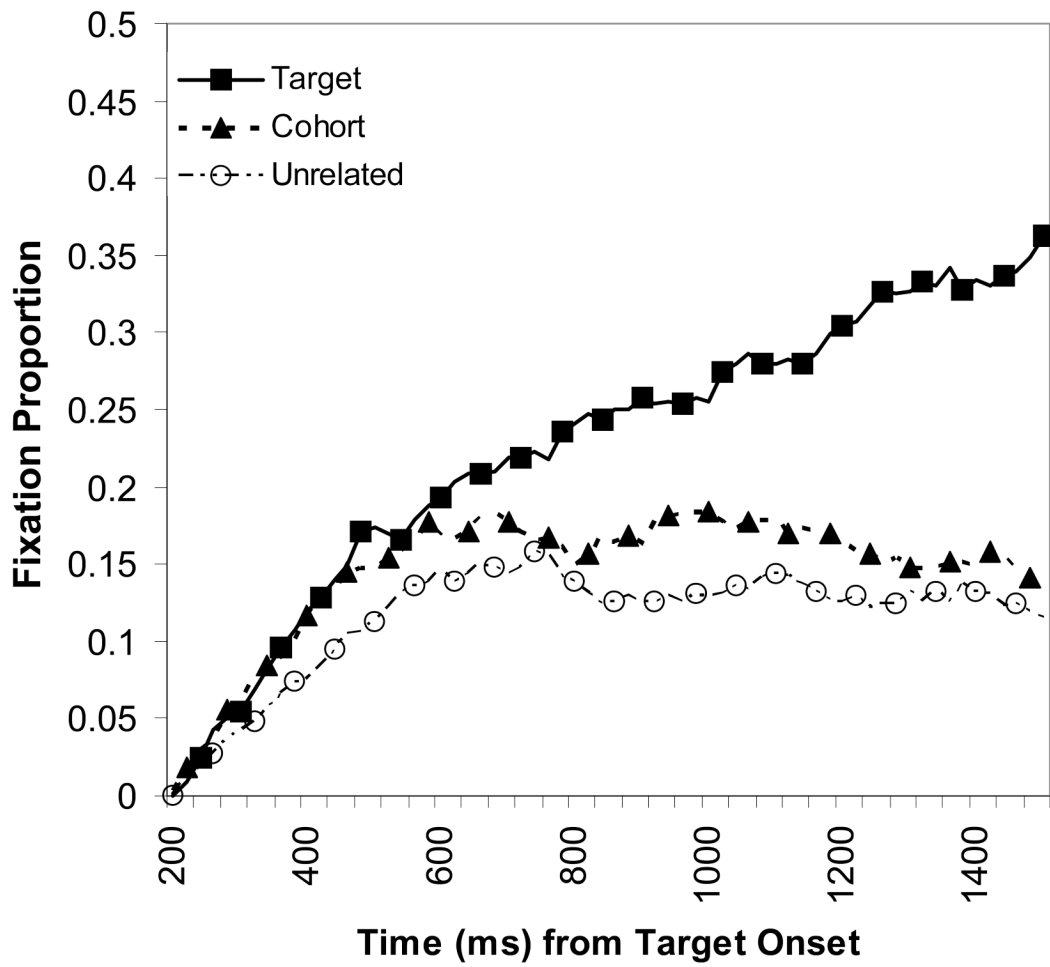


Figure 6. Fixation proportions over time to target, cohort (collapsed across consistency), and unrelated distracter objects in Experiment 2. Only fixations beginning 200ms after the onset of the object target word are included.

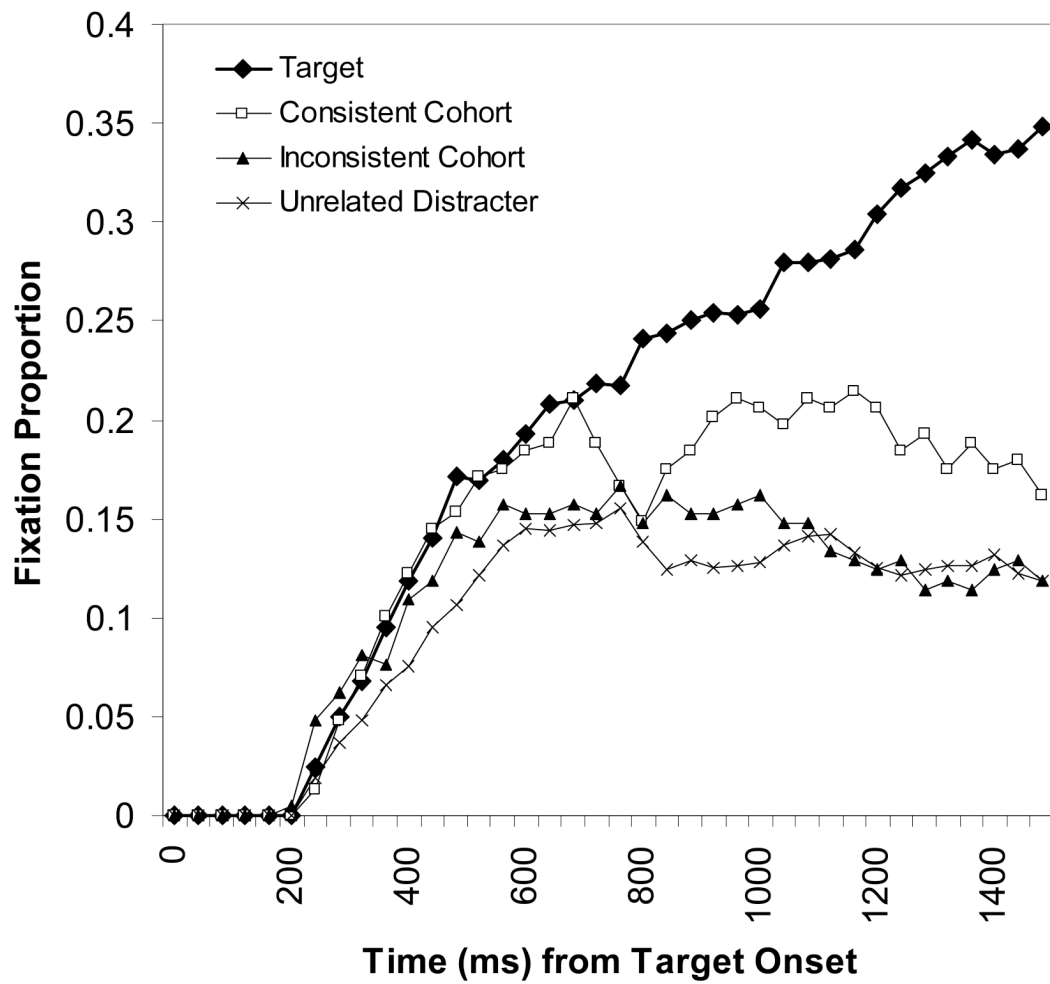


Figure 7. Fixation proportions over time to target, consistent cohorts, inconsistent cohorts, and unrelated distracter objects in Experiment 2. Only fixations beginning 200ms after the onset of the object target word are included.

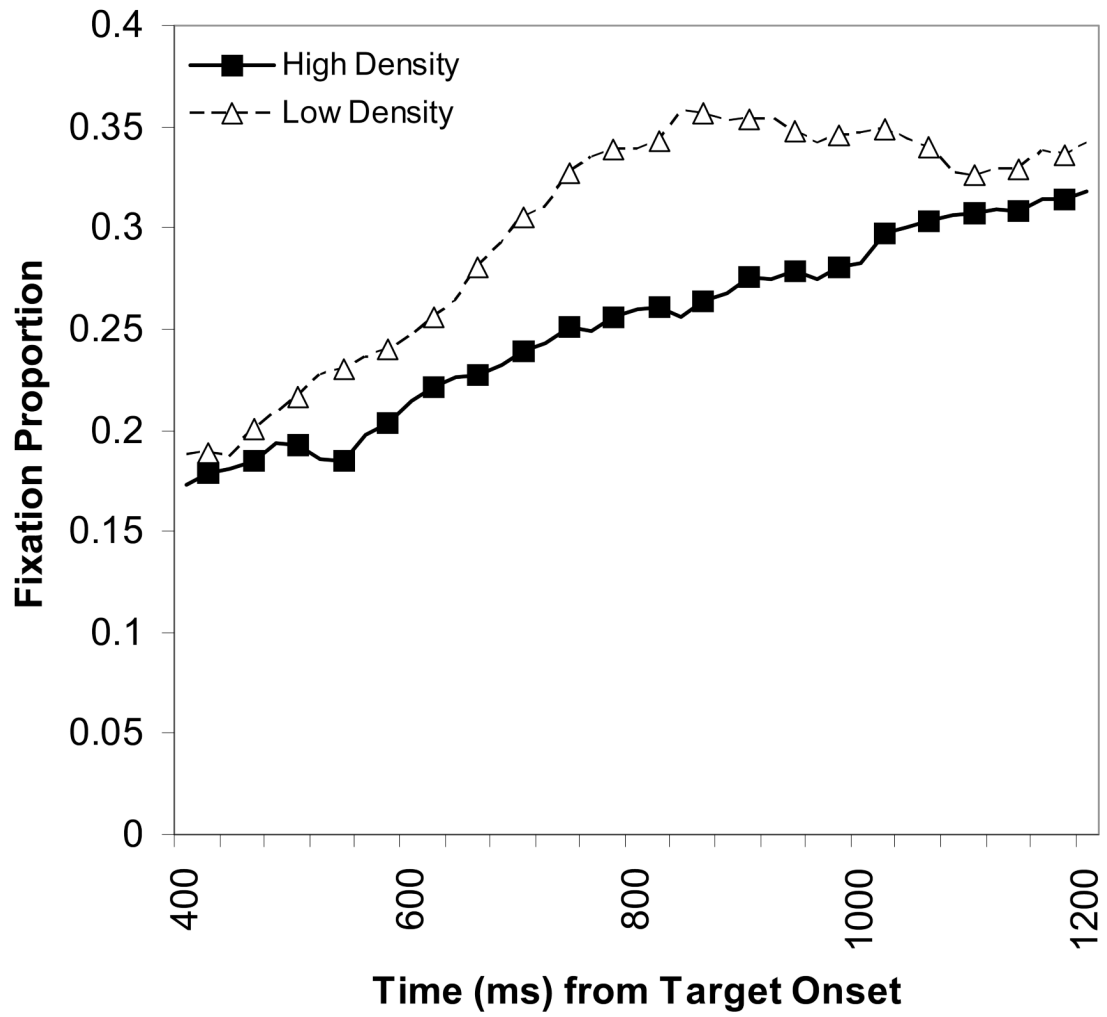


Figure 8. Fixation proportions over time to targets in dense (1 phonological neighbor) and sparse (0 neighbors) neighborhoods in Experiment 2.

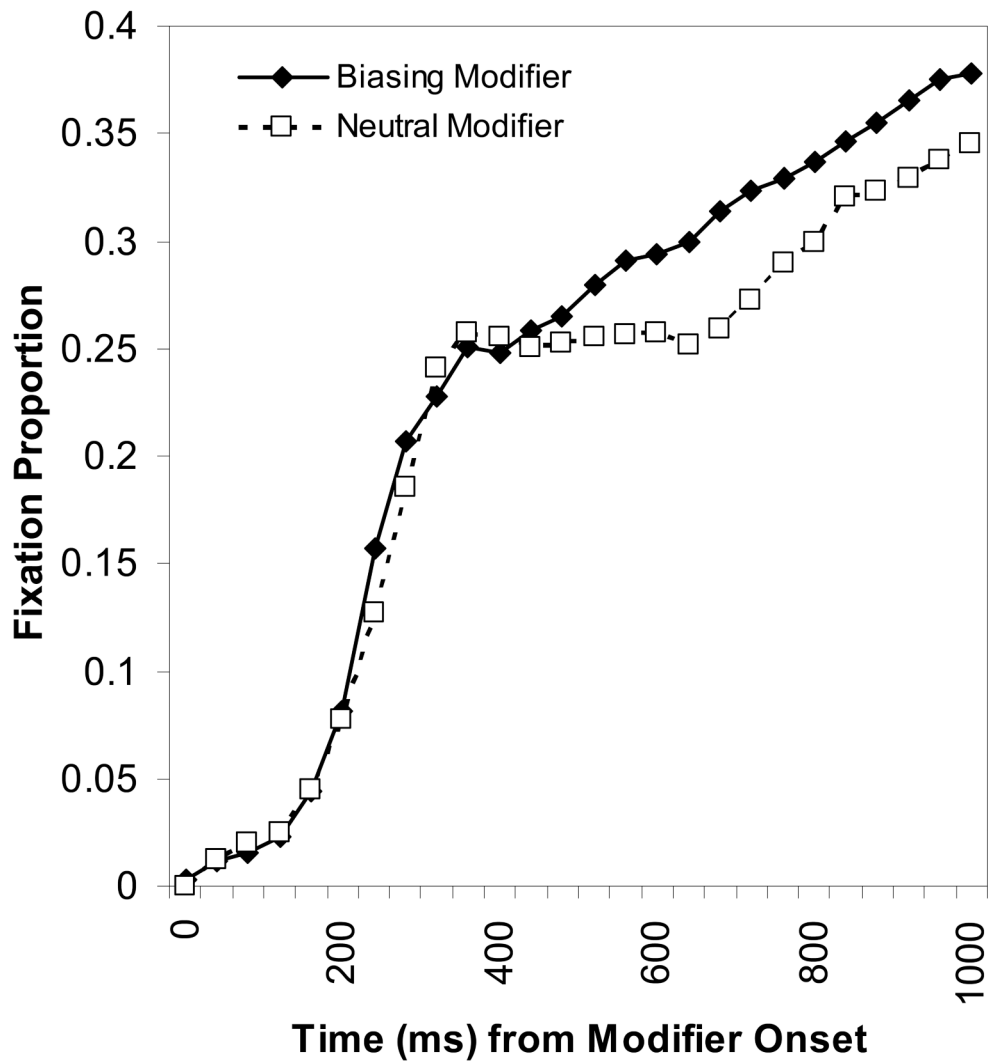


Figure 9. Experiment 3 anticipatory fixation proportions over time to objects in the eventual target family during the window in which only a consistent or inconsistent modifier is heard.

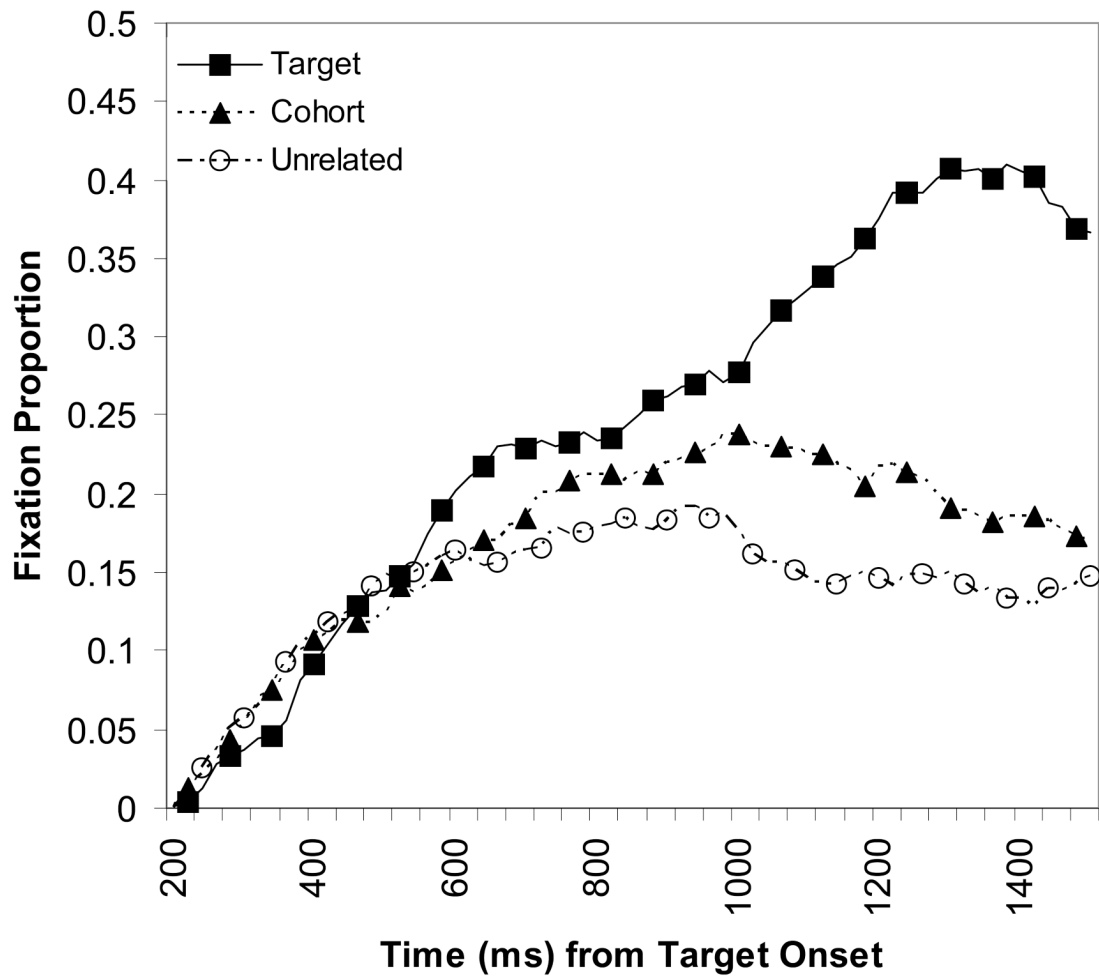


Figure 10. Fixation proportions to target, cohort (collapsed across consistency), and unrelated distracter objects in Experiment 3. All fixations beginning 200ms after object target word onset are included.

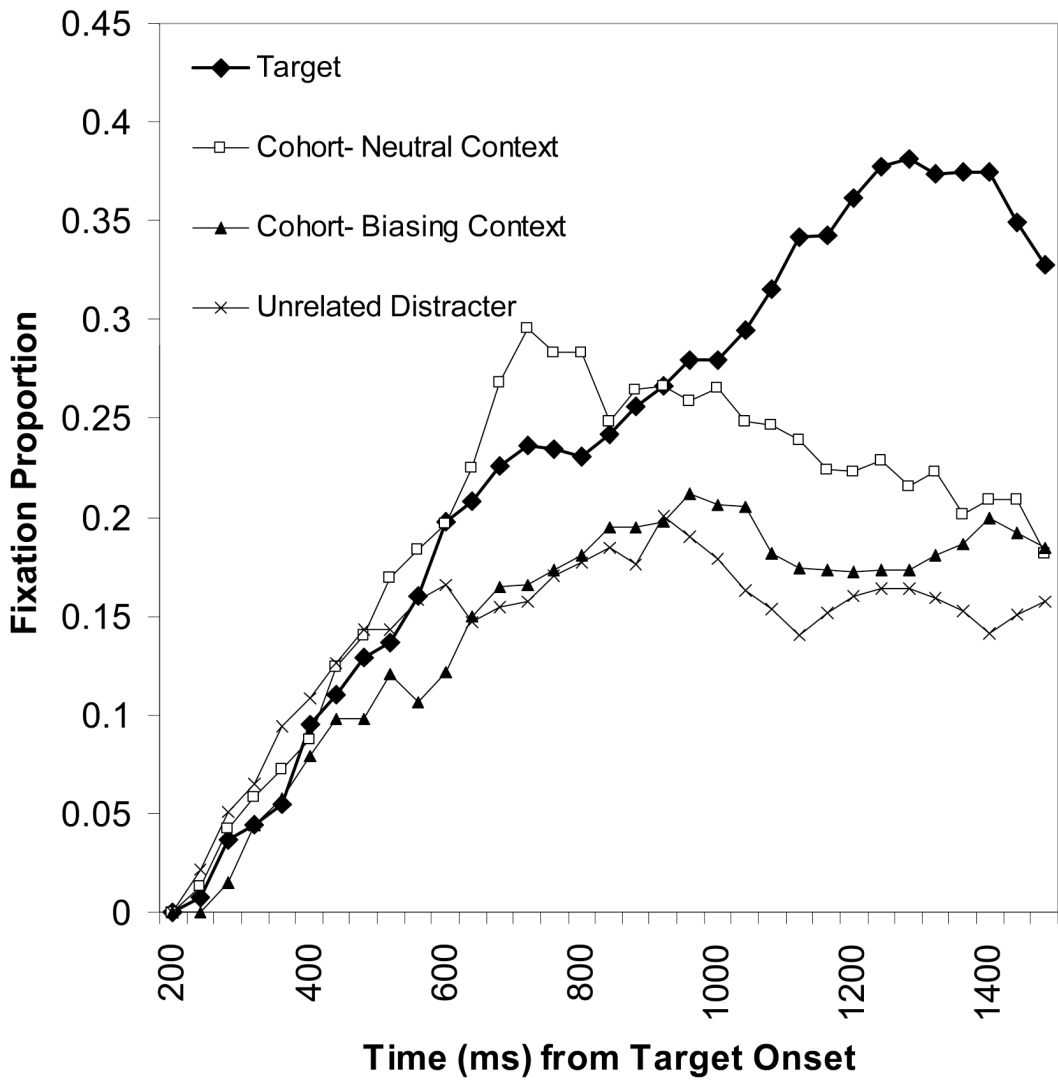


Figure 11. Experiment 3 fixation proportions over time to target, opposite family cohort, and distracter in biasing and neutral contexts. Only fixations beginning 200ms after the onset of the object target word are included.

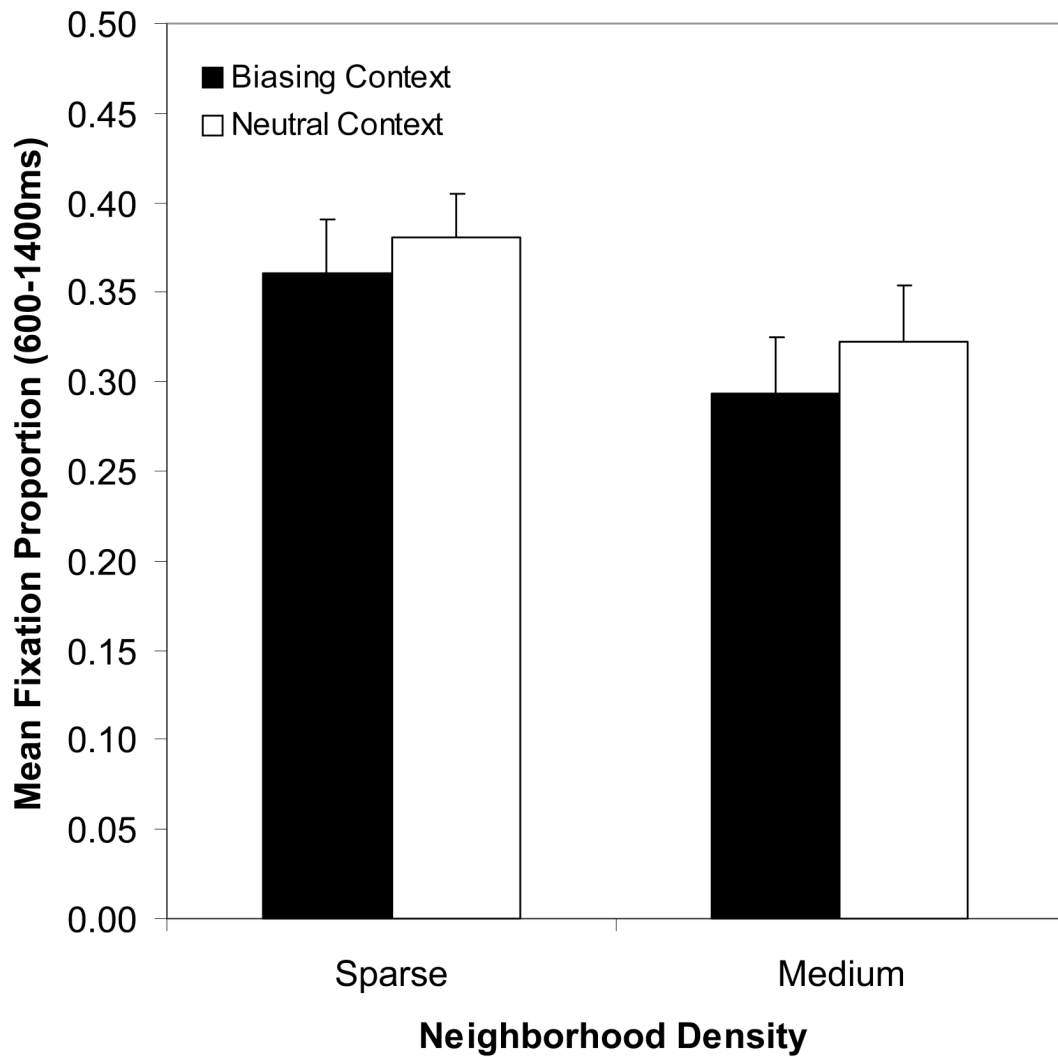


Figure 12. Mean target fixation proportions (600-1400 ms after target onset) for Experiment 3 cohort absent trials. Error bars represent SEM.

Table 1

Experiment 1 Lexicon

Icon Words	Object Words
/ku'bera/	/bupa/
/ku'bero/	/bado/
/gi'lapo/	/dibo/
/gi'lape/	/doti/
/ka'doti/	/pibu/
/ka'dote/	/pota/
/go'reba/	/tadu/
/go'rebi/	/tupi/

Table 2**Experiment 2 Lexicon**

Icon Words	Object Words
/ba'kopi/	/bupa/
/de'mito/	/bupo/
/go'deka/	/dota/
/ki'tedo/	/doti/
/mi'pare/	/pibo/
/pe'goma/	/pibu/
/ra'bige/	/tadi/
/to'rabi/	/tadu/
	/bigo/
	/bomu/
	/daga/
	/dimi/
	/para/
	/poku/
	/tiri/
	/tuko/

Table 3

Experiment 3 Lexical Items

Icon Words	Object Words
/bado/	/ku'roba/
/dira/	/ku'robi/
/guko/	/ku'robo/
/kotu/	/ku'robu/
/masi/	/ta'piga/
/pabu/	/ta'pigi/
/ripa/	/ta'pigo/
/sogi/	/ta'pigu/
/tumi/	/mi'tudo/
	/mi'tuda/
	/do'maki/
	/do'maku/
	/bi'damo/
	/go'bamu/
	/pi'gura/
	/ro'dupi/