Published in final edited form as:

Cogn Sci. 2017 April ; 41(Suppl 4): 706–747. doi:10.1111/cogs.12401.

Learning during processing Word learning doesn't wait for word recognition to finish

Keith S. Apfelbaum^{a,1} and Bob McMurray^b

^aDepartment of Psychology, Ohio State University, 1835 Neil Ave, Columbus, OH 43210

^bDepartment of Psychological and Brain Sciences, Department of Communication Sciences and Disorders and DeLTA Center, University of Iowa, E11 SSH, Iowa City, IA 52242

Abstract

Previous research on associative learning has uncovered detailed aspects of the process, including what types of things are learned, how they are learned, and where in the brain such learning occurs. However, perceptual processes, such as stimulus recognition and identification, take time to unfold. Previous studies of learning have not addressed when, during the course of these dynamic recognition processes, learned representations are formed and updated. If learned representations are formed and updated while recognition is ongoing, the result of learning may incorporate spurious, partial information. For example, during word recognition, words take time to be identified, and competing words are often active in parallel. If learning proceeds before this competition resolves, representations may be influenced by the preliminary activations present at the time of learning. In three experiments using word learning as a model domain, we provide evidence that learning reflects the ongoing dynamics of auditory and visual processing during a learning event. These results show that learning can occur before stimulus recognition processes are complete; learning does not wait for ongoing perceptual processing to complete.

Keywords

Word learning; associative learning; processing dynamics; temporal processes; lexical access

A century of research on learning offers detailed mechanistic understanding of how learners come to link various representations (what we broadly refer to as associative learning). This work has elucidated *what* sorts of things can be learned. Major branches of cognitive science study what types of associations form most easily and what constraints control the things that are learnable (e.g., Goldstone & Landy, 2010; Munakata & O'Reilly, 2003; Thiessen, 2011). Research has also uncovered much about *where* in the brain associative linkages occur. We have complex understanding of how different neural circuits accommodate different types of information, and how this information is stored within neural circuits (J. H. Freeman, 2010; O'Reilly & Norman, 2002). Further, we know *how* associations are built; that is we have clear theories and models of how information is combined to strengthen and weaken associations (Hebb, 1949; Rescorla & Wagner, 1972; Rumelhart, Hinton, &

¹Corresponding author: apfelbaum.3@osu.edu, (614) 248-1242.

Williams, 1986) and of the complex emergent products that arise when many such associations are built simultaneously (O'Reilly & Munakata, 2000; Seidenberg & McClelland, 1989; Wasserman, Brooks, & McMurray, 2015).

However, we do not yet know *when* learning occurs. It is clear that learning does not link raw perceptual inputs (if there is such a thing) to each other; rather stimuli are processed, encoded, represented, and/or categorized either during learning or prior to it. Such processes unfold dynamically over time (Spivey, 2007). What is not clear then is when—during these dynamically unfolding processes—associative linkages are formed. Are learning systems intelligent enough to wait until this processing is complete before forming associations? Or is learning much simpler and “always on,” forming associations while dynamic processing is ongoing?

Much prior work on associative learning has employed unambiguous stimuli such as lights and tones or basic visual stimuli like Gabor patches. For these stimuli, processing is extremely rapid, and processing dynamics are relatively simple. As a result, theories of learning based on these results did not need to address the dynamics of stimulus processing – such theories can assume immediate stimulus identification without compromising their fit to the data, or they can assume that any processing dynamics simply add noise to the system.

However, more complex stimuli require more protracted, complex processing, especially when these stimuli are themselves temporally extended. When perceiving sequential stimuli like musical sequences or words, learners take in information across time, with multiple representations competing for recognition as the input is received (Bharucha, 1987; McClelland & Elman, 1986). Even for simple perceptual decisions, single-unit recording studies with non-human primates show ongoing competition between multiple representations at the neuronal level (Shadlen & Kiani, 2013), and human studies using techniques like mouse-tracking and ERPs show similar unfolding dynamic competition (Huettenlocher & McMurray, 2010; McKinsty, Dale, & Spivey, 2008; Miller & Hackley, 1992). From the earliest moments of stimulus processing, the decision making system considers multiple candidate interpretations of the input in parallel.

This temporally extended processing creates a challenge for theories of associative learning: a unitary or discrete interpretation of a stimulus (e.g., a single choice out of multiple competing choices) may only be possible late in processing (if at all). How does the learning system know when processing is done so that representations can be linked? If learning occurs before these recognition processes are complete, the learner might form multiple partial associations with partially activated representations. Whether learning occurs before processing is complete or whether it waits could ultimately have a large impact on *what* learners end up linking, because different representations may be active at different times.

Spoken words offer a clear example of this problem. At early moments of lexical processing, multiple words can match the input. For example, after hearing the acoustic sequence “ba-”, the input is consistent with *bag*, *battle*, *ballast*, and so on. Listeners do not wait until ambiguity resolves to begin activating words. Instead, multiple words are activated in parallel, and identification proceeds through competition between these multiple candidate

words as more input is received (Allopenna, Magnuson, & Tanenhaus, 1998; Gaskell & Marslen-Wilson, 2002; Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, 1987; McMurray, Clayards, Tanenhaus, & Aslin, 2008; McQueen, Norris, & Cutler, 1994). This competition can be altered by fine-grained detail in the bottom-up acoustic signal (Andruski, Blumstein, & Burton, 1994; Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Marslen-Wilson & Warren, 1994; Salverda et al., 2006), expectations set up by preceding knowledge (Salverda, Kleinschmidt, & Tanenhaus, 2014), knowledge of the phonotactics of the language (Magnuson, Dixon, Tanenhaus, & Aslin, 2007), top-down information about likely words in a given context (Connine, Blasko, & Hall, 1991; Dahan & Tanenhaus, 2004), and active inhibition among competing words (Dahan et al., 2001; Luce & Pisoni, 1998). Word recognition thus exhibits complex dynamics as multiple information sources affect the dynamics of the competition process.

This miasma of active lexical candidates raises problems for learning how to map new words onto their referents. If we assume similar dynamics of processing for new word-forms that are being learned, at what point in this processing can these word-forms be linked to referents? Does the learner only form/update the linkages between word-forms and referents after all competition is resolved, even if the learning event encourages earlier learning? How does associative learning cope with multiple activated representations whose activations fluctuate across time? Learners appear to have a moving target of association, as the correct word-form is rarely active in isolation. If learning occurs while competing word forms are co-active, this could result in spurious mappings as the phonological competitors are linked to the incorrect meaning. Delayed learning occurring after competition resolves would avoid such spurious mappings. Because the items that are active change over time, *when* learning occurs might shape *what* is actually learned.

For example, while a learner is acquiring words like *gonu* and *goba* (as in the studies below), at the onset of the word, when only “go-” has been heard, she will briefly activate both of these word-forms before the disambiguating information late in the word arrives, conforming to the normal principles of spoken word recognition elucidated for familiar words. However, if a referent is available at this time (e.g. a *gonu* is displayed), and if learning occurs before this competition resolves², she may link this referent to *both* of the activated word-forms. Practically, only after the whole word has been heard can she determine that the object should be correctly mapped to the word form *gonu* and not *goba* but if learning does not wait for this, it may be problematic. Learning before competition resolves would lead to a correct association between the *gonu* and the lexical item “gonu,” but also a spurious association between the *gonu* and the competitor (“goba”).

Certain word learning paradigms or situations may allow learning to unfold during competition, which might encourage such spurious associations. For example, word learning is often thought to invoke unsupervised or observational learning (Akhtar, Jipson, &

²Note that for explanatory purposes here we are assuming some version of a localist account in which word forms are represented as complete units of some sort (e.g., a lexeme or lemma as in (Chen & Mirman, 2012; McMurray et al., 2012). As we describe in the General Discussion however, this logic also applies to more distributed models (e.g., Gaskell & Marslen-Wilson, 1997; Harm & Seidenberg, 1999) which link meanings directly to sound fragments (phonemes / syllables). We are making no strong representational claims here.

Callanan, 2001; McMurray, Horst, & Samuelson, 2012; Medina, Snedeker, Trueswell, & Gleitman, 2011; Ramscar, Dye, & Klein, 2013; Siskind, 1996; Yu & Smith, 2007). In such unsupervised learning schemes, learning could occur at any time after the stimuli are presented. In the simplest forms of unsupervised learning (e.g., McMurray et al., 2012; Yu & Smith, 2012), associations form between any co-active representations, without regard for the state of competition within the system. Such systems predict that word-object links would form during periods of competition, when multiple word-forms are partially active, leading to multiple erroneous linkages between competing lexical items and a referent. For example, when acquiring phonologically similar novel words like *gonu* and *goba* (as in the experiments reported below), the learner will partially activate both word-forms after hearing *go-*. If learning occurs before this parallel activation resolves, each word-form will become associated with whatever referents are present.

More intelligent forms of unsupervised learning may incorporate mechanisms to attenuate or eliminate this effect. For example, learners might rely on hypothesis generation and testing. This type of learning suggests that learners form only a single hypothesized referent for a given word-form; however, after encountering evidence that this mapping is wrong (e.g., the hypothesized referent is not present when a word is heard), the learner completely overwrites the previous association. Under these schemes, the emphasis on a single hypothesis would suggest that it is likely that the lexical competition would need to resolve prior to the generation and/or evaluation of hypotheses. Consequently, learning would likely only occur when a single word-form and a single referent are selected (Medina et al., 2011; Trueswell, Medina, Hafri, & Gleitman, 2013). The strong form of this hypothesis-generation form of word learning thus predicts that parallel activation of multiple word-forms during word recognition is tangential to the word-referent learning that occurs, and that learning would likely wait for lexical competition to complete before learning can begin.

Some models using more associative forms of unsupervised learning also incorporate mechanisms that could halt learning until competition resolves, eliminating the threat of spurious associations during ongoing competition resolution. For example, activation thresholds (Samuelson, Smith, Perry, & Spencer, 2011; Schutte, Spencer, & Schöner, 2003) or competition monitors (Griffiths, Steyvers, & Tenenbaum, 2007) could be used to track the state of real-time processing to determine whether activation is stable enough for learning to proceed. These mechanisms would delay learning while multiple items compete for activation, thereby avoiding spurious linkages; learning in these architectures is only on the basis of post-competition stimulus representations. A learner utilizing such competition monitors would only learn on the basis of the items that win the competition, as learning is blocked earlier. However, there is as yet no empirical basis that would support or refute these forms of competition monitoring.

It is unknown whether word learning (or other forms of learning) contains such gating mechanisms to block learning during parallel activation, or if instead learning can occur earlier during stimulus recognition processes. The learning process might be a smart or gated system that prevents parallel associations from forming while processing dynamics are under way. Alternatively, it might operate in real-time, with associations beginning to form as soon as potential targets of association are available. Under this view, which we term

“dynamic learning” in reference to the real-time dynamics of processing, learning can occur as soon as associative targets are available even if multiple targets are coactive, and it may link many briefly considered stimuli.

The current studies assess these possibilities by investigating how associative learning between words and objects interacts with unfolding activation dynamics of word recognition. Word learning offers an ideal domain to investigate the interactions between stimulus processing and learning, as the dynamics of lexical activation are complex, but also well studied. The answer to how learning operates in the lexical domain has important implications for learning more generally; if processing dynamics and learning interact directly in word learning, it suggests a general tendency for the associative learning system to rely on very simple learning that occurs throughout processing even for complex stimuli.

We ask when during processing the links between words and referents are formed: can learning occur while multiple stimuli are active, or is it confined to operate only after competition resolves when one-to-one mappings are more readily available? We addressed this with an adaptation of a common unsupervised word-learning paradigm (cross-situational learning) that allows control over when learners have the opportunity to form associations during a trial. Participants learned partially overlapping words (*goba/gonu*) to provide a learning situation in which learners can co-activate competing word-forms during word recognition. We then manipulated the onset of the visual referents to either occur simultaneously with the word-form (providing the opportunity for both active word-forms to be linked to the referents during stimulus recognition) or after competition has completed (so that only the representation that emerged after competition resolved can be linked to the referents). After learning, we used a common variant of the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) to determine whether lexical competition was increased by simultaneous presentation, which would indicate spurious learned associations with competing word-forms. Specifically, if spurious associations were formed during learning, participants should show greater consideration of competitor items (as indicated by more fixations to competitors), even when correctly recognizing a target word. If a learner has formed a spurious association between the referent of *goba* and the word-form “gonu,” then they should show increased looks to the *goba* object when hearing “gonu.”

Prior work using a similar paradigm (Roembke & McMurray, 2016) has established that this combination of eye-tracking and cross-situational learning can detect partial spurious associations between words and objects (although this study did not investigate the dynamics of lexical competition). In this study, even as participants were clicking on the correct referent of a newly learned word, they simultaneously showed increased fixations to other plausible associates. This suggests that learners maintain multiple potential referents for a given word (see also, Dautriche & Chemla, 2014), and that both can influence behavior within a single trial as they are simultaneously clicking on one referent and showing increased fixations to another. Here we employ a similar logic to show that spurious associations between phonological competitors are indeed partially formed additional associations. If a learner is selecting the correct referent but simultaneously showing increased fixations to a phonological competitor, it would suggest that the learner has laid

down two associations. That is, this logic establishes the parallelism of the learning: two word-forms that are momentarily active in parallel during learning lead to two associations that simultaneously shape behavior in parallel at test.

Across experiments, we show evidence that learners form associations when multiple word-forms are active and competing if referents are available. Further, we show that this learning interacts with a number of different aspects of ongoing processing dynamics.

Experiment 1: Manipulating the timing of associative learning

Experiment 1 tested our primary hypothesis that learning could occur for multiple competing word-forms before lexical competition resolves. We manipulated when the visual referents were presented in a word learning task to alter when the learning system could begin forming associations. In the simultaneous condition, visual referents were presented simultaneously with the auditory words. If learning is always on, this should allow associative links to be potentially laid down immediately while competition is ongoing and multiple word forms are partially active. This learning in the simultaneous condition was compared with a delayed condition in which visual referents were presented several hundred milliseconds after the auditory words, after competition has presumably resolved on a single word-form interpretation. In this condition no associations could be formed during periods of competition, as no referents were present. This should block the spurious associations formed in the synchronous condition. Both conditions present the same information, and the duration of both auditory and visual presentation is the same for both; only the relative onset of the auditory and visual information differs between conditions.

After learning we assessed whether spurious associations with competing word-forms had developed using a variant of the visual world paradigm (VWP) to assess how strongly different items were considered (fixated) in the presence of a newly learned word. A learning system that delays learning until after competition resolves predicts no difference in the types of associations that are formed as a function of when referents are presented relative to auditory stimuli; whether competition has resolved on its own before the referent appears or the learning system waits, no associations are formed until a one-to-one mapping is available. Consequently at test, learners in both conditions should occasionally fixate the competitor (since a target like *gonu* is temporarily ambiguous with its competitor, *goba*). However these fixations should not differ between conditions as they are completely driven by the bottom up input, not by associations formed during learning.

In contrast, a dynamic learning system predicts that in the simultaneous condition learners may form parallel associations between multiple partially active word-forms and the visual referent. However, in the delayed condition, they should only form one-to-one mappings. Thus, this predicts that simultaneous presentation during training should lead to increased fixations to the competitor at test – driven by both the bottom up input and the spurious associations. We tested whether these associations had formed by comparing competitor effects after learning; spurious associations with competitors should increase the amount that competitors are considered during word recognition. Testing in this version of the VWP allows measurement of partial associations with multiple items (see also Roembke &

McMurray, 2016): if learners simultaneously click the correct referent, and while showing increased fixations to the competing referent, it would suggest that the result of simultaneous training is to establish (at least) two associations.

Methods

Participants

Participants were volunteers from the University of Iowa community who received course credit or monetary compensation for their participation. All participants were monolingual native speakers of English, and they reported no vision, hearing or neurological disorders. Data from 60 participants were included in the analysis: 30 from the *synchronous* group and 30 from the *delay* group. Data from an additional 21 participants were not included in analysis because these participants failed to learn the correct word-referent mappings during training (these participants performed worse than 75% correct during the VWP trials for either *onset* or *offset competitor* trials). The accuracy data were quite bimodal; included participants averaged over 98% accuracy, whereas excluded participants averaged 50.0%. Eight of the excluded participants were in the *delay* group, while the remaining 13 were in the *synchronous* group ($\chi^2(1)=0.85$, $p=.34$). Due to a computer error, two participants in the *delay* group received fewer than the full 320 VWP trials.

This fairly conservative exclusion criterion was necessary to interpret the eye-tracking data. Because our analysis of competitor effects is predicated on consideration of competitors when a word is recognized, we need to be certain that the correct word is being activated. For participants with low accuracy, even correct trials may be unreliable; if a participant is performing near chance, correct responses may arise because of guessing, so looking dynamics would be unrelated to accurate lexical recognition. Critically, by conditioning the analysis of fixations on an accurate mouse click decision, we can establish whether learners are actively considering two potential referents at test (the one that they clicked on, and the one they are showing enhanced fixations to).

Materials

Auditory stimuli were two-syllable consonant-vowel-consonant-vowel non-words that adhered to the phonotactic rules of English, but did not have meanings. These words included two pairs of onset competitors sharing the first CV (*busa/bure*; *goba/gonu*), as well as two pairs of offset competitors sharing the second CV (*jafa/mefa*; *pacho/lucho*). Auditory stimuli were recorded from a male native speaker of English in a sound-treated room. Stimuli were adjusted to have consistent peak amplitudes, and 100 ms of silence was added to the beginning and end of each. Twenty-six exemplars of each word were used. For each participant, a random set of 16 exemplars of each stimulus was selected for the training trials; the remaining 10 exemplars were used in testing. This meant that participants only heard each exemplar a few times over the course of the experiment, preventing them from anticipating which word they were hearing on the basis of surface variability between productions.

Visual stimuli were color photographs of eight hard-to-identify objects (Figure 1), excised from background context. The word-object pairings were randomized for each participant.

Procedure

Participants learned associations between the eight novel words and visual referents. The experiment included three phases: a word-form learning phase, a cross-situational word-referent mapping phase and a VWP testing phase. Before beginning the experiment, a head-mounted Eyelink II eye-tracking was placed on the participants head, and a standard nine-point calibration procedure was conducted.

Experiments began with a phoneme monitoring task to familiarize participants with the word-forms used in the study before the word-referent learning (Gaskell & Dumay, 2003; Kapnoula, Gupta, Packard, & McMurray, 2015). This was done to make it more likely that even by the beginning of the word-referent associative learning trials, competitor word-forms would be learned well enough to be coactive; from the first trials, when hearing *go-*, learners knew that both *gonu* and *goba* were potential completions of the item. Participants received exposure to the auditory form of the words without any accompanying meaning, and they were told that the words from this pre-exposure would be used in subsequent learning tasks. Participants heard one of the words from the study, and they were instructed to press the space bar if the word contained an “O” sound, and to do nothing if it did not. Half of the novel words used in the study had “O” sounds, and half did not. Participants completed 64 trials, hearing each word eight times in a random order. Participants had 2000 ms to register a response before the next trial began.

Next participants completed the word-referent associative learning trials. Words were trained using a passive, cross-situational learning paradigm (Yu & Smith, 2007), which allowed us to manipulate when (during lexical processing) the referents were available for association. This paradigm typically requires multiple exposures to each word in order for words to be fully learned. This provides increased opportunity for spurious associations with competitors to strengthen. On each trial, participants saw two candidate referents and heard a word. One visual stimulus was the referent of this word; the other was drawn randomly from the other stimuli, except that it was never the referent of the target’s phonological competitor; that is, if the auditory stimulus for a trial was “gonu,” a *gonu* was on the display, but the foil item was *never* a *goba*. This ensured that no bottom-up co-occurrence statistics supported mappings between a referent and the phonological competitor (they never co-occurred).

Participants initiated each trial by clicking a central point. After this, the specific events on a given trial were controlled by the timing condition, as described below. We used a variant of the classic cross-situational learning in which participants made no overt response on any of the learning trials (unlike, for example, recent studies by Trueswell et al., 2013; and Roembke & McMurray, 2016). Participants were told that each trial, one of the displayed images matched the spoken word, and that they should try to learn the correct pairings. Training included 32 blocks of eight trials (each word with its correct referent once). The location of the correct referent and the identity of the foil were randomly assigned for each trial. The timing of visual and auditory presentation varied between conditions, providing different times when participants could form associations.

For participants in the *synchronous* group, visual and auditory stimuli were presented simultaneously 100 ms after the start of the trial. The visual stimuli then remained on the screen for 800 ms. In the *delay* condition, the auditory stimulus began 100 ms after the dot was clicked, but the screen remained blank for another 1000 ms, at which point the visual stimuli were displayed. Because the longest auditory stimulus was 740 ms (including 100 ms of silence at the onset of the file), the visual stimuli always appeared well after word offset in the delay condition. As in the synchronous group, the visual stimuli then remained on the screen for 800 ms. In both groups there was an inter-trial interval of 550 ms.

After every fourth block of training trials, participants in all conditions completed four interim testing trials in which a single visual stimulus was presented along with an auditory word. Participants were told to indicate whether this word was the correct label for the visual stimulus. These trials were included to ensure that participants maintained attention throughout the experiment and to provide a coarse measure of learning during training. Each word was played four times for interim testing throughout the experiment: twice with the correct referent, and twice with a randomly-selected mismatching referent (never the referent of the phonological competitor). Participants were alerted before these trials to ensure that they were prepared to make responses. For these trials, the visual and auditory stimuli were presented simultaneously for all participants. Note that these trials provided participants in the *delay* group with a small number of synchronous trials; this could weaken differences between groups, meaning that any effects found might underestimate effect sizes that would occur with fully differentiated training. Participants clicked on “match” and “mismatch” buttons on the display to indicate whether they believed an accurate pairing was given.

After learning, participants performed a four-alternative VWP task (Tanenhaus et al., 1995) to gauge which associations (both correct and spurious) were learned. Participants were presented with four visual stimuli and a central fixation dot. Each trial contained two pairs of phonological competitors: one *onset competitor* pair and one *offset competitor* pair; pairings were randomly chosen for each participant and were consistent throughout the test trials. They clicked a dot to initiate the trial, at which point the dot disappeared. A word was played 100 ms later, and the participants clicked the referent of the word to signal what object they believed the word referred to.

The timing of the VWP test trials was identical for participants in both *synchronous* and *delay* conditions. Fixations were tracked throughout these trials to determine the degree that the target, its competitor and the two unrelated foils were considered during the trial. The dynamics of familiar word recognition predict that participants should fixate both targets and competitors (Allopenna et al., 1998). However, if word/object learning occurred before competition resolved, spurious associations with competing word-forms formed during periods of competition should elicit even greater competitor fixations, as these fixations would reflect both phonological overlap and associative linkages between the referent and the competitor word. On the other hand, if word/object learning waits for competition to resolve, then fixations should be driven by phonological overlap alone, and we should observe similar performance for participants in the *synchronous* and *delay* timing conditions of the word-referent associative learning trials; for both conditions, learning only occurs

after competition settles on a single interpretation. Item location was randomized for each trial. Participants completed 40 blocks of trials, with each word presented once per block (320 total trials).

Analysis

Data were analyzed with mixed effects models (Baayen, Davidson, & Bates, 2008; Jaeger, 2008) using the lme4 package (version 1.1–6) in R (Bates & Sarkar, 2011). We used the lmerTest function (version 2.0–6) to compute p -values. Analysis of accuracy data used a binomial linking function, and data were entered as binary values signaling accuracy on each trial.

Eye-movement data were analyzed with linear models. For analyses of target looks, the DV was the proportion of looks to the target from 500 to 1500 ms after auditory stimulus onset,

transformed with the empirical logit function ($\log \left[\frac{p}{1-p} \right]$). This onset time was selected because it coincided with the beginning of consistent competitor fixations; the stimuli contained 100 ms of silence at onset, and it takes 200 ms to program an eye-movement so this onset signals looks generated after approximately 200 ms of stimulus had been heard. The offset of the window (1500 ms) coincided with where looks to competitors and unrelated objects significantly declined. Use of other time windows showed very similar results.

For analysis of competitor fixations (our primary measure) we used a different transformation (see also Roembke & McMurray, 2016). The empirical logit (as we used for target fixations) only takes into account fixations to one object (in this case, the target). However, for looks to the competitor, it was important to evaluate the degree to which competitor looks exceeded fixations to unrelated objects to account for possible overall changes in looking behaviors. These unrelated looks are not independent of competitor looks (because if the participant is looking at the competitor she can't look at the unrelated object). Thus, we derived a closely related scaling, by computing log-odds ratios of competitor to unrelated looks ($\log \left[\frac{p_{comp}}{p_{unr}} \right]$). This shows the degree to which the competitor looks exceed the looks to the unrelated. Because there were two unrelated items in the display and only one competitor, we used the average looks to both unrelated items. Note that this analysis does not consider the number of looks to the target items in the display, as we were interested in how competitor fixations compared to fixations to items completely unrelated to the auditory stimulus.

This log-odds ratio was used as the DV for analyses of competitor effects. Log-odds ratios greater than 0 signify greater consideration of the competitor than the unrelated objects. Because this measure treats competitor looks relative to unrelated looks, it estimates competition with respect to the overall tendency to fixate any item on the display. This ensures that effects from learning are not artifacts of more general changes in the willingness to fixate any visual stimuli. Because non-zero values are needed in both the numerator and denominator of the odds ratio to calculate a meaningful log-odds ratio, we computed these log-odds ratios using the mean of competitor and unrelated looks across all repetitions of

each stimulus for each participant for this analysis (individual trials have a much higher likelihood of a participant not looking at one of the item types).

DVs for the eye-movement data were computed for each stimulus for each subject and used in a linear mixed effects model. *P*-values for individual coefficients were estimated either with the Wald Z statistic (for binomial models) or using the T-statistic using the Satterwaithe approximation for the d.f. (SAS Institute Inc., 1978), which was implemented in the lmerTest package of R. In every model, the fixed factors were timing condition during training (*synchronous*: $-.5$; *delay*: $+.5$) and word type (*onset competitor*: $-.5$; *offset competitor*: $+.5$). Participant and auditory word were included as random intercepts. This model structure was chosen based on a comparison of models of varying complexity for the primary analysis (the analysis of log-odds ratios of the competitor effect). These models started with random intercepts of both participant and item, and added slopes for participant on word-type (timing condition was between subject and so could not be a random slope), and for item on timing-condition (items varied between word types, so could not be a random slope). More complex models did not offer a better fit to the data by χ^2 test (all $p > .6$) – this pattern held throughout all experiments.

Results

Learning

The primary analysis of interest was whether competition during the VWP trials increased in *synchronous* participants relative to *delay* participants. However, to ensure that such effects result from spurious associations rather than from overall poorer learning, we needed to establish that the two groups learned effectively. We first considered the accuracy in the interim trials throughout training to determine whether there were differences in the speed or effectiveness with which participants learned the words (Figure 2) which shows that participants reached a high level of accuracy early in training and maintained it.

Statistical analyses are presented in Table 1. For our analysis of the interim trials, fixed effects in the model included *block* (coded as raw block number), *condition* (*synchronous*/*delay*), and *word-type* (*onset*/*offset*-competitor) as fixed factors in the model. These analyses showed a significant main effect of block ($p < .00001$), as performance improved throughout the training blocks. There was also a significant interaction of block and timing condition ($p = .013$); this interaction reflected an earlier advantage for *synchronous* participants in the first few blocks and a slight advantage for *delay* participants at later blocks. No other main effects or interactions approached significance (all $p > .16$). During the final block of test, both groups reached high levels of performance ($M_{\text{synchronous}} = .98$; $M_{\text{delay}} = 1.0$), suggesting that both groups learned the words to ceiling before testing began. Because of the consistently high accuracy and low variance across participants, attempts to model performance on this block produced erratic results.

To further examine whether learning differed between groups, we analyzed the accuracy of identifying the correct target during the VWP trials (Table 1). Participants were overall quite accurate (recall that inaccurate participants – those with less than 75% correct on either word type – were excluded from analysis); the remaining participants in both groups were

quite close to ceiling ($M_{\text{synchronous}}=.99$; $M_{\text{delay}}=.98$). The groups did not differ ($p=.39$). However, there was a significant effect of word-type ($p=.014$), as the *onset competitors* ($M=.98$) were slightly less accurately identified than the *offset competitors* ($M=.99$). This accuracy difference may arise because of overall greater competition between words that overlap at onset than those that overlap at offset (Alloppenna et al., 1998; Marslen-Wilson & Zwitserlood, 1989). There was no interaction of timing condition with word type ($p=.080$). This confirmed that the two groups did not differ in how well they learned the words.

Finally, we analyzed looks to the target items on the VWP task to determine whether participants showed slower or faster latencies to locate the target across the two timing conditions (an indicator of efficiency gains that may result from learning). Difficulty identifying the target could be indicative of poorer learning despite accurate responding. Figure 3 displays the average looks to the target across time during the trial. For analysis, the empirical logit of these looks was taken for the average number of looks between 500 and 1500 ms transformed with the empirical logit function. There was a marginally-significant effect of timing condition ($B=-.14$, $SE=.076$, $t(58.0)=-1.88$, $p=.066$), as the *synchronous* group showed greater fixations to the target, counter to predictions of poorer learning by the *synchronous* group. There was also a main effect of word type ($B=.093$, $SE=.018$, $t(6.0)=5.02$, $p=.0024$), with greater target looks for words with *offset competitors* than for words with *onset competitors*. There was no interaction ($B=-.048$, $SE=.029$, $t(411.7)=-1.62$, $p=.11$). Across analyses of accuracy and looking time, there was no evidence that participants with *synchronous* timing struggled to effectively learn the words. This ensures that analysis of competitor effects can proceed without concerns that enhanced competition would arise from poor acquisition of target-referent pairings. This analysis does raise the possibility, however, that the *synchronous* group may have ultimately learned somewhat better than the *delay* group (although this conflicts with the slightly higher accuracy in the *delay* group). We address this concern in greater detail after the results of Experiment 3 and in the Online Supplement, showing that these small differences in learning are unrelated to the magnitude of competitor effects.

Competition

Our primary analysis concerned the effect of timing condition during learning on competition during the VWP testing phase. Figure 4 shows the looks to the competitor and unrelated items as a function of training condition over time. It suggests that changes occurred in both competitor and unrelated looking and justifying our relative measure, the log-odds ratio of competitor looks relative to looks to unrelated objects. In a very few cases (5 out of 480; 4 from *delay* participants, 1 from *synchronous* participants), no looks occurred to either the competitor (3 cases) or to the unrelated objects (2 cases) during the time window; these cases resulted in log-odds ratios of negative or positive infinity. These cases were excluded from the analysis³. Recall that in this analysis, log-odds ratios greater than 0 signify greater consideration of the competitor than the unrelated objects. Figure 5A and B

³Mixed effects models deal well with missing data so exclusion does not impede analysis; replacing these cells with the mean log-odds ratio for that participant in that overlap location condition yielded highly similar results.

present these log-odds ratios over time; Figure 5C shows the average log-odds ratio across the time window for each condition.

This analysis (Table 2) revealed that participants were more likely to fixate the competitor than the unrelated items (as indicated by the intercept; $p=.00016$). Critically, this effect was magnified for participants in the *synchronous* condition relative to the delay group ($p=.0075$); participants who learned the words with visual referents presented during auditory processing of the words looked to the competitor more than those who saw the visual referents after auditory word offset.

The effect of word type approached significance ($p=.051$); as would be expected given previous research, there were greater competitor effects for *onset competitors* than for *offset competitors*. The interaction was not significant ($p=.76$), signaling that the increased competitor effect for the *synchronous* group was consistent for the two types of competitors.

Discussion

The increased competitor fixations in the *synchronous* group suggest that learning can occur before competition resolves, resulting in spurious associations with co-active word forms. This occurred despite the referent of a target word never co-occurring with that of its competitor; there is no basis in physical co-occurrence for these spurious associations. Instead, it appears that these associations arise because competitor words are activated during real-time lexical processing during training. When stimulus timing allows learning to occur before this competition resolves, the competing word-forms are both linked to the presented referents. Experiment 1 thus supports the dynamic-learning hypothesis that associative learning can proceed before competition is complete. Rather than waiting until competition resolves, associative links are formed while multiple representations are active, leading to multiple parallel associations.

The lack of interaction between timing condition and word type is somewhat surprising. Although the expected decrease in competition from offset competitors (Alloppenna et al., 1998) was seen (across both timing conditions), the increase in competitor effects for the *synchronous* group was equivalent for *onset* and *offset competitors*. One might predict that *offset competitors* should elicit fewer spurious competitors; during recognition, the decreased activation of these competitors should limit the strength of spurious associations that can form. Evidence from previous studies of novel word learning, however, suggests that *offset competitors* may show greater activation early in learning (Magnuson, Tanenhaus, Aslin, & Dahan, 2003), when we predict spurious associations are most likely to form. Magnuson and colleagues found that when learning novel words, rhyme competitors showed quite large competition early in learning, with very early effects showing nearly equivalent competition to cohort competitors. Because spurious associations are predicted to form through coactivation early in learning, such large *offset competitor* effects may arise because of these early strong rhyme activations.

As a whole these results appear to support a dynamic learning hypothesis. However, an alternative explanation is that these effects were simply driven by differences in how well

the items were learned which could affect the degree of competition. It is unclear if this is a viable explanation. On the interim testing trials, participants in the *delay* condition were slightly more accurate by the end of testing, though this difference did not appear on the accuracy in the VWP trials. However, target fixations (a measure of efficiency) rose marginally faster in the *synchronous* condition, providing evidence that the *delay* group may have learned more poorly. Thus, evidence is mixed concerning a difference in overall learning. We address whether differences in how well the words were learned could elicit the competitor effects seen here shortly after Experiment 3, as well as with extensive analyses reported in the Online Supplement. In short, these supplementary analyses provided no evidence that differences in the strength of learning impacted the degree of competitor effects.

Experiment 2: Updated learning during ongoing processing

Learners in Experiment 1 showed evidence that if referents were available during the presentation of the word, learning occurred during periods of lexical processing (competition). However, processing does not end at stimulus offset; instead, it continues until competition resolves. Indeed, in the *delay* condition of Experiment 1, all learning is based on the outcome of this continued processing, as no referents were available prior to offset. In the *synchronous* condition, little learning could occur during these late processing periods, as the next trial initiated shortly after the auditory stimulus ended. However, the potential for additional competition resolution could alter what associations are formed. First, if a dynamic learning system continues to update learned representations as the activation profile of the stimuli being processed continues, the spurious associations formed during early lexical coactivation could be ameliorated after competition resolves. Similarly, if learning always occurs at the end of the learning trial, additional post-trial time to continue processing can allow learned representations to reflect the post-competition activation.

Learning during this post-stimulus period would require that learners form and update associative links on the basis of working memory representations (or simply sustained activation states) of both the visual and auditory stimuli; after the stimuli complete, some memory of their forms must be retained as the learner will receive no more sensory information to support them. Note that in invoking working memory here we are not making the claim that listeners are offloading representations from a lexical processing system to a separate working memory system; rather working memory is more likely embedded within lexical processing by just sustaining activation for the word (e.g., McClelland & Elman, 1986).

More importantly, however, to the extent that the activation of the stimulus endures, this would provide an opportunity to continue learning on the basis of this activation. This working-memory-based learning is the central form of learning in the *delay* condition of Experiment 1 as well. Experiment 2 extends this to ask whether such learning based on post-competition working-memory representations of the visual stimuli overcomes the formation of spurious associations that arise during synchronous presentation from multiple co-active candidates. This is accomplished by adding a delay to the end of synchronous trials, after auditory and visual stimulus offsets. During this delay, if the visual stimuli are being

maintained in memory, they will provide the opportunity to update associations as competition resolves. As competition resolves, learning can operate on the post-competition word-form activations that were the basis of learning in the *delay* condition. This was not possible in Experiment 1 because the onset of the next trial would have disrupted any updated learning.

Methods

Participants

Thirty participants from the same participant pool as Experiment 1 contributed data to the analysis for Experiment 2. Data from an additional eight participants were not included in the Experiment 2 data because these participants fell below the 75% accuracy threshold for either *onset* or *offset competitor* trials during the VWP portion of the study.

Procedure

Experiment 2 was identical to Experiment 1 in all facets except for the timing during the word-referent training trials. As in the *synchronous* condition of Experiment 1, auditory and visual stimuli were presented simultaneously, starting 100 ms after the participant initiated the trial; the visual stimuli remained on the screen for 800 ms, after which the screen was blank. However, unlike the *synchronous* condition of Experiment 1, in Experiment 2 the blank screen period lasted 1,550 ms (which was 1,000 ms longer than in Experiment 1). This afforded participants a long period to update their learned associations after competition resolves – nearly twice as much time elapsed after visual stimulus offset as during visual presentation. All other aspects of all tasks remained the same between the experiments, and the same stimuli were used in both.

Analysis

Results from Experiment 2 were compared to the *synchronous* participants in Experiment 1. This comparison is ideal as these two groups had identical stimulus presentation until the post-trial interval, and because the *synchronous* group showed enhanced competitor effects, making them the logical comparison to see whether continued processing decreased these spurious associations. All analyses followed the same format as those in Experiment 1, and the same model structure was used (the fixed factors were timing condition during training (*synchronous*: $-.5$; *Experiment 2*: $+.5$) and word type (*onset competitor*: $-.5$; *offset competitor*: $+.5$), and participant and auditory word were included as random intercepts – more complex models including random slopes did not provide a better fit to the data by χ^2 test; $p > .1$).

Results

Overall learning

As in Experiment 1, we began by examining the overall efficacy of learning to ensure that potential changes in competitor effects could not arise because of poorer overall learning in one group or the other. Figure 6 suggests on interim test trials, participants learned the word/object mappings very well, and perhaps slightly better than the *synchronous* group of

Experiment 1. A logistic mixed effects model (Table 3) showed a main effect of block ($p < .0001$), as participant performance improved throughout training. The interaction of word type and block was marginally significant ($p = .064$), as *offset competitors* reached ceiling slightly more quickly. No other effects approached significance (all $p > .15$). Performance at the last block was very close to ceiling in all conditions ($M = .98$); attempts to model accuracy for this final block resulted in erratic behavior, as there was very little variability in performance.

Analysis of accuracy during the VWP trials (Table 3) showed that both groups were quite close to ceiling ($M_{\text{synchronous}} = .99$; $M_{\text{Experiment 2}} = .99$); the groups did not reliably differ ($p = .55$); and there was no interaction of timing condition with word type ($p = .37$). As in Experiment 1, there was a significant effect of word type ($p = .0041$), as participants were less accurate for *onset competitors* ($M = .98$) than for *offset competitors* ($M = .99$).

Analysis of target looks (Figure 7) showed greater and more robust looking to the target in the synchronous condition than in this experiment. Again we used the empirical logit transform of average looks to the target across the analysis window (500–1500 ms). This analysis revealed a main effect of timing condition ($B = -.16$, $SE = .069$, $t(58.0) = -2.26$, $p = .028$), with fewer target looks by participants in Experiment 2 than by the *synchronous* group in Experiment 1. There was also a main effect of word type ($B = .11$, $SE = .019$, $t(6.0) = 5.80$, $p = .0012$), with the expected increase in target looks for *offset competitor* items. There was no interaction ($B = -.011$, $SE = .029$, $t(411.7) = -.38$, $p = .70$). As in Experiment 1, the *synchronous* group showed greater looks to the target, suggesting that any increase in competitor fixations for this group cannot be explained by poorer learning of the correct word-referent associations.

Competition

As in Experiment 1, the critical analysis focused on the degree to which participants considered the referent of the phonological competitor of the target word during the VWP trials. We computed log-odds ratios of looks to the competitor item compared to the average of looks to the two unrelated items from 500 to 1500 ms after stimulus onset, exactly as in Experiment 1.

In this analysis (Figure 8, Table 4), the *synchronous* group exhibited greater competitor effects than participants in Experiment 2 ($p = .030$); the extra processing time mitigated spurious associations that could have formed from incremental learning during lexical processing. There was also the expected main effect of word type ($p = .038$), with greater competitor effects for *onset competitors* than for *offset competitors*. However, as in Experiment 1, the interaction with competitor type was not significant ($p = .53$).

Discussion

Experiment 2 demonstrated reduced competitor effects when training included additional time over which learners could continue to suppress competitors, compared to training in which the following trial began fairly quickly after stimulus offset, before competitors could be fully suppressed (the *synchronous* condition from Experiment 1). This pattern suggests

that although learning can occur during periods of competition, additional time to update learned associations as competition resolves can reduce spurious associations between phonological competitors and the visible referent. After the offset of the stimulus, the learner can continue suppressing competitors and altering learned associations to reflect the later state of processing. Paired with the results of Experiment 1, this suggests that while learning can occur during competition (when multiple associative targets are coactive), it can also continue as processing proceeds, on the basis of working memory representations of the stimuli. Unsupervised associative learning may not occur a single time per learning event, but instead may happen across the events unfolding both in the world and in working memory. Increased processing time allows competition to resolve and reduces spurious learning.

Experiment 3: Interactions of visual and auditory processing

Experiments 1 and 2 emphasized the way that auditory words are processed across time. Meanwhile, visual information may also influence the way auditory stimuli are processed. As a person scans a visual scene, they can use the information to constrain possible interpretations of upcoming auditory information (Chen & Mirman, 2012; Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010). For example, early word learners use mutual exclusivity to identify potential referents of new words (Markman & Wachtel, 1988); if a child sees a stuffed dog (and knows the word *dog* already) and an unknown toy and is told to choose the “dax,” the child is likely to pick the unknown toy – her knowledge that the dog is not called a “dax” lets her interpret the novel word as the name of the unknown toy. Ramscar and colleagues (2010) showed that this type of prediction of the upcoming name of an item can help learners acquire word-referent mappings more quickly (although this paper did not address real-time competition processes).

Such expectations could play a similar role in constraining the activation of competitor words in a novel word learning situation. If the learner sees two objects (e.g. a *goba* and a *pacho*) and has some idea of their labels, she can boost activation of *goba* and *pacho* before any auditory information is received. If the auditory stimulus then starts with *go-*, activation for *goba* has a boost relative to *gonu*, allowing more rapid resolution of the ambiguity. The context of the task thus provide clues to the upcoming auditory stimulus, as long as the learner has some knowledge of the correct word-referent mappings (Yurovsky, Fricker, Yu, & Smith, 2014). This speeded competition resolution will thus improve as learning proceeds – the better the learner knows the correct word-referent mappings, the better she can pre-activate the possible upcoming stimuli, and the quicker she can suppress activation of the competing word-form.

In the previous experiments, pre-activation of the upcoming words was not possible: the visual stimuli never preceded the auditory, so participants did not have the requisite time to start forming expectations. In Experiment 3, during the learning phase, we presented the visual stimuli first, to allow time for the learners to take in this information to form expectations about the upcoming auditory stimulus. That is, the timing of presentation in this experiment allowed time for the dynamics of visual processing more time to unfold, to maximally affect lexical processing dynamics.

Methods

Participants

Data from 30 participants selected from the same participant pool as the previous experiments were included in analysis of Experiment 3. Data from an additional seven participants were excluded because they did not reach the 75% accuracy threshold for either *onset* or *offset competitors* during the VWP trials. Due to computer error, three included participants did not receive the full 320 VWP trials.

Procedure

The stimuli and design were identical to the previous experiments in all aspects except the timing of the word-referent training trials. On these trials, visual stimulus presentation began 100 ms after the participant clicked the dot to initiate the trial. After 1000 ms of visual display, the auditory stimulus began. The visual stimuli remained on the screen for an additional 800 ms after auditory stimulus onset, and then the screen was blank for a 550 ms inter-trial interval.

Analysis

Data from Experiment 3 were compared against data from the *synchronous* condition from Experiment 1; other than the visual pre-scan, these two groups had identical stimulus timing, including visual presentation present during auditory stimuli, and a rapid transition to the succeeding trial after auditory offset. Additionally, this comparison affords the clearest means of investigating whether visual processing dynamics mitigate the formation of spurious associations during lexical activation. The same model structure was used as in previous studies; more complex models did not improve model fit (all $p > .13$).

Results

Overall learning

As in the prior experiments, we began by analyzing learning during the interim test trials. As Figure 9 shows learners were highly accurate and did not differ substantively from the synchronous group of Experiment 1. This was confirmed with a logistic mixed model. This showed a main effect of block ($p < .0001$), with performance improving throughout training. No other main effects or interactions approached significance (all $p > .15$), signaling a quite similar timecourse of learning between the groups and the word types. On the final block of interim trials, accuracy was extremely high ($M = .995$); participants in Experiment 3 were all at ceiling in this block, as were all *synchronous* participants for *offset competitor* items. These highly accurate levels of performance precluded statistical analysis of the final block.

Accuracy during the VWP trials was also high for both groups ($M_{\text{synchronous}} = .99$; $M_{\text{Experiment 3}} = .97$). This difference approached significance ($p = .064$); however, the Experiment 3 participants performed more poorly than the *synchronous* participants from Experiment 1, so any increased competition by the *synchronous* group can't be attributed to poorer overall learning. There was also a main effect of *word type* ($p = .022$), with higher accuracy for *offset competitors*. The interaction was not significant ($p = .37$).

Analysis of target looks (Figure 10) used the empirical logit transform of average looks to the target during the analysis window (500–1500 ms), as in the previous experiments. The two groups did not differ in target fixations ($B = -.13$, $SE = .086$, $t(58.0) = -1.55$, $p = .13$), counter to predictions that the *synchronous* group might have learned the correct target-referent associations more poorly. There was a main effect of word type ($B = .11$, $SE = .017$, $t(6.0) = 6.65$, $p = .00056$), with more target looks for *offset competitor* trials than for *onset competitor* trials. There was no interaction ($B = -.0078$, $SE = .030$, $t(410.8) = -.26$, $p = .79$).

As in the previous experiments, there was no evidence that the *synchronous* group learned the words more poorly than participants in Experiment 3. Thus increased competitor effects for this group could not emerge as a result of poorer learning.

Competition

Figure 11 shows the critical analysis of competitor effects, again showing increased fixations in the *synchronous* group. Our statistical analysis (Table 6) again used the log-odds ratio of looks to the competitor item versus looks to the average of the two unrelated items during the analysis window (500–1500 ms). Two cases from Experiment 3 produced log-odds ratios of positive or negative infinity, and so were excluded from analysis.

This analysis revealed a significant effect of timing condition ($p = .045$), with greater competitor effects for the *synchronous* participants from Experiment 1 than for the Experiment 3 participants. That is, the visual pre-scan period decreased the magnitude of competitor effects. The expected significant effect of word-type approached significance ($p = .069$), with marginally greater competitor effects for *onset competitors*. As in the prior experiments, the interaction was not significant ($p = .48$).

Discussion

When learners were given an opportunity to preview and process the visual stimuli before auditory stimulus presentation, the magnitude of competitor effects was decreased relative to participants with no such pre-scan. This decrease emerged despite the simultaneous availability of words and visual referents in the preview group (which should have promoted lexical competition was unfolding). Additionally, both groups received the same amount of time post-stimulus before the next trial began, so the Experiment 3 participants had no opportunity to continue resolving competition. Instead, it appears that the presence of the visual stimuli before auditory onset impacted auditory processing, such that competitor words were more readily suppressed.

This finding argues for an interaction between visual processing and auditory processing (Ramscar et al., 2010). Specifically, as a learner processes a visual display, visual information constrains auditory processing to emphasize the relevant word forms (see also, Chen & Mirman, 2012). The learner can form expectations that the upcoming auditory stimulus will be the name of one of the two displayed items, allowing rapid suppression of a competitor word that does not match the display. However, these expectations require that the learner has some idea what the word-referent mappings are; without this knowledge, she has no way to predict that “gonu” is a more likely auditory stimulus than “goba” on a trial

where the *gonu* is present. This requires a nuanced interplay between timescales of learning and processing: the state of knowledge of the correct word-referent mappings affects the way that expectations can be formed about upcoming auditory stimuli; these expectations affect the way that competition dynamics play out during auditory processing; and the nature of these auditory processing dynamics impact the updated learned associations between words and referents (see Yurovsky, Fricker, Yu, & Smith, 2014 for a similar account of partial knowledge constraining word learning though without the critical real-time interactions we demonstrate here). Early in learning, learners would be predicted to have little ability to form useful expectations of the upcoming auditory stimulus, and so should be more susceptible to spurious associations with competitors. The present design does not allow direct investigation of this effect, as analysis is predicated on having primarily correct trials by test. Instead, further research with other testing regimes is needed in the future.

Combined Statistical Analysis

The analyses of Experiments 2 and 3 compared the participants in those experiments against the data from the *synchronous* condition of Experiment 1. However, these may be better considered as separate conditions of a single experiment, as all comparisons are conducted against the same reference group. The experiments were presented separately as they address different theoretical questions. However, an omnibus analysis including all conditions together may be more statistically appropriate. We conducted this analysis to confirm that the effects presented held when using this more standard approach. For this analysis, we combined the log-odds ratios for competitor effects across conditions. We included condition (*synchronous*, *delay*, Experiment 2 and Experiment 3; dummy coded with *synchronous* as the reference group) and word-type (*onset competitor*: $-.5$; *offset competitor*: $+.5$) as factors, and random intercepts of participants and auditory word. As in the independent models used for each experiment, more complex models did not improve model fit by χ^2 test ($p > .85$).

This analysis confirmed the findings from the independent models (Table 7). Specifically, every condition showed a significantly larger competitor effect compared to the *synchronous* group (all $p < .033$), and there was a significant effect of word-type, with greater competitor effects for *onset competitors*. The interaction of condition and word-type was not significant for any comparison. This analysis thus confirms the findings detailed above, and ensures that effects were not a result of the statistical approach used.

Alternatively, effects from Experiments 2 and 3 could be compared against the *delay* condition of Experiment 1; this would allow investigation of whether the different timing manipulations showed the same impact on competitor effects, and if they led to different patterns of effects for the two timing conditions. We conducted the same model as above, but using the *delay* condition as the reference group. Neither Experiment 2 nor Experiment 3 differed significantly from the *delay* group, and no interactions with word-type were attested (all $t < 1$). Further, comparisons between Experiment 2 and Experiment 3 showed no difference in the size of competitor effects and no interaction with word-type. Thus the *delay* condition and Experiments 2 and 3 led to a similar reduction in the competitor effects, and similar effects for both word types (perhaps each of these manipulations reduced spurious

associations to floor, leading to the extremely similar levels of effects). Future research is needed to determine whether these manipulations differentially impact the formation of spurious associations. Specifically, we predict that the visual preview manipulation of Experiment 3 may lead to a different timecourse of the elimination of competitor effects, as it relies on partial knowledge of the mappings to block coactivation.

Differences in the Quality of Learned Representations

An important concern across studies is whether these differences in competition simply reflect differences in how well words were learned in the two conditions. Our initial hypothesis was that more poorly learned words would exhibit more competition. Thus, in each of the previous experiments, care was taken to ensure that the *synchronous* group's increased competitor effects did not arise because they had learned the words more poorly than the *delay* group. However, it is also possible that *better* learning by the *synchronous* group could elicit larger competitor effects⁴. If learners in this group established stronger word-referent mappings, then the facility with which they fixated referents matching the acoustic stimulus during the VWP trials should be improved. This would lead to more ready target fixations (as was attested in Experiments 1 and 2); however it may also lead to increased consideration of partially matching referents early in the activation process. That is, participants in the *synchronous* group might have shown larger competitor effects because their stronger word-referent mappings sped the activation dynamics.

Looking across experiments at our three measures of learning quality (accuracy on the interim testing trials, accuracy on the VWP trials, and fixations to the target on the VWP trials) there is mixed evidence for differences in the overall quality of learning (Table 8). The two accuracy measures either show a numerically small disadvantage for *synchronous* training (Exp. 1, interim testing), or a marginal advantage for it (Exp. 2, VWP accuracy). The target fixation measure, however, is slightly more consistent, showing significantly or marginally more fixations for synchronous learners in two experiments. Thus, it is important to address the possibility that synchronous words were learned better, which in turn gives rise to greater competitor fixations.

To investigate this possibility, we examined the basic assumption that better learning leads to more competition. Thus, we examined the relationship between the level of learning and the degree of competitor effects across the three studies. If effects arise because of degree of learning rather than because of timing condition, then timing condition should be irrelevant to competitor effect size, but better learning should yield stronger competitor effects. We conducted several additional analyses to determine whether those who learned better also showed greater competitor effects. These analyses used accuracy on the interim trials and/or the VWP trials as our measure of learning. This was done (rather than the target fixations measure) for several reasons. First, it is the most direct measure of learning – fixations to the target also reflect oculomotor and visual-cognitive factors (Farris-Trimble & McMurray,

⁴This influence of learning hasn't been comprehensively examined. However, Magnuson, Tanenhaus, Aslin, & Dahan (2003) found that rhyme competitor effects decrease with learning, and found no change in cohort competitor effects. These results suggest that better learning is unlikely to result in the differences in competitor effects seen in our experiments.

2013). Second, target fixations are not independent of competitor fixations; if learners are looking at the target, they cannot be simultaneously looking at competitors.

We present a representative analysis below. Several additional analyses are available in the Online Supplement. These analyses uniformly showed no such relationship; even when comparing the poorest to the best learners, the degree of learning was substantially less predictive of competitor effect sizes than was timing condition, and in many cases better learners showed smaller competitor effects.

Representative Analysis—Determining how well a learner has acquired a word-referent mapping is difficult; despite two learners reaching ceiling performance, they may have sub-threshold differences in how well they have learned the mappings. Although we can't state unequivocally that all participants who are at 100% accuracy learned equally well, we can be fairly sure that those at ceiling learned better than those off ceiling – those making multiple errors clearly haven't learned the mappings as well as those making none. The primary analyses presented above (as well as those presented here) excluded participants who make a large number of errors for practical reasons: we need a large number of accurate trials for reliable eye-movement data. However, even among those who passed our accuracy threshold, there was variability in performance, allowing us to examine whether participants who learned the mappings better showed greater competitor effects than participants who learned the mappings more poorly.

We thus adopted an extreme groups design, comparing the very best learners to the very worst learners on the basis of accuracy during the VWP trials, and ignoring those in the middle. This split of the participants ensures that we are examining the greatest possible difference in learning. Across all conditions, there were 22 participants who made no more than one error during the VWP trials ($M=.36$ errors; $SD=.48$); we dubbed these participants the High Learners (HL). There were 19 participants who made at least nine errors during these trials ($M=16.63$; $SD=9.15$); these we dubbed the Low Learners (LL). The distribution of which conditions these participants were drawn from is presented in Table 9. The pairings were relatively evenly distributed across conditions (hinting that overall learning may not be a function of condition).

We then examined the competitor effect for members of these two groups. The HL group did indeed show slightly greater competitor effects than the LL group for onset competitors (Figure 12). However, this effect was small and non-significant (by two-tailed t-test; $t(39)=-.61$, $p=.55$). More convincingly, for offset competitors, the HL group showed *smaller* competitor effects than the LL group (though this too was non-significant by two-tailed t-test: $t(39)=-1.77$, $p=.085$)⁵. If anything, these results suggest that better learning may somewhat inhibit competition from offset competitors, running counter to the finding of the increased competitor effects seen by the *synchronous* group. The large differences in

⁵The difference in effects between the two types of competitors is somewhat surprising. A possible explanation is that, as mentioned in the Discussion for Experiment 1, activation of *offset competitors* may be much larger early in learning than late in learning (Magnuson et al., 2003). Low learners may be retaining these large offset activations for a greater number of trials, leading to their relatively large competitor effects.

competitor effects seen in the experiments are thus very unlikely to have arisen because of differences in learning.

A number of additional variants of this analysis are reported in the Online Supplement. These include treating accuracy as a continuous measure, dividing groups within timing condition, and using interim testing trials as a measure of learning. None of these analyses support the hypothesis that the strength of learning is related to the strength of competition; learners who quickly reached ceiling and remained there showed no greater competitor effects than those whose learning lagged, and in some cases seemed to show smaller competitor effects. This suggests that our effects on competition derive primarily from differences in the timing of the visual stimuli, not from a spurious difference in learning.

General Discussion

Word learning requires mapping auditory word forms to referents. However, as we show here, learning does not occur independently of stimulus processing. To form an association between a word and referent, both must be identified or categorized, a process that is usually conceptualized as dynamically activating the correct item or set of features from an array of competing interpretations. These experiments suggest the real-time processes of encoding word-forms and identifying objects are intricately linked to the somewhat slower processes of building associative links between them (learning).

It is well known that familiar word recognition proceeds incrementally: as words are heard, several competing word-forms are activated in parallel, and these activations are updated to reflect later arriving information. Similar findings are also observed for newly learned word forms (Creel, Aslin, & Tanenhaus, 2006; Kapnoula et al., 2015; Magnuson et al., 2003). Here we show that this real-time competition is linked to learning. Although learning could wait for this competition to resolve, Experiment 1 suggests that learning occurs during these periods of ambiguity, at least in some circumstances. Critically, in these experiments competing referents of the target word-form were *never* displayed in conjunction with it – these spurious associations between the target word form and a competing referent could only arise as a result of the learner's implicit activation of the competing word-form during the learning phase. In contrast, when learning was delayed by withholding visual presentation until after auditory stimulus offset, participants had time to resolve lexical competition, and the spurious associations were reduced.

The interactions between learning and processing are not solely bound to the moments when words are being heard. The dynamics of processing extend beyond stimulus offset, as competition continues to resolve, and they can be shaped by information received before the stimulus. Experiment 2 shows that dynamics continue to impact learning as ongoing competitor suppression after stimulus offset reduces spurious associations. Importantly, this learning—even the simple form of unsupervised learning implemented here—uses working memory representations of stimuli, and does not solely rely on perceptual information available in the moment. Learning is thus affected not just by the initial stimulus recognition and identification processes, but also by maintenance processes that keep representations of

stimuli active. This finding suggests that individual differences in memory functionality will influence the forms that associations take during learning.

Experiment 3 shows that not only do auditory processing dynamics affect learning, but the visual scene does as well (Chen & Mirman, 2012; Ramscar et al., 2010) as visual processing also unfolds dynamically. As a learner views a visual scene, she can form expectations about upcoming information, and these in turn can augment the way that auditory stimuli are processed, perhaps by biasing the competition between lexical items. The visual information allows learners to begin activating potential word forms that accord with the scene; when the auditory information arrives, the top-down activations from the visual stimuli merge with the bottom-up word recognition (akin to forming resonances with different sources of information as in Adaptive Resonance Theory; Grossberg & Myers, 2000), thereby shaping how processing proceeds. These visual effects require at least partial associations between the correct word form and referent, demonstrating a cyclical interaction between learning and processing: the degree of learning affects how competition resolves, which in turn affects further learning.

These experiments suggest a sophisticated interplay of real-time processing and slower timescales of learning. However, this interplay is ultimately undergirded by an unsophisticated form of learning, which simply links whatever representations are simultaneously active. Learners can form associations as stimulus processing unfolds, even if competition has not yet resolved, leading to parallel associations during periods of competition.

There are several potential empirical concerns worth addressing. First is the possibility that learners in one or the other group simply learned the words less well and this led to differences in competition. As Table 8 reveals, the evidence for this across studies is mixed at best. Moreover, as we report here and in the Online Supplement we found little evidence that the degree of learning is related to the degree of competition in this paradigm. While we cannot fully rule this explanation out, we see little support for it in our data.

A second potential concern is that the *synchronous* timing condition may have simply encouraged participants to more rapidly seek out the target during the testing portion. All conditions except the *synchronous* group had trials with identical total length, while the *synchronous* group had substantially shorter total trial times. Such speed could explain both the increased competitor looks and the faster looks to the target. However, several aspects of the data make this interpretation unlikely. First, competitor looks were considered relative to unrelated looks. If participants speed up looks under synchronous presentation while being more cautious with a delayed presentation, differences should emerge for all item types; instead, the *synchronous* group looked less at the unrelated items (e.g., Figure 4), suggesting that their increased competitor effects are not a result of more liberal looking during the trials. Second, participants in Experiment 3 likely had a similar response pressure to locate the targets after auditory onset. From the onset of the auditory stimulus, these participants had the same time-course of the trial as the *synchronous* participants – these participants had to quickly map the auditory information to the visual display before the next trial began. The additional time these participants received came before auditory stimuli were presented,

making it unlikely that time pressures were eliminated relative to the *synchronous* group. However, these participants did not show the same competitor effects as the *synchronous* participants, but instead showed decreased competitor effects at test. Despite the fast transition between trials after the auditory stimuli, these participants did not show the same degree of increased competition. As such, it appears unlikely that the increased competitor effects for the *synchronous* presentation arise because the time pressure during training encourages them to start seeking more liberally at trial onset than the other groups.

We've loosely framed the patterns of learning in these studies in terms of associations between something akin to a localist lexeme representation of the word-form and its referent (e.g., Chen & Mirman, 2012; McMurray, Horst, & Samuelson, 2012). However, this is not the only possible framing for this result. An alternative is that lexical entries do not have single unitary representations, but rather emerge out of direct association between distributed sound patterns (e.g., linking phonemes, syllables or other "fragments" of the word) directly to meanings. Indeed, this is the assumption of distributed models like the distributed cohort model (Gaskell & Marslen-Wilson, 1997) or the triangle model (Harm & Seidenberg, 1999). Under these assumptions, rather than forming mappings with partially activated word-forms, learners in the *synchronous* condition formed associations between the referents and lexical fragments (i.e. syllables). When hearing "go-," (the onset of *gobu*), learners may form associations directly between those phonological representations and the available referents. The consequence is that whenever they hear "go-" (whether as part of *gobu* or *gona*) that referent is more active. In the delayed condition, this tendency is reduced because the syllable and the picture are not co-present, and the remainder of the word is heard before the referent appears.

This explanation differs from the loose framing we've offered primarily in the nature of lexical representations. Our goal is not to address this classic debate here. In fact, at a broader level this alternative approach also accords with the ideas of dynamic learning that motivated our study. If learners are directly mapping referents to individual syllables, phonemes or fragments, it means that learners have started forming associations but before they've finished resolving the lexical competition and before binding these syllables into words. If learners are more likely to form strong associations with syllables in the *synchronous* group, it suggests that stimulus timing impacts which associations are formed. This approach thus also suggests that learning operates continually alongside processing; listeners can form word-referent mappings before competition resolves, either between referents and partially activated word-forms, or between referents and unbounded syllables of words.

Implications for our general understanding of learning

Despite a rich history of research on the mechanisms of learning, the precise timing of when associative traces are formed or updated has not been investigated. These studies offer initial evidence that associations can form before competition resolves, when multiple interpretations of stimuli (or working memory representations) are available. The associations are not contingent on having a single stimulus active or on competition being

complete; they form in parallel with all active representations, without using any form of competition monitor to gate learning.

This has implications for various domains that are characterized by dynamic competition mechanisms, including music perception (Bharucha, 1987), decision making (McKinstry et al., 2008; Shadlen & Kiani, 2013), spatial reasoning (Thelen, Schöner, Scheier, & Smith, 2001), and person perception (J. B. Freeman & Ambady, 2011). Additionally, beyond competition, other real-time processes like decay in working memory (Zhang & Luck, 2009), may alter representations that are ultimately linked via learning. Indeed, classic work on the phonological loop shows implicates working memory capacity as critical for word learning (Baddeley, Gathercole, & Papagno, 1998). More complex real-time processes like visual search, sentence parsing, or motor control could have even more complex consequences of learning.

The interaction of processing dynamics and learning may play out differently in other forms of learning. Feedback after a response is given provides a signal for when to initiate learning, and it may therefore guide learning toward more precise mappings. However, the timing of feedback, the availability of responses prior to the stimulus (e.g., Experiment 3), or response deadlines placed on the learner could still give rise to similar interactions as seen here.

For unsupervised forms of associative learning, understanding the precise dynamics of the learning environment is critical for understanding what associations are formed. Learning is enmeshed in stimulus processing, with bidirectional interactions. Rather than treating learning and stimulus recognition as independent, the two should be considered as integral components of the same system. These interactions can deepen descriptions of learning in populations with differences in dynamic processes. For example, differences in the dynamics of word recognition have been noted among those with dyslexia (Desroches, Joanisse, & Robertson, 2006), specific language impairment (Dollaghan, 1998; McMurray, Samelson, Lee, & Tomblin, 2010; Montgomery, 2002), individuals who use cochlear implants (Farris-Trimble, McMurray, Cigrand, & Tomblin, 2014), and over typical development (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Rigler et al., in press). These changes in processing dynamics could leave a mark on the nature of the mappings that are formed during learning.

A particularly relevant question is whether learning during ongoing competition *always* occurs, or if certain aspects of our experiment design led to a special form of processing. We argue for the former, for several reasons. First, as we describe below, more intelligent learning systems that block learning until competition is resolved may be unnecessary; natural variability in learning events may overcome instances of learned spurious associations. Moreover, even in the cases where such associations were most encouraged (i.e. the *synchronous* group of Experiment 1), participants were able to learn the correct word-referent mappings, suggesting a relatively small magnitude of the effects. The capacity for simple learning mechanisms to overcome spurious associations without specialized mechanisms obviates the pressure for such mechanisms to develop.

The second argument that this may be a more general effect is that our experiments create a perhaps uniquely useful opportunity for learners to engage such more intelligent mechanisms to block spurious associations: when given highly overlapping words and training with simultaneous audio-visual presentation, learners would clearly benefit from gating learning until competition resolves. The presence of learned spurious associations in the face of such conflict suggests that dynamic learning is the default. In natural learning situations, where immediate competition can be limited by the use of context, changes in temporal structure and greater variety in the information being learned, there would be less necessity to engage gating mechanisms.

Implications for word learning

These experiments suggest that real-time processes need to be considered in mechanistic models of word learning. Typical descriptions of word learning present the learner's task as a one-to-one mapping task of linking stimuli together after these stimuli are identified. Word recognition is treated as orthogonal to learning. However, these experiments demonstrate that the processes of word recognition are intertwined with learning.

In several domains within word learning, researchers have started to bridge these timescales of learning. Ramscar and colleagues (2010) showed that predictions formed by adults during a word learning event impact the efficacy of learning, and that manipulating the relative timing of auditory and visual events can impact these predictions (and through them, learning). However, in these studies the real-time processing (prediction) was in the service of learning, whereas the processing dynamics in the present study serve comprehension. The studies were not concerned with the dynamics during activation, and instead focused on how the ability to predict an upcoming word or an upcoming referent impacted the facility with which learning occurred. More specifically, this study demonstrated that manipulating the temporal order of stimuli can affect how easily words are learned, but it did not address how the temporal structure of a learning event impacts *what* is learned. The present study thus builds on Ramscar and colleagues' work, showing that temporal order has both global effects on acquisition of mappings, as well as local effects on the structure of learned representations.

Interactions among timescales are also of growing importance in theories of lexical development. Fernald and colleagues have shown that infants' efficiency of processing familiar words predicts later vocabulary development (Fernald & Marchman, 2012; Fernald, Perfors, & Marchman, 2006). Similarly, phenomena like fast-mapping that have been classically considered learning processes may actually be sophisticated forms of real-time inference (Horst & Samuelson, 2008), and these real-time processes change over development (Bion, Borovsky, & Fernald, 2013). However, the present study is the first to manipulate real-time processing and show consequences for *what* is learned online. Thus, it provides critical evidence for this broader theoretical thrust toward understanding word learning as the interaction of two timescales.

Interactions among timescales are also a key component of recent computational models of word learning. For example in the dynamic associative model (McMurray et al., 2012; McMurray, Zhao, Kucker, & Samuelson, 2013), word recognition is conceived as a real-time

competition process among competing referents, and mappings between word-forms and referents are simultaneously and gradually updated using simple Hebbian associative learning. In many ways this study represents strong empirical evidence of several foundational assumptions of this model: real-time competition and learning occur simultaneously and throughout lexical processing; and learning is served in part by simple, (perhaps dumb) mechanisms of Hebbian learning. At the same time, the computational instantiation of this theory has not yet addressed the complex competitive dynamics arising from phonological overlap (indeed this model abstracts away from any phonological representation), and in this way our empirical results prompt necessary extensions of this approach.

These results also demonstrate that word learning can rely on extremely simple associative learning mechanisms. Researchers have argued that word learning demands complex, non-associative learning mechanisms (Waxman & Gelman, 2009). In contrast, our results present evidence for the simplest imaginable form of associative learning. This provides at least an existence proof that word learning need not rely on complex learning systems (see also, Wasserman, Brooks, & McMurray, 2015), even as it may be embedded in much richer real-time processing systems (Roembke & McMurray, in press). Indeed it may be this interplay of time scales that allows word learning to appear so efficient and intelligent despite relatively simple associative mechanisms – real-time processes can allow children to make rapid and intelligent inferences about novel words, even as learning is slow and gradual. Despite the simplicity of the learning described in this manuscript, sophisticated learning of difficult, overlapping mappings can emerge – a learning system need not be complicated to yield complicated outcomes.

Our results argue against forms of learning that exclusively assign one-to-one mappings (like propose-but-verify; Medina et al., 2011). Although listeners eventually resolve competition, we show that learning can occur before a single word form is identified. Learners don't always form associations on the basis of post-competition activation of a single stimulus, but instead can form multiple, parallel associations. At least in some circumstances, learners entertain a many-to-many mapping, with the strength of these mappings varying for different stimuli.

These findings raise the possibility that learners form spurious associations throughout language acquisition. Words are often learned when their referents are present; the situation created in the *delay* condition of Experiment 1 (that appear to reduce such associations) is likely uncommon. However, there are several reasons to believe that such spurious associations may not substantively affect the course of word learning in children. First, natural variability in word learning events allows learning to occur at several points during competition resolution, including situations analogous to Experiment 2 (longer processing time) and Experiment 3 (the opportunity to “pre-activate”). The spurious associations seen in the *synchronous* group here may have been exaggerated by the rapid transition to the next trial after stimulus offset. Second, learners may not typically learn overlapping words in parallel; that is, children may know *sandwich* by the time they encounter *sandal*. Although the current experiments do not address coactivation of already-known words, Experiment 3

provides some hints that learners may bootstrap off the visual scene to constrain auditory processing to not include known words.

This builds on previous work showing that learners gradually come to learn word meanings, and use their partial knowledge to guide further learning (Yurovsky et al., 2014). Yurovsky and colleagues showed that word learning proceeds by gradual accumulation of knowledge from several exposures to word-referent pairings, and that as knowledge of one word-referent mapping is learned, it allows learners to bias later learning against other pairings with this word. A similar mechanism could explain why exposure to visual stimuli before auditory stimuli prevents the formation of spurious mappings; if learners have partial knowledge of a referent's label, they can winnow the candidates that are activated for later mappings. Additionally, few natural learning situations entail the rapid structure of the *synchronous* trials of Experiment 1. Instead learners can likely refine their learned associations after the learning event as processing proceeds, as in Experiment 2.

Finally, our use of highly overlapping words may exaggerate the degree of competition that would typically arise in natural word learning situations. We selected items that would engender maximum coactivation; typical learning situations likely rarely include such a dense collection of words being learned simultaneously. Our manipulation may thus magnify the formation of spurious associations. However, this does not diminish the importance of this finding. As discussed above, learning such a dense set of words would be the primary situation in which to engage intelligent learning systems to wait until competition resolves. Finding evidence of learning during competition resolution in this condition provides extremely strong evidence that learning is always occurring. Even if dynamic learning is unlikely to exert a cost in natural word learning situations, it likely has some drawbacks. Could there be any benefits that might outweigh these costs? First is rapidity – in running speech, there is typically little time between the recognition of a target word and the subsequent word. If the system can only learn when competition is fully resolved, there may just not be enough time to do any useful learning in the context of running speech.

Second, there is also a deeper theoretical issue that arises in the context of running speech. Many of the early debates on segmentation of familiar words make the point that it is not always clear when a word ends – given the large number of embedded words (e.g., *car* in *cargo*) there can be considerable uncertainty as to when words end. Indeed as many as 40% of words in a corpus of natural speech cannot be identified until several phonemes after their offset (Bard, Shillcock, & Altmann, 1988). A classic solution to this issue in segmentation is continuous mapping (McClelland & Elman, 1986)—the system continuously compares multiple possible lexical interpretations and lets competition sort it out later. This solves the problem without requiring an explicit marker of a word boundary. In the context of learning, segmentation is likely to be even more difficult given that children may not know all of the words in an utterance. Here dynamic learning can be seen as a form of continuous mapping operating in learning -- the system greedily links whatever it can to the available referents and lets later learning situations sort it out.

As we have described, natural word learning is unlikely to result in massive spurious associations, and there may be benefits to the rapidity of dynamic learning. However, the

point of this work was not to capture realistic aspects of word learning. Rather, our goal was more basic. Our evidence of the continuous and dynamic nature of associative learning offers a critical insight into the nature of learning. When forming associations, learning is not fully gated until competition resolves. Instead, learning occurs as stimulus processing unfolds, before the learner has identified a singular interpretation of the stimulus.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Prahlad Gupta, Eliot Hazeltine, Karla McGregor, Larissa Samuelson and Bruce Tomblin for helpful discussions, Dan McEchron for technical help, and Vladimir Sloutsky and Caitlin Hilliard for comments on an earlier draft. This work was supported by the Ballard and Seashore Dissertation Year Fellowship (to K.S.A.) and National Institutes of Health Grant DC-0008089 (to B.M.).

References

- Akhtar N, Jipson J, Callanan MA. Learning words through overhearing. *Child Development*. 2001; 72(2):416–430. [PubMed: 11333075]
- 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(4):419–439. DOI: 10.1006/jmla.1997.2558
- Andruski JE, Blumstein SE, Burton MW. The effect of subphonetic differences on lexical access. *Cognition*. 1994; 52(3):163–187. [PubMed: 7956004]
- Baddeley A, Gathercole S, Papagno C. The Phonological Loop as a Language Learning Device. *Psychological Review*. 1998; 105(1):158–173. DOI: 10.1037/0033-295X.105.1.158 [PubMed: 9450375]
- Bard EG, Shillcock RC, Altmann GTM. The recognition of words after their acoustic offsets in spontaneous speech. *Perception & Psychophysics*. 1988; 44(5):395–408. [PubMed: 3226889]
- Bharucha J. Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*. 1987; 5(1):1–30.
- Bion, RaH, Borovsky, A., Fernald, A. Fast mapping, slow learning: disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*. 2013; 126(1):39–53. DOI: 10.1016/j.cognition.2012.08.008 [PubMed: 23063233]
- Chen Q, Mirman D. Competition and cooperation among similar representations: toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*. 2012; 119(2):417–30. DOI: 10.1037/a0027175 [PubMed: 22352357]
- Connine CM, Blasko DG, Hall M. Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraint. *Journal of Memory and Language*. 1991; 30(2): 234–250. DOI: 10.1016/0749-596X(91)90005-5
- 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):1–19. DOI: 10.1016/j.jml.2005.09.003
- 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(5/6):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(2):498–513. DOI: 10.1037/0278-7393.30.2.498

- Dautriche I, Chemla E. Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2014; 40(3):892–903. DOI: 10.1037/a0035657
- Desroches AS, Joanisse MF, Robertson EK. Specific phonological impairments in dyslexia revealed by eyetracking. *Cognition*. 2006; 100(3):B32–42. DOI: 10.1016/j.cognition.2005.09.001 [PubMed: 16288732]
- Dollaghan C. Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics*. 1998; 19(02):193. doi: 10.1017/S0142716400010031
- Farris-Trimble A, McMurray B. Test – Retest Reliability of Eye Tracking in the Visual World Paradigm for the Study of Real-Time Spoken Word Recognition. *Journal of Speech, Language, and Hearing Research*. 2013; 56(August):1328–1346. DOI: 10.1044/1092-4388(2012/12-0145)group-level
- Farris-Trimble A, McMurray B, Cigrand N, Tomblin JB. The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology. Human Perception and Performance*. 2014; 40(1):308–27. DOI: 10.1037/a0034353 [PubMed: 24041330]
- Fernald A, Marchman Va. Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*. 2012; 83(1):203–22. DOI: 10.1111/j.1467-8624.2011.01692.x [PubMed: 22172209]
- Fernald A, Perfors A, Marchman Va. Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*. 2006; 42(1):98–116. DOI: 10.1037/0012-1649.42.1.98 [PubMed: 16420121]
- Fernald A, Pinto JP, Swingley D, Weinberg A, McRoberts GW. Rapid Gains in Speed of Verbal Processing by Infants in the 2nd Year. *Psychological Science*. 1998; 9(3):228–231. DOI: 10.1111/1467-9280.00044
- Freeman JB, Ambady N. A dynamic interactive theory of person construal. *Psychological Review*. 2011; 118(2):247–279. DOI: 10.1037/a0022327 [PubMed: 21355661]
- Freeman, JH. Developmental neurobiology of cerebellar learning. In: Blumberg, MS, Freeman, JH., Robinson, SR., editors. *Oxford Handbook of Developmental Behavioral Neuroscience*. Oxford: Oxford University Press; 2010.
- Gaskell MG, Dumay N. Lexical competition and the acquisition of novel words. *Cognition*. 2003; 89(2):105–132. DOI: 10.1016/S0010-0277(03)00070-2 [PubMed: 12915296]
- Gaskell MG, Marslen-Wilson WD. Integrating Form and Meaning: A Distributed Model of Speech Perception. *Language and Cognitive Processes*. 1997; 12(5/6):613–656. DOI: 10.1080/016909697386646
- Gaskell MG, Marslen-Wilson WD. Representation and competition in the perception of spoken words. *Cognitive Psychology*. 2002; 45(August 1997):220–266. [PubMed: 12528902]
- Goldstone RL, Landy D. Domain-Creating Constraints. *Cognitive Science*. 2010; 34(7):1357–1377. DOI: 10.1111/j.1551-6709.2010.01131.x [PubMed: 21564250]
- Griffiths TL, Steyvers M, Tenenbaum JB. Topics in semantic representation. *Psychological Review*. 2007; 114(2):211–244. DOI: 10.1037/0033-295X.114.2.211 [PubMed: 17500626]
- Grossberg S, Myers C. The resonant dynamics of speech perception: interword integration and duration-dependent backward effects. *Psychological Review*. 2000; 107(4):735–767. [PubMed: 11089405]
- Harm MW, Seidenberg MS. Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*. 1999; 106(3):491–528. [PubMed: 10467896]
- Hebb, DO. *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley and Sons; 1949.
- Horst JS, Samuelson LK. Fast mapping but poor retention by 24-month-old infants. *Infancy*. 2008; 13(2):128–157. DOI: 10.1080/15250(xH)701795598
- Huette S, McMurray B. Continuous dynamics of color categorization. *Psychonomic Bulletin & Review*. 2010; 17(3):348–54. DOI: 10.3758/PBR.17.3.348 [PubMed: 20551357]
- Kapnoula E, Gupta P, Packard S, McMurray B. Immediate lexical integration of novel word forms. *Cognition*. 2015; 134:85–99. [PubMed: 25460382]
- Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*. 1998; 19(1):1–36. [PubMed: 9504270]

- Magnuson JS, Dixon JA, Tanenhaus MK, Aslin RN. The dynamics of lexical competition during spoken word recognition. *Cognitive Science*. 2007; 31(1):133–156. [PubMed: 21635290]
- 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(2):202–227. [PubMed: 12825637]
- Markman EM, Wachtel GF. Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*. 1988; 20(2):121–157. DOI: 10.1016/0010-0285(88)90017-5 [PubMed: 3365937]
- Marslen-Wilson WD. Functional parallelism in spoken word-recognition. *Cognition*. 1987; 25(1–2): 71–102. [PubMed: 3581730]
- Marslen-Wilson WD, Warren P. Levels of Perceptual Representation and Process in Lexical Access: Words, Phonemes, and Features. *Psychological Review*. 1994; 101(4):653–675. DOI: 10.1037/0033-295X.101.4.653 [PubMed: 7984710]
- Marslen-Wilson WD, Zwitserlood P. Accessing Spoken Words: The Importance of Word Onsets. *Journal of Experimental Psychology: Human Perception and Performance*. 1989; 15(3):576–585.
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology*. 1986; 18(1):1–86. [PubMed: 3753912]
- McKinstry C, Dale R, Spivey MJ. Action dynamics reveal parallel competition in decision making. *Psychological Science*. 2008; 19(1):22–4. DOI: 10.1111/j.1467-9280.2008.02041.x [PubMed: 18181787]
- McMurray B, Clayards Ma, Tanenhaus MK, Aslin RN. Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin & Review*. 2008; 15(6):1064–71. DOI: 10.3758/PBR.15.6.1064 [PubMed: 19001568]
- McMurray B, Horst JS, Samuelson LK. Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*. 2012; 119(4):831–77. DOI: 10.1037/a0029872 [PubMed: 23088341]
- McMurray B, Samuelson VM, Lee SH, Tomblin JB. Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*. 2010; 60(1):1–39. DOI: 10.1016/j.cogpsych.2009.06.003 [PubMed: 19836014]
- McMurray, B., Zhao, L., Kucker, S., Samuelson, LK. Pushing the envelope of associative learning: Internal representations and dynamic competition transform association into development. In: Hollich, G., Gogate, L., editors. *Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence*. Hershey, PA: IGI Global; 2013. p. 49-80.
- McQueen JM, Norris D, Cutler A. Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1994; 20(3):621–638. DOI: 10.1037/0278-7393.20.3.621
- Medina TN, Snedeker J, Trueswell JC, Gleitman LR. How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(22):9014–9. DOI: 10.1073/pnas.1105040108 [PubMed: 21576483]
- Miller J, Hackley Sa. Electrophysiological evidence for temporal overlap among contingent mental processes. *Journal of Experimental Psychology: General*. 1992; 121(2):195–209. DOI: 10.1037//0096-3445.121.2.195 [PubMed: 1318354]
- Montgomery JW. Examining the Nature of Lexical Processing in Children with Specific Language Impairment: Temporal Processing or Processing Capacity Deficit? *Applied Psycholinguistics*. 2002; 23(3):447–470. DOI: 10.1017/S0142716402003077
- Munakata Y, O'Reilly RC. Developmental and computational neuroscience approaches to cognition: The case of generalization. *Cognitive Studies*. 2003; 10:76–92.
- O'Reilly, RC., Munakata, Y. *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press; 2000.
- O'Reilly RC, Norman KA. Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *Trends in Cognitive Sciences*. 2002; 6(12):505–510. [PubMed: 12475710]
- Ramscar M, Dye M, Klein J. Children value informativity over logic in word learning. *Psychological Science*. 2013; 24(6):1017–23. DOI: 10.1177/0956797612460691 [PubMed: 23610135]

- Ramscar M, Yarlett D, Dye M, Denny K, Thorpe K. The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*. 2010; 34(6):909–57. DOI: 10.1111/j.1551-6709.2009.01092.x [PubMed: 21564239]
- Rescorla, RA., Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH., Prokasy, WF., editors. *Classical Conditioning II*. Appleton-Century-Crofts; 1972. p. 64-99.
- Rigler H, Farris-Trimble A, Greiner L, Walker J, Tomblin JB, McMurray B. The slow developmental timecourse of real-time spoken word recognition. *Developmental Psychology*. n.d
- Roembke TC, McMurray B. Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*. 2016; 87:105–127. DOI: 10.1016/j.jml.2015.09.005 [PubMed: 26858510]
- Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986; 323:533–536.
- Salverda AP, Dahan D, Tanenhaus MK, Crosswhite K, Masharov M, McDonough J. Effects of prosodically modulated sub-phonetic variation on lexical competition. *Cognition*. 2006; doi: 10.1016/j.cognition.2006.10.008
- Salverda AP, Kleinschmidt D, Tanenhaus MK. Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language*. 2014; 71(1):145–163. DOI: 10.1016/j.jml.2013.11.002 [PubMed: 24511179]
- Samuelson LK, Smith LB, Perry LK, Spencer JP. Grounding word learning in space. *PloS One*. 2011; 6(12):e28095.doi: 10.1371/journal.pone.0028095 [PubMed: 22194807]
- Schutte AR, Spencer JP, Schöner G. Testing the dynamic field theory: working memory for locations becomes more spatially precise over development. *Child Development*. 2003; 74(5):1393–417. [PubMed: 14552405]
- Seidenberg MS, McClelland JL. A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*. 1989; 96(4):523–568. [PubMed: 2798649]
- Shadlen MN, Kiani R. Decision making as a window on cognition. *Neuron*. 2013; 80(3):791–806. DOI: 10.1016/j.neuron.2013.10.047 [PubMed: 24183028]
- Siskind JM. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*. 1996; 61:39–91. [PubMed: 8990968]
- Spivey, MJ. *The Continuity of Mind*. Oxford: Oxford University Press; 2007.
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. *Science*. 1995; 268(5217):1632–1634. DOI: 10.1126/science.7777863 [PubMed: 7777863]
- Thelen E, Schöner G, Scheier C, Smith LB. The dynamics of embodiment: a field theory of infant perseverative reaching. *The Behavioral and Brain Sciences*. 2001; 24(1):1–34. discussion 34–86. [PubMed: 11515285]
- Thiessen ED. Domain general constraints on statistical learning. *Child Development*. 2011; 82(2):462–470. DOI: 10.1111/j.1467-8624.2010.01522.x [PubMed: 21410909]
- Trueswell JC, Medina TN, Hafri A, Gleitman LR. Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*. 2013; 66(1):126–156. DOI: 10.1016/j.cogpsych.2012.10.001 [PubMed: 23142693]
- Wasserman EA, Brooks D, McMurray B. Pigeons acquire multiple categories in parallel via associative learning: A parallel to human word learning? *Cognition*. 2015
- Waxman SR, Gelman SA. Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*. 2009; 13(6):258–263. [PubMed: 19447670]
- Yu C, Smith LB. Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*. 2007; 18(5):414–420. DOI: 10.1111/j.1467-9280.2007.01915.x [PubMed: 17576281]
- Yu C, Smith LB. Modeling cross-situational word-referent learning: prior questions. *Psychological Review*. 2012; 119(1):21–39. DOI: 10.1037/a0026182 [PubMed: 22229490]
- Yurovsky D, Fricker DC, Yu C, Smith LB. The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review*. 2014; 21:1–22. DOI: 10.3758/s13423-013-0443-y [PubMed: 23702980]

Zhang W, Luck SJ. Sudden death and gradual decay in visual working memory. *Psychological Science*. 2009; 20(4):423–8. DOI: 10.1111/j.1467-9280.2009.02322.x [PubMed: 19320861]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.
Examples of visual stimuli used in the experiments (images in the experiment were in color).

Figure 2. Proportion of correct responses during interim testing trials for Experiment 1. A) Onset competitor trials. B) Offset competitor trials.

Figure 3. Raw proportion of looks to the target items in Experiment 1. A) Onset competitor trials. B) Offset competitor trials.

Figure 4. Raw proportion of looks to competitor and unrelated (UR) items in Experiment 1. Unrelated items reflect the mean of looks to the two unrelated items in the display. A) Onset-competitor trials. B) Offset-competitor trials.

Fig. 5. Competitor effects in Experiment 1. Looks to the competitor are compared to the mean of looks to the two unrelated items. A) Log odds ratio of looks for onset competitors across time. B) Log odds ratio of looks to the offset competitors divided by looks to the unrelated objects across time. C) Mean log odds ratio during analysis window (500–1500 ms) by competitor type and timing condition. Error bars are the standard error of the mean.

Figure 6. Proportion of correct responses during the interim testing trials across blocks of training. A) Onset competitor items. B) Offset competitor items.

Figure 7.

Raw proportion of looks to the target object across time, comparing participants from the synchronous condition of Experiment 1 to participants in Experiment 2. A) Onset competitor trials. B) Offset competitor trials.

Figure 8. Log odds ratios of competitor effects in Experiment 2. A) Log odds ratio for onset competitor trials across time. B) Log odds ratio for offset competitor trials across time. C) Mean log odds ratio during analysis window by competitor type and timing condition. Error bars are the standard error of the mean.

Figure 9. Proportion of correct responses during the interim testing trials, comparing participants in the synchronous condition of Experiment 1 to participants in Experiment 3. A) Onset competitor trials. B) Offset competitor trials.

Figure 10.

Raw proportion of looks to the target object across time, comparing participants in the synchronous condition of Experiment 1 to participants in Experiment 3. A) Onset competitor trials. B) Offset competitor trials.

Figure 11. Log odds ratios of competitor effects for Experiment 3. A) Log odds ratio for onset competitor trials across time. B) Log odd ratio for offset competitor trials across time. C) Mean log odd ratios during analysis window by competitor type and timing condition. Error bars signify the standard error of the mean.

Figure 12. Log odds ratios of competitor effects for different learner groups (LL: Low learners; HL: High learners). Error bars signify the standard error of the mean.

Table 1
Results from analyses of accuracy during interim and VWP trials. Significant effects are in bold.

Model	Comparison	<i>B</i>	<i>SE</i>	<i>Z</i>	<i>p</i>
Accuracy during interim trials for Experiment 1	Block	.72	.062	11.63	<.0001
	Timing condition	-.64	.46	-1.40	.16
	Word type	-.37	.35	-1.05	>.2
	Block × timing	.30	.12	2.48	.013
	Block × type	.094	.12	.78	>.2
	Timing × type	-.054	.70	-.076	>.2
Accuracy during VWP trials for Experiment 1	Block × timing × type	-.022	.24	-.091	>.2
	Timing condition	-.27	.32	-.85	>.2
	Word type	.38	.15	2.46	.014
	Timing × type	-.41	.23	-1.75	.080

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2
Results from analyses of competitor effect during VWP trials. Significant effects are in bold.

Comparison	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	.13	.024	12.1	5.40	.00016
Timing condition	-.10	.038	57.0	-2.77	.0075
Word type	-.096	.039	6.0	-2.43	.051
Timing × type	.016	.052	406.1	.30	>.2

Table 3

Results from analyses of accuracy during interim and VWP trials, comparing Experiment 2 to the *synchronous* group from Experiment 1. Significant effects are in bold.

Model	Comparison	B	SE	Z	p
Accuracy during interim trials for Experiment 2	Block	.55	.057	9.77	<.0001
	Timing condition	.71	.50	1.42	.16
	Word type	-.29	.37	-.78	>.2
	Block × timing	-.045	.11	-.40	>.2
	Block × type	.21	.11	1.85	.064
	Timing × type	.10	.72	.14	>.2
	Block × timing × type	.20	.22	.92	>.2
Accuracy during VWP trials for Experiment 2	Timing condition	-.15	.25	-.60	>.2
	Word type	.48	.17	2.87	.0041
	Timing × type	-.22	.25	-.90	>.2

Table 4

Results from analyses of competitor effect during VWP trials comparing Experiment 2 to synchronous participants from Experiment 1. Significant effects are in bold.

Comparison	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	.14	.020	11.7	7.11	<.0001
Timing condition	-.077	.035	57.5	-2.22	.030
Word type	-.088	.033	6.0	-2.65	.038
Timing × type	.033	.052	406.6	0.63	>.2

Table 5

Results from analyses of accuracy during interim and VWP trials, comparing Experiment 3 to the *synchronous* group from Experiment 1. Significant effects are in bold.

Model	Comparison	B	SE	Z	p
Accuracy during interim trials for Experiment 3	Block	.65	.065	10.02	<.00001
	Timing condition	.48	.49	.96	>.2
	Word type	-.58	.42	-1.40	.16
	Block × timing	-.17	.13	-1.31	.19
	Block × type	.13	.13	1.01	>.2
	Timing × type	.63	.77	.81	>.2
Accuracy during VWP trials for Experiment 3	Block × timing × type	-.14	.25	-.55	>.2
	Timing condition	-.58	.31	-1.85	.064
	Word type	.70	.31	2.29	.022
	Timing × type	.20	.22	.89	>.2

Table 6

Results from analyses of competitor effect during VWP trials comparing Experiment 3 against the synchronous group of Experiment 1. Significant effects are in bold.

Comparison	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	.15	.023	10.8	6.36	<.0001
Timing condition	-.072	.035	58.2	-2.05	.045
Word type	-.086	.039	6.0	-2.22	.069
Timing × type	.035	.050	406.6	.70	>.2

Table 7

Results from analyses of competitor effect during VWP trials comparing all experiments in a single omnibus model. The synchronous group from Experiment 1 served as the reference group. Significant effects are in bold.

Comparison	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	.18	.026	48.8	7.0	<.0001
Sync vs. delay	-.10	.033	114.0	-3.2	.0020
Sync vs. Exp 2	-.077	.033	113	-2.3	.022
Sync vs. Exp 3	-.072	.033	113.8	-2.2	.033
Word-type	-.11	.040	66.2	-2.8	.0064
Delay × word type	.015	.049	818.4	.31	.75
Exp 2 × word type	.033	.049	816.5	.67	.50
Exp 3 × word type	.054	.052	551.7	1.05	.30

Table 8

Summary of results on quality of learning across three experiments. Significant effects are marked with *, marginal effects with +.

	Accuracy		
	Interim Testing	VWP	Efficiency Target Fixations
Exp 1	Sync poorer *	∅	Sync faster +
Exp 2	∅	∅	Sync faster *
Exp 3	∅	Sync better +	∅

Table 9

Distribution of participants from different conditions into learning groups.

Condition	HL	LL
Sync	8	5
Delay	6	6
Exp 2	5	1
Exp 3	3	7