Summer 2013

Real-Time Competition Processes in Word Learning

Keith S. Apfelbaum
University of Iowa

REAL-TIME COMPETITION PROCESSES IN WORD LEARNING

by

Keith S. Apfelbaum

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

August 2013

Thesis Supervisor: Associate Professor Bob McMurray

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Keith S. Apfelbaum

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Psychology at the August 2013 graduation.

Thesis Committee: Bob McMurray, Thesis Supervisor

Prahlad Gupta

R. Eliot Hazeltine

Karla McGregor

Larissa Samuelson

J. Bruce Tomblin

You keep using that word. I do not think it means what you think it means.
Inigo Montoya
The Princess Bride

ACKNOWLEDGMENTS

Well before beginning an endeavor as daunting as writing a dissertation, an aspiring scientist needs to find a home for his research. In this, I am luckier than most, as I very nearly didn't get started. After making a foolhardy choice to attend a competitor university rather than Iowa, I came shuffling back with my tail between my legs, asking for a second chance. I owe a great debt of gratitude to the Psychology Department at Iowa, and particularly to my advisor, Bob McMurray, for rescuing me from an academic abyss and bringing me into a program that perfectly met my academic and personal needs.

This generosity from Bob and the Psychology Department has carried through my time at Iowa, providing me incredible support in becoming an independent researcher. I've been constantly pushed to challenge the status quo in order to seek the clearest truths in my science; in a department surrounded by accomplished and intimidating scientists, I was never expected to just go along with an idea without first being convinced. Thanks to Bob and others for keeping me honest and insisting I return the favor.

While working with Bob, I've had the distinct pleasure and good fortune of working with many top-notch colleagues within the MACLab. I haven't space to list all who have helped my development at Iowa, but a few labmates deserve particular credit for their contributions to this dissertation, other work I've done at Iowa, and my general intellectual growth. Ariane Rhone and Ashley Farris-Trimble were founts of knowledge and constant sources of support, both academically and when I needed a friend to keep me forging ahead in my studies. Marcus Galle was my constant partner in crime, and our frequent academic competitions helped me accomplish much during my career at Iowa. Joe Toscano, Gwyneth Rost, Tanja Roembke, Cheyenne Munson, Effie Kapnoula, Hannah Rigler, Michelle Quinn and Kayleen Hannaway were wonderful sources of inspiration (and valuable sources of criticism). Countless undergraduate RAs helped keep the lab running smoothly and happily; particular thanks go to Aimee Marino, Natasha Bullock-Rest, Kelsey Wiggs and Sarah Standley.

My dissertation committee of Prahlad Gupta, Eliot Hazeltine, Karla McGregor, Larissa Samuelson and Bruce Tomblin has been a constant sounding board as my dissertation project (and my previous failed attempt at a dissertation project) has developed. Their advice and amazing patience, on this project and on numerous others, have made this process all the more worthwhile.

Several academic groups at Iowa have proven foundational in keeping me thinking hard about my work and other important topics in cognitive and developmental psychology. I especially want to thank members of the Delta Center, the Modeling and Language Discussion Groups, and the Cognition and Perception and Developmental Journal Club for insightful discussions and frequent reminders of the fun side of science.

Of course, completing a PhD doesn't come cheap. The projects in this dissertation were funded by NIH grant DC008089 to Bob McMuray and the Ballard-Seashore Dissertation Year Fellowship to me. I also want to thank the department for my years of funding and travel support to keep my intellectual drive going.

Family and friends are essential to survive graduate school. To my family: heartfelt thanks for your patience and support, even when you (and I) had no idea what I was doing. Victoria Erhardt has been amazingly generous and patient in helping me reach my goals, even when they've proven crazy or scary, as have many friends both in Iowa City and across the world.

When deciding to become a scientist, there's an image of putting together easy experiments to prove hypotheses and generate new knowledge. It seems like a straightforward, linear process. The most important thing I've learned in graduate school is that science is never linear, and is rarely easy. Yet despite the frustrating, confounding, often infuriating nature of this pursuit, the people around my at Iowa have kept it exciting, intriguing and fun.

ABSTRACT

Perceptual processes take time to unfold. Whether a person is processing a visual scene, identifying the category an object belongs to, or recognizing a word, cognitive processes involving competition across time occur. These ongoing competitive processes have been ignored in studies of learning. However, some forms of learning suggest that learning could occur while competition is ongoing, resulting in the formation of mappings involving the competing representations. This dissertation uses word learning as a test case to determine whether such learning exists. In a series of five experiments, participants were taught words under different stimulus and task conditions to encourage or discourage learning during periods of lexical competition. These studies reveal a complex relationship between ongoing lexical competition processes and word learning. Specifically, in cases where learners rely on unsupervised associative learning, they present evidence of learning that is continuous in time, starting during periods of lexical competition and continuing throughout the course of its resolution. These studies offer insight into the nature of associative learning, into the forms of learning that occur when learning new words, and into the ways that task and stimulus structure impinge on how a learner forms new associations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

BACKGROUND

1.1 Temporal dynamics of perceptual processing

Most aspects of perception require some degree of complex processing to recognize or interpret. Although cognition proceeds at impressive speeds, these processes take time to play out. This temporal component is present across a huge array of perceptual processes: when searching for a visual referent in a scene with many distracting objects, the viewer often spends a protracted period seeking the correct visual target (Treisman & Gelade, 1980; Treisman, 1986; Wolfe, 1994); distinguishing between colors that are similar elicits competition between representations, which requires some time to resolve (Bornstein & Korda, 1984; Huette & McMurray, 2010); auditory words are processed across time, with several word-forms active in parallel until sufficient information is received to disambiguate the forms (Allopenna, Magnuson, & Tanenhaus, 1998; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994; Zwitserlood, 1989). For example, when hearing the word "sandal," the listener initially activates both *sandal* and *sandwich*. As more of the word is heard, *sandwich* is inhibited, however *candle* becomes active, as it overlaps with later portions of the word. Thus a central aspect of much of cognitive processing is the temporal dynamics of how this processing unfolds.

The temporal dynamics of processing not only offer important research questions about general cognitive capacities, but they also raise important challenges to understanding *learning* within these domains. As information comes in across time and requires time to process, at what point does the organism learn[1]? When multiple stimuli

[1] The term "learning" is quite ambiguous, as it refers to both changes within a single learning event and more long-term changes across a series of events. In this case, the question is how learning within a single instance occurs; the critical question investigated throughout this dissertation is the time at which associations, weights or connections are adjusted during a learning event.

are active in parallel and competing for activation – as in color perception and word recognition – how does the learning system cater learning to the correct stimuli instead of the briefly co-active competitors?

Learning in temporally-extended domains could take two broad forms. First, learning could rely on some mechanism that determines when the processing of the relevant stimuli is complete and only then initiate learning. Such a system would gate any learning to wait until competition has resolved so that only the representations that win the competition are learned. This could be achieved either through an explicit monitor of activation, or through more implicit means, such as initiating learning when activation reaches some threshold. A learner relying on such a competition monitor would avoid forming any representations based on briefly active competitors, as no learning occurs until after competition processes have resolved. In the case of learning the mappings between words and their referents, for example, while lexical activation processes are ongoing and several word-forms are competing in parallel for activation, the learner could wait until one word-form has suppressed the activation of the competitors before forming any representations; they would not learn until *sandwich* had suppressed *sandal*.

Alternatively, learning could proceed continuously in time; whenever stimuli are active, the learning system may immediately begin to form representations. For word learning, the formation of word-referent mappings would begin as soon as both word-form and referent stimuli have received any degree of activation. This would lead to representations mapping the parallel-activated word-forms onto referents, leading to an array of spurious mappings; the learning would initiate when both *sandal* and *sandwich* are active. However, because the appropriate word-form's activation with respect to the referent is both more consistent and longer-lasting than that of the competitors, the correct mapping could nonetheless come to dominate learning.

Although both of these forms of learning are capable of accommodating temporal processing dynamics, they do so in quite different ways. Determining whether learning relies on signals indicating the presence or absence of competition can help elucidate a fundamental facet of learning. Yet little work has investigated the timing of when learning occurs relative to stimulus processing. Instead, many studies focus on broad questions about the nature of learning. Those that emphasize more specific mechanisms of learning systems often use stimuli that obviate concerns of processing dynamics in order to ask alternative questions. In both these approaches, temporal processing dynamics are external to the questions of interest, leading researchers to treat such factors as noise to be controlled for as best as possible in their designs.

1.1.1 Questions on the nature of learning

A large volume of learning research has been devoted to uncovering broad principles about the nature of learning. By determining what tools are available to learners, what broad forms of representation are used in learning and what forms of information are most learnable, these approaches offer deep insights into the learning system. However, these approaches operate at a fundamentally higher level than investigations of online processing dynamics and learning. They thus provide little leeway into investigating these more microstructure questions about how learning proceeds in individual learning events.

Debates about the mechanisms that enable learning have been ongoing for millennia. Plato and Aristotle debated whether humans learned new information by unlocking innate knowledge or by filling in a blank slate of the mind. This debate has evolved into modern psychological comparisons, pitting rationalist approaches, which argue for inborn capacities or knowledge and more rational or inferential processes driving the course of learning, against emergentist theories, which argue that learning proceeds using domain-general systems that learn through experience by tracking patterns in the environment. In language development, rationalists argue for innate devices that are tuned to learning linguistic structures (Fisher & Marcus, 2006; Fodor, Gaskell, & Brill, 1975; Hirsh-Pasek et al., 1987; Kanwisher, 2010; Marcus, Pinker, & Ullman, 1992). This approach is supported by hypotheses that the structure of language is too complex to learn using domain-general pattern-recognizing mechanisms (Fodor & Pylyshyn, 1988; Marcus et al., 1992; Marcus, 1993; Pinker & Prince, 1988), and that linguistic input is impoverished relative to the vastness of the language that is learned (Berwick, Pietroski, Yankama, & Chomsky, 2011; Marcus, 1993; though see, Hsu & Chater, 2010). The emergentist approach counters these arguments by emphasizing the capacity for learners to extract complex information from natural linguistic input (Chater & Christiansen, 2010; Christiansen, Allen, & Seidenberg, 1998; McClelland, St John, & Taraban, 1989; Plaut, 1997; Plunkett & Marchman, 1993; Seidenberg & McClelland, 1989), and by demonstrating that quite simple domain-general learning is capable of learning to abstract beyond properties of the stimulus (Apfelbaum & McMurray, 2011; Christiansen et al., 1998; McClelland, 2010; Plunkett & Marchman, 1993; Thiessen, 2011; Toscano & McMurray, 2010). Many such accounts rely on associative learning as a powerful learning mechanism that is capable of acquiring detailed information without specialized mechanisms. More nuanced approaches suggest that the distinction between rationalist and emergentist theories is false, as the two forms are fully inseparable. Instead, learning is always a combination of influences from different capacities, and trying to partition learning between them is an impossible task (Chang, Dell, & Bock, 2006; Chang, 2002; Spencer, Blumberg, et al., 2009; Spencer, Samuelson, et al., 2009). Theories about the degree of rationalist vs. emergentist mechanisms used in learning are focused on high-level discussion about the nature of learning, with little focus on the dynamics of each individual learning event. As such, questions of perceptual processing are wholly external to this debate.

The form of perceptual representation also drives major debates regarding the nature of learning. Studies of categorization contrast theories that learners store a huge number of individual exemplars of categories, and use these exemplars as a basis for future stimulus identification (Creel, Aslin, & Tanenhaus, 2008; Goldinger, 1998; Kruschke, 1992; Nosofsky, Kruschke, & McKinley, 1992; Nosofsky, 1986; Pierrehumbert, 2003; Regier, 2005; Zaki & Nosofsky, 2007) against theories that learners utilize a form of stimulus averaging to determine a prototypical category member, and then compare future items against this prototype (Barsalou, 1985; Minda & Smith, 2002; Posner & Keele, 1968; J. D. Smith, 2002). These theories differ greatly in their presumption of how learning occurs. Exemplar theories include a mechanism for learning each exemplar and applying a category tag to it, whereas prototype learning requires a mechanism for adjusting representations to reflect new prototypes. However, this debate again disregards questions of temporal processing dynamics; the fundamental question is the nature of storage and stimulus comparison after stimuli are identified, rendering issues of the timecourse of stimulus recognition less important.

Uncovering task and stimulus structures that improve learning also drives research on the nature of learning. For example, a range of studies have explored the role that variability in irrelevant features of stimuli plays in helping learners identify the aspects of stimuli that are important for differentiating categories; this benefit of variability holds across grammar learning (Gómez, 2002), motor skill learning (Catalano & Kleiner, 1984), word learning (Rost & McMurray, 2009, 2010), word segmentation (Bortfeld & Morgan, 2010; Singh, White, & Morgan, 2008; Singh, 2008), reading (Apfelbaum, Hazeltine, & McMurray, 2012), and even learning to land planes (Huet et al., 2011). This suggests that stimulus variability is a general aid to learning. Uncovering the domain-generality of this principle of learning expands understanding of the nature of learning and how stimulus characteristics affect this learning. However, as with other questions about the nature of learning, these studies of stimulus characteristics remain agnostic about the role of temporal processing dynamics in learning; instead, these studies emphasize the use of comparing stimuli across trials, and are thus more focused on longer time scales than processing within an individual learning event. Temporal processing dynamics are again treated as external to the main question regarding the nature of learning.

High-level investigations into the nature of learning are essential for a comprehensive understanding of the learning process. Yet these broader questions of learning concern the macro-structure of the learning system, and thus are unconcerned with more detailed aspects of individual learning events. Although these approaches offer impressive insight into how learning occurs, they are not as useful for understanding how learning interacts in real-time with ongoing perceptual processes.

1.1.2 Questions on the mechanisms of learning

More specific studies on the mechanisms of specific forms of learning provide a detailed view of the learning process. However, these studies typically rely on stimuli that are designed to limit the contribution of temporal processing dynamics in order to more directly focus on other aspects of learning. Classic approaches to studying learning often rely on stimuli are incredibly simple – for example lights and tones – and that may require very little in the way of processing. When investigating how organisms learn to link stimuli and responses, relying on stimuli that have highly-predictable (and often automatic) responses is quite useful. For example, in plotting the neural circuitry underlying associative motor learning, Freeman and colleagues rely on eye-blink conditioning (Freeman & Steinmetz, 2011; Ng & Freeman, 2012); the stimuli in such a technique are simple tones or lights that are unambiguous and quite rapidly processed. The timecourse of this stimulus processing is fairly irrelevant to the investigation of the neural pathways of the learning, because the relevant processing is fairly automatic and rapid.

The use of stimuli designed to have minimal online processing difficulty also aids studies of learning in more complex domains. Studies exploring fast-mapping abilities in young children, for example, rely on stimuli with little phonological overlap to emphasize the relevant aspects of the study (specifically, how well children can apply mutual exclusivity to identify referents in a task and encode these associations into memory; e.g., Horst & Samuelson, 2008; Spiegel & Halberda, 2010). Words like "cheem" and "fode" are quite distinct, and elicit minimal competition from one another. Using such distinct stimuli can be necessary to identify effects with a locus in learning, as opposed to in real-time processes that are engaged at study or test (e.g., Zhao, 2013). Similarly, studies aimed at detailing visual categorization often rely on quite simplified stimuli that can be recognized quite quickly, with little competition (Booth & Waxman, 2002; Nosofsky, 1988; Posner & Keele, 1968). More complex stimuli are unnecessary to answer questions about the nature of representation or the types of information learners use to identify new stimuli, and so are an effective choice for these studies.

Such simplified stimuli are quite valuable for highlighting the effects of interest in these studies, and they help researchers arrive at quite impressive understanding of learning. However, these stimuli are also designed to reduce real-time processing demands for participants during learning, and thereby block investigation of interactions between online processing and learning.

1.2 Word learning as a model domain

Word recognition is a key domain in which temporal processing dynamics could impact learning. When hearing a word, the perceptual signal comes in across time. The listener is often unable to identify what word they are hearing until the end of the word (or sometimes even well after the offset of the word; Bard, Shillcock, & Altmann, 1988; Connine, Blasko, & Hall, 1991; Grosjean, 1985). Yet listeners begin making predictions about the word they are hearing from early points in the acoustic signal; listeners activate multiple word-forms that are consistent with the acoustic signal, and then continuously update these activations to reflect incoming information (Gaskell & Marslen-Wilson, 1999; Marslen-Wilson, 1987). Ambiguity abounds throughout this process, and often learners must rely on auditory information later in the word to interpret what they have already heard (Connine et al., 1991; Ganong, 1980; McMurray & Jongman, 2011; McMurray, Tanenhaus, & Aslin, 2009). Word recognition is thus a temporally-extended process with competition occurring throughout.

Because word recognition processes are complex and protracted over time, learning the mappings between words and their meanings may present a particularly compelling way to examine the issues of how real-time processes and learning relate. If word learning initiates before competition has resolved, it must cope with many competing word-forms. As such, word learning offers a strong test case for the timing at which learning occurs.

A detailed understanding of these word recognition processes allows precise predictions about how they may affect word learning processes. The following sections detail the temporal dynamics of word recognition, highlighting aspects of incremental processing of incoming acoustic information, graded and parallel activation of word-forms, and competition and inhibition between word-forms that are active in parallel. These facets of processing are then discussed with reference to different forms of learning to determine how such processing dynamics might influence the representations formed during word learning.

1.2.1 Temporal dynamics of word recognition

The nature of lexical processing is such that word recognition is a complex temporal process, and individual word-forms are rarely active independently. Ongoing ambiguity and competition throughout lexical processing challenge the capacity for the learner to identify the appropriate word-form to form new mappings. These temporal processing dynamics are well-understood and offer insight into the different predictions of how word learning may proceed if learning is occurring continuously during lexical processing versus waiting until processing is complete.

1.2.1.1 Incremental processing

From the earliest portions of the speech signal, listeners begin to activate words that are consistent with the input (Allopenna et al., 1998; Magnuson, Dixon, Tanenhaus, & Aslin, 2007; Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, 1987; McMurray, Clayards, Tanenhaus, & Aslin, 2008; Toscano, Anderson, & McMurray, 2013; Zwitserlood, 1989). This activation begins quite early; within 200 ms of the onset of the word, listeners begin preferentially making eye movements to words that are compatible with the acoustic signal over words with mismatching phonological forms.

As more of the word is heard, some of those words that were initially activated are suppressed (Allopenna et al., 1998; Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, 1987), seeming to suggest that listeners access the appropriate word-forms by tracking which words are fully compatible with the input to that point (Gaskell & Marslen-Wilson, 1997; Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, 1987; Norris, 1994). However, as processing proceeds, words that mismatched at onset but are compatible with later portions of the word are activated (Allopenna et al., 1998; Connine, Blasko, & Titone, 1993; Marslen-Wilson, Moss, & van Halen, 1996; McMurray et al., 2009), particularly if these other words are phonetically close to the heard speech signal (Andruski, Blumstein, & Burton, 1994; Creel, 2012). Listeners are thus not relying on a purely left-to-right phonological match between the signal and words in the lexicon to determine lexical identity, but instead exhibit more flexibility in online lexical activation.

These incremental activation processes are sensitive to quite subtle aspects of the acoustic signal, including sub-categorical details that reflect upcoming segments. Listeners use small acoustic change to make predictions about what words they are hearing, and use these predictions to narrow the scope of words they consider (Dahan, Magnuson, Tanenhaus, & Hogan, 2001). Similarly, sub-categorical changes in the duration of speech segments help listeners determine the syllabic structure of words in order to more accurately segment the speech stream (Salverda et al., 2006; Salverda, Dahan, & McQueen, 2003).

Words are not activated individually, nor are they activated in an all-or-none manner. Instead, many words are activated in parallel, and these words' activations reflect their degree of fit to the input and the level of competition from other words. Activation grows as more acoustic information is encountered and competition suppresses words that do not match the input. Throughout lexical access, a variety of words are accessed to varying degrees.

1.2.1.2 Graded/parallel activation

Zwitserlood (1989) showed that listeners activate multiple lexical candidates matching the acoustic signal. She used a cross-modal priming task to determine whether sentential context biases listeners to preferentially activate candidates consistent with the input. Instead, she found that even in contexts that were strongly biasing, listeners showed semantic priming to multiple senses of the word; upon hearing the onset "capt-," listeners activated semantic relatives of both *captain* and *captive*. The visual world paradigm offers further evidence of parallel activation; when given a display with phonologically-related items, listeners fixate both targets and competitors throughout the trial, time-locked to when the signal is consistent with the signal (Allopenna et al., 1998). Critically, these fixation patterns map closely onto models of lexical processing that posit parallel activation (Allopenna et al., 1998; McClelland & Elman, 1986). Marslen-Wilson showed that the degree of activation varies in part as a function of word frequency; high frequency words elicit greater priming of related words than do low frequency words, suggesting that the high frequency words were activated more strongly (Marslen-Wilson, 1987).

While multiple words are activated, these activations are graded, such that words that more closely match the acoustic input are activated more strongly. Listeners hearing mispronounced words activate the semantic networks of the correctly-pronounced form in accordance with the degree of mismatch; when hearing *gat*, listeners show greater priming for *dog* than when hearing *wat* (Andruski et al., 1994; see also, Creel & Dahan, 2010; White & Morgan, 2008; White, Yee, Blumstein, & Morgan, 2013). Similarly, sub-categorical mismatches show graded patterns of recovery from mispronunciations; when hearing *barakeet*, listeners are faster to access the lexical form *parakeet* if the word-initial /b/ is closer to the b/p acoustic boundary (McMurray, Tanenhaus, & Aslin, 2002; see also, Utman, Blumstein, & Burton, 2000). Listeners maintain graded activation of words to the extent that these words are viable referents.

1.2.1.3 Competition between parallel-active words

Having several words active in parallel with varying degrees of activation results in competition during lexical access. In order to determine which word is being heard, the listener must suppress the coactive words. These competitive processes affect both the speed and accuracy with which an item is recognized; words with greater competition are less efficiently accessed (Dahan, Magnuson, Tanenhaus, et al., 2001; Goldinger, Luce, & Pisoni, 1989; Luce & Pisoni, 1998; Magnuson et al., 2007, 2003).

Words whose phonological competitors are high frequency exhibit slowed recognition relative to words with low-frequency neighbors (Dahan & Gaskell, 2007; Dahan, Magnuson, & Tanenhaus, 2001; Luce & Pisoni, 1998; Magnuson et al., 2007; Prabhakaran, Blumstein, Myers, Hutchison, & Britton, 2006; though see, Morrison & Ellis, 1995). Competition is also elevated for words with many neighbors, particularly if those neighbors are high frequency (Chen & Mirman, 2012; Cluff & Luce, 1990; Goldinger et al., 1989; Luce & Pisoni, 1998; Vitevitch & Luce, 1998, 1999). Words that reside in such high density phonological neighborhoods are more difficult to recognize in noise (Luce & Pisoni, 1998), are more slowly identified as words (Luce & Pisoni, 1998) and are more slowly linked to visual referents (Magnuson et al., 2007, 2003).

Competitor effects during lexical access are closely time-locked to the acoustic signal. Words with a large number of phonological competitors at word onset show very early changes in activation, and this competition persists throughout activation. Meanwhile, words whose competitors are more evenly distributed across the word show overall later competition effects than those with primarily onset competitors (Magnuson et al., 2007). Sub-categorical mismatch studies also exhibit immediate competition when conflicting acoustic information occurs; when cross-splicing the onset of the word *neck* onto the final consonant of the word *net*, Dahan and colleagues found decreased fixations to a visually-presented net, even if no image of a neck was present (Dahan, Magnuson, Tanenhaus, et al., 2001). However, this competition only emerged if the onset portion spliced onto the final consonant came from another word (that is, the onset of *nep* does not elicit competition), suggesting that competition emerges on the basis of word-word inhibition, rather than on the basis of acoustic mismatch.

Computational descriptions of lexical access suggest that competition processes are not confined to activation of the word-form, but instead cascade through processing (Chen & Mirman, 2012; Gaskell & Marslen-Wilson, 1997, 1999). When competition slows or weakens activation, access to semantic information is also impaired. Gaskell and Marslen-Wilson (1997, 1999) developed the Distributed Cohort Model to investigate how phonological information propagates to semantic access. In this model both phonological and semantic representations are distributed, with no mediating lexical units. During processing of a word, competition is conceptualized as a blending of the distributed representations of the various active words. When several phonological word-forms are simultaneously active, semantic access becomes more diffuse. Phonological competition thereby affects the facility with which semantic representations are accessed.

Chen and Mirman (2012) built on this conceptualization of interaction between phonological and semantic interaction, but used a very different model architecture and included more explicit dynamics in their model. This model simulates competition via lateral inhibition between localist units rather than averaging across distributed representations. However, they come to similar conclusions as Gaskell and Marslen-Wilson (1997, 1999) regarding the cascading influence of phonological information. The addition of dynamics in this model also leads to predictions that the cascading competition is closely time-locked to the phonological competition; as multiple word-forms become active and compete, this competition results in immediate changes in semantic activation. The Chen and Mirman model further adds consideration of competition between semantic representations, and argues that this form of competition affects phonological processing.

Apfelbaum, Blumstein and McMurray (2011) found empirical evidence for cascading competition (see also, Zwitserlood, 1989). This study used the visual world paradigm to measure consideration of semantic relatives of target words as a proxy for semantic activation (e.g., Yee & Sedivy, 2006). They found that words from dense phonological neighborhoods exhibited reduced semantic activation. These changes began quite early, signaling immediate semantic effects of phonological competition. Competitive processes are thus not encapsulated, but affect numerous components of the language processing stream.

1.2.1.4 Summary

The word recognition process is characterized by continuous (in both time and activation) adjustments in activations of several candidate word-forms. Listeners begin activating words from the earliest moments in the acoustic signal, and then update these activations throughout hearing the word. Several words are active in parallel throughout this process, and these words compete with each other for recognition. Although this processing typically leads to a single word dominating processing, throughout the process many other words receive brief, partial consideration. These parallel activations could impact the formation of links between sounds and meaning during word learning, depending on whether learning the word-referent mappings occurs before competition is resolved.

1.3 What determines the timing of word learning?

If learning is conceptualized as the formation of links between potential word-forms and referents (as in an associative learning account), learning that is continuous in time from the earliest points of lexical activation may lead to mappings between a referent and a variety of co-active word-forms. During periods of competition, all active word-forms will be linked to the same referent. However, if learning waits until competition is resolved, lexical processing dynamics pose little threat to the formation of appropriate word-referent mappings; mappings are only formed when a single word-form is fairly unambiguously active.

Given the rather protracted timecourse of word recognition, it may be possible to determine whether learning occurs from the earliest points of the acoustic signal. That is, investigating whether learners form mappings during periods of temporary lexical ambiguity offers insight into the timing of lexical encoding. This would further understanding of both how people learn words and how learning operates more generally. In the following sections, different theories of word learning are contrasted to determine their individual predictions about whether learning could entail parallel associations with competing word-forms. Different forms of learning are then contrasted to determine what features are necessary in a learning system to elicit learning before competition is resolved.

1.3.1 Theories of learning

Much of the debate over the nature of word learning centers on the degree to which associative learning is used to map a word-form to its referent. Evidence of learning during parallel lexical activation may offer some headway in this debate. Although virtually everyone agrees that at least some degree of associative learning is necessary, at least for early word learning (e.g. Golinkoff & Hirsh-Pasek, 2006; Waxman & Gelman, 2009; Werker & Curtin, 2005), many argue that this form of learning plays a limited role, especially as learners become more sophisticated (Fennell & Waxman, 2010; Golinkoff & Hirsh-Pasek, 2006; Namy, 2012; Waxman & Gelman, 2009). Instead, these theorists posit that learners rely on more explicit learning systems and domain-specific word learning tools to acquire word-meaning mappings. These theories advocate that learners use word learning constraints (Golinkoff, Mervis, & Hirsh-Pasek, 2008; Markman & Hutchinson, 1984; Markman, 1990; Mervis & Bertrand, 1994), hypothesis testing (Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013) and social inferences (Fennell & Waxman, 2010; Goldstein, King, & West, 2003; Golinkoff & Hirsh-Pasek, 2006; Hall, Williams, & Belanger, 2010; Johnson, Ok, & Luo, 2007) that are specifically tuned to helping the child learn language. In contrast, others argue that associative accounts are capable of quite impressive learning and abstraction (Apfelbaum & McMurray, 2011; Gliozzi, Mayor, Hu, & Plunkett, 2009; McMurray, Horst, & Samuelson, 2012; Samuelson & Smith, 1998; Sloutsky & Robinson, 2013; L. B. Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; L. B. Smith, Jones, Yoshida, & Colunga, 2003; Thiessen, 2011). Such accounts rely on domain-general mechanisms which elicit learning that appears domain-specific; as the learning system encodes statistical information from the environment, it develops abstract representations that appear specialized despite deriving from general mechanisms.

In this section, these two broad approaches to word learning are contrasted. Much of the work debating the form that word learning takes emphasizes the problem of referential ambiguity (Quine, 1960; Siskind, 1996), so this section relies heavily on how the different learning approaches deal with this challenge. When a learner hears a novel word, she must identify the referent from an infinite array of possible choices. For example, if an explorer encounters a speaker of an unknown language pointing to a rabbit running across a field and saying “gavagai,” there is no way to be sure that the speaker meant the rabbit – he could just as easily have meant to indicate the scene, the grass, the act of running, the rabbit's ears, the collection of parts that comprise the rabbit, or an infinite number of other possible referents (Quine, 1960). Given a single instance of hearing a word and seeing a scene, identifying the correct referent is impossible without some additional constraints on how to learn. As words are rarely provided to children in isolated naming instances, learners must rely on additional forms of information to extract word-referent mappings. Both explicit/constraint-based theories and implicit/associative theories have been proposed to account for this central difficulty in word learning; depending on which theory is adopted, very different predictions for the timing of learning emerge.

1.3.1.1 Constraint-based theories

The classic solution to the problem of referential ambiguity is the use of word learning constraints to narrow the scope of potential referents. These constraints are often conceptualized as domain-specific operators that guide infants to correct word-referent mappings (Golinkoff et al., 2008; Markman & Hutchinson, 1984; Mervis & Bertrand, 1994). The whole-object constraint biases learners to presume that a novel word refers to a whole object rather than its constituent parts (Markman, 1990). When the listener hears "gavagai," this constraint eliminates theories that this refers to the ears of the rabbit or to "undisconnected rabbit parts." The novel name - nameless category (Mervis & Bertrand, 1994) constraint similarly biases the learner to presume that when a novel referent is heard, it refers to an object in the display that does not yet have a name. This has been suggested as the means by which children fast map, where they reliably select the referent of a novel word when only one referent in the display is novel (Carey & Bartlett, 1978; Mervis & Bertrand, 1994); children use the knowledge that most objects only have one name to infer that the novel name must apply to whatever in the scene does not yet have a name. The mutual exclusivity constraint operates in tandem with these other constraints, biasing the learner to presume that each object has only a single name (Houston-Price, Caloghiris, & Raviglione, 2010; Markman & Wachtel, 1988; Markman, 1990). When a novel word is encountered, the learner knows to ignore objects that they already have names for as potential referents.

Although rarely explicitly discussed, these constraint-based approaches operate at a high-level of abstraction; constraints operate based on the learner identifying a word-form, then determining whether this word-form is novel, before determining what its referent is. For such operations to occur, a fairly concrete decision as to the form of the word must be made; if competition is ongoing, the learner is unable how to apply the constraints to the current word. This reliance on abstract representations thus necessitates that much of the dynamics of lexical activation have settled before any learning proceeds.

1.3.1.2 Associative and statistical theories

Associative theories of word learning suggest that referential ambiguity is accommodated by tracking co-occurrence between words and possible referents across instances (Siskind, 1996; K. Smith, Smith, & Blythe, 2011; Yu & Smith, 2007, 2012); if the learner tracks that typically when the word *dog* is heard, a furry, four-legged creature is present, she can deduce that this animal is the intended referent of the word. Although a single trial is insufficient to accurately learn the correct word-referent mapping, the learner combats referential ambiguity within a learning instance by learning across several instances. As more word-referent pairings are learned, later word learning situations are easier as fewer items are unnamed, and thus few are potential referents for a novel name (Yu & Ballard, 2007; Yurovsky et al., in press).

This form of learning functions quite differently from constraint-based approaches; rather than the learner using explicit tools to determine word-referent mappings, the learning is mostly implicit. Tracking the statistics in the input is thought to be fairly automatic, without any need for the learner to form explicit hypotheses. A potential way to maintain these various statistics is through associative learning, where co-occurrence statistics are maintained in weights between representations. Across trials, learners must maintain several partial associations linking the words and pictures that have co-occurred. This parallel set of weak mappings is quite distinct from the constraint-based approach, which often highlights specific hypotheses about individual word-referent pairings.

Associative accounts are not incompatible with constraints, however the method of instantiating these constraints is quite different in associative models. Using domain-general mechanisms, such as competition and interactivity, many word-learning constraints naturally emerge from associative systems (Kachergis, Yu, & Shiffrin, 2012; McMurray et al., 2012; Namy, 2012; L. B. Smith et al., 2002; Yu & Smith, 2012). Thus the presence of effects that are predicted by constraints is not evidence that the learner is relying on constraints. However, emergent constraints leave room for more nuanced application of these constraints (and may offer more natural extension into cases where constraints must be loosened, as in bilingualism: Byers-Heinlein & Werker, 2013; Curtin, Byers-Heinlein, & Werker, 2011; Houston-Price et al., 2010).

Many formal models of associative word learning have been proposed (Gliozzi et al., 2009; Mayor & Plunkett, 2007; McMurray et al., 2012; Regier, 2005; Samuelson, 2002; Schafer & Mareschal, 2001; Westermann & Mareschal, 2004; Yu & Smith, 2012). These models vary greatly in terms of architecture, forms of representation, learning principles and the fundamental problems they are trying to solve. However, none of these models incorporate the temporal processing dynamics of word recognition. Instead, these models rely on simplified representations of word-forms, with most models representing lexical activation in a word learning event by unambiguously activating a single lexical representation (Gliozzi et al., 2009; McMurray et al., 2012; Yu & Ballard, 2007). This simplification is adopted either because these models are not concerned with the timing of lexical processing, or because not enough is known about how the temporal dynamics of processing affect learning to make clear predictions. Yet this focus on learning issues that are external to temporal processing dynamics may miss important interactions during learning.

Although often described as “dumb” process, capable only of thoughtlessly forming links between co-present stimuli, and without the capacity to show higher-level abstraction (Fennell & Waxman, 2010; Golinkoff & Hirsh-Pasek, 2006; Waxman & Gelman, 2009), many forms of associative learning are quite sophisticated and powerful (Kachergis et al., 2012; Rescorla, 1988; Sloutsky, 2009; L. B. Smith et al., 2003; Wasserman & Miller, 1997). Associative learning is far from a monolithic process, and instead includes a range of forms of learning. Depending on the learning environment, timing of stimuli and presence of feedback, a learner may rely on different associative methods to form appropriate mappings between stimuli. This richness of forms of learning makes associative learning a flexible tool for encoding information in a range of domains. However, it also complicates consideration of how processing dynamics might interact with an associative system. Considering the primary dimensions of variability across forms of associative learning can help clarify how the dynamics of lexical activation would be predicted to affect different associative word learning systems.

Associative learning thus may predict that learning occurs continuously during lexical competition processes; however, associative learning includes a range of different forms of learning systems. In the following section, two classes of associative models are contrasted: supervised and unsupervised models. This major distinction in forms of associative learning is critical in forming predictions about the timing of learning. Depending on which of these forms of associative system is predicted to occur in a word learning task, learning may or may not be predicted to occur continuously in time.

1.3.2 Supervised vs. unsupervised associative learning

Although associative learning offers the potential for interaction between online lexical activation processes and word learning, such interaction depends on the form of associative learning that takes place when learning words. Associative learning is a broad term often used to refer to a wide variety of learning mechanisms that may only be loosely related by the fact that they all form associations between elements. However, they also vary on many dimensions. The presence or absence of supervision is among the most pertinent dimensions of variation between forms of associative learning; learning that relies on feedback to update activations and mappings functions quite differently than learning without feedback, and often leads to quite different solutions to mapping problems. Critically, this feature of associative systems also alters predictions about whether and how online processes impinge on the learning process.

1.3.2.1 Supervised learning

Supervised associative learning develops a representation of some structure using some form of feedback signal that guides the learner to a precise, functional encoding of the relationship between stimuli. In the classic model, this feedback is an explicit error signal that corrects the learner when a mistake is made (Rescorla & Wagner, 1972; Rumelhart, Hinton, & Williams, 1986). Such an error signal not only informs the learners that they have made a mistake, but also provides the correct answer. This allows rapid adjustment of associations in order to develop a more functional set of links. Such a trial-and-error learning process is quite powerful, and it is capable of learning mappings that are not linearly-separable when using an appropriately structured system (e.g. with intermediate representations; Rumelhart, Hinton, & Williams, 1986). However, there are doubts about the plausibility of such a system as a primary determiner of learning because of its reliance on a teaching signal that not only signals whether the learner was correct, but also what the correct answer was (De Sa & Ballard, 1998; O'Reilly, 1996, 2001). Such feedback is often not available to the learner, casting doubt on the value of using an external teaching signal as a guide to learning. This is particularly the case in language learning, where parents often fail to correct misproductions made by their children (Marcus, 1993; though see, Chouinard & Clark, 2003; Gruendel, 1977).

A more subtle form of supervised learning uses the exact same learning rules, but relies on prediction to form internally-generated error signals. As the cognitive system processes information, it constantly forms predictions about upcoming information (Clark, in press; Elman, 1990). These predictions can then be validated against the form that this later information takes. Although this provides a form of error signal, it differs qualitatively from a classic teaching signal. Prediction error doesn't tell the network how to act, merely what to expect. This can be accomplished using standard feedback mechanisms that detail the correct pattern for all outputs, or it can be done using feedback that only signals whether the learner is correct or incorrect, without detailing the correct output for all possible stimuli.

Nonetheless, this approach to supervision has been offered as a way to rectify the implausibility of classic error-driven approaches' reliance on complex teaching signals (Elman, 2009; O'Reilly, 1996). Such a learning system is able to form very precise predictions, and is even adept at accommodating linguistic tasks (Elman, 1990). Forming predictions and adjusting representations based on prediction error has even been suggested as a foundational principle of human information processing (Clark, in press).

Many consider prediction-based learning to be unsupervised rather than supervised, as the feedback signal is internally generated. The category boundary between supervised and unsupervised learning systems is quite fuzzy, as there is no consensus over what defines supervision. I place prediction models into the supervised case as the way that associations are changed is somewhat akin to that of truer supervised models (O'Reilly, 1996): both forms of learning update representations using later-occurring information to alter mappings and minimize error. In both external error-driven learning and prediction learning, the learner relies on an error signal to gauge whether an output is correct, and uses this error signal as the basis of adjusting associations or weights. A more thorough analysis of supervised vs. unsupervised learning could elucidate several dimensions of variability and a gradient of supervision; however, such a treatment is outside the scope of this dissertation.

Prediction-based supervised models have a more nuanced relationship with temporal processing dynamics. These models are inherently temporal, as predictions are typically formed before confirming or disconfirming evidence is encountered. Because the error signal is contingent on the later-occurring information, there is an inherent delay in the prediction system. To some extent, the degree to which a prediction-based system would be influenced by processing dynamics is contingent on when this later occurring information is presented. However, in most models relying on prediction error, predictions are made on the basis of fairly high-level information; when making predictions about word-referent mappings, predictions may take the form of specific hypotheses of which words pair with which referents, or which referent should be seen given a word (akin to hypothesis testing approaches to word learning, Medina et al., 2011; Trueswell et al., 2013). If the learner is forming predictions on the basis of single word – single referent pairings that are determined post-competition, the temporal dynamics of processing likely would play little role in how these associations are formed. These predictions can then be gauged against high-level lexical information when the word-referent pairing is again encountered; the learner need not maintain multiple partial word-referent mappings, and the predictions can be quite precise.

In both the above forms of supervised learning, the timing of learning is unambiguous: learning occurs when a supervisory signal is received. Whether this signal is a teacher detailing the correct answer or later-occurring information signaling whether a prediction was accurate, there is a distinct time when representations can be updated. These error signals typically arrive after processing dynamics are relatively settled. For a teaching signal, the learner must resolve competition enough to make a response to a single candidate, after which the error signal is given. For a prediction learning system, competition within the system must resolve to a sufficient point that the learner can form a prediction; by the time the later-occurring information is received to determine whether the prediction was accurate, the learner is likely to have very few competitors active.

Supervised learning systems thus do not allow continuous learning – this form of learning updates representations at a given time late in processing, when much of the dynamics of ongoing processing is resolved. Supervised learning predicts little impact of representations activated in parallel on the learning process.

1.3.2.2 Unsupervised learning

Supervised learning appears to obviate concerns that the temporal dynamics of lexical activation impinge on the formation of new word-referent mappings. In each form of supervised learning, the system has some signal for when to update representations based on the time when the feedback is received. In most cases this occurs after the stimuli have been processed, and signals how to adjust representations using these post-competition representations. Unsupervised learning operates quite differently; the learner links whichever stimuli or representations are co-present (Hebb, 1949). This does not require a teaching signal, instead using repeated learning instances to form accurate mappings; things that consistently co-occur are linked strongly, whereas things that only co-occur sometimes receive much weaker links (Apfelbaum & McMurray, 2011; Yu & Ballard, 2007; Yu & Smith, 2012). This gradually leads to the appropriately paired representations to dominating the learned mappings.

Many descriptions of word learning suggest that some form of unsupervised learning occurs. Although this form of learning lacks the explicit teaching signal that helps form fast associations, given the number of words children learn and the rapidity with which they learn them, some form of learning without a teacher is likely necessary.

However, few models of word learning offer true analogs of unsupervised learning (though see, McMurray, Horst, & Samuelson, 2012; Yu & Ballard, 2007; Yu & Smith, 2012). Instead, many rely on some form of supervised learning. Regier's (2005) exemplar model of word learning, for example, utilizes a formal error-driven teaching signal to train the model (see also, Schafer & Mareschal, 2001). This choice is justified by commenting that this learning rule is associative, and thus is appropriate for modeling associative forms of word learning. Other classes of associative models endorse unsupervised learning, but use training regimes that highly constrain the associations that can form; many such models only present a single stimulus at a time, eliminating any forms of referential ambiguity during training (e.g., Gliozzi et al., 2009; Mayor & Plunkett, 2007). Although these models rely on unsupervised learning rules, this presentation format provides a form of soft feedback for the model, as the word-referent mappings are always unambiguous. This single-object/single-referent presentation is quite distinct from the word learning situation that most engenders theories of unsupervised associative learning: discovering word meanings across instances, when many possible referents are available.

There have been few true forays into studying word learning as a truly unsupervised associative process. A notable exception is work on cross-situational word learning (K. Smith et al., 2011; Yu & Smith, 2007, 2010). In this task, learners are predicted to rely on unsupervised associations across trials in order to learn accurate word-referent mappings (though they may also be using some form of prediction). Because several potential referents are available in any display, the learner has to track several word-referent pairings in parallel throughout many instances of hearing the word identified. The learner never receives an error signal to correct her mapping, but instead forms this though associating whichever referents are present when a word-form is encountered (though see, Medina et al., 2011; Trueswell et al., 2013). Formal models of this word learning task show that this form of learning is accommodated nicely by models that rely solely on co-occurrence to form associative links (Frank, Goodman, & Tenenbaum, 2009; McMurray et al., 2012; Yu & Smith, 2012).

Unsupervised learning is often implemented using a mechanism analogous to Hebbian learning between neurons (McMurray et al., 2012; Munakata & O'Reilly, 2003; O'Reilly, 2001). In Hebbian learning, neurons that are coactive augment their connection strength, and those whose activation is in antiphase decrement this strength. This implements a form of unsupervised associative learning, wherein the learner strengthens connections between co-present stimuli or representations. As such, unsupervised associative word learning could be thought of as a form of Hebbian learning, where the neurons being linked represent lexical or referential information (and where the learning is between populations of neurons rather than individual neurons).

However, temporal processing dynamics may complicate this link; Hebbian learning does not have a built-in signal for when to learn, so learning may occur throughout an event. Most models of word learning that suggest unsupervised learning have not attempted to capture the temporal dynamics of the learning event. Instead, models typically make the simplifying assumption that a single word-form is fully activated throughout the learning trial, and that this post-competition representation of the word-form is mapped onto a post-competition representation of the referent. Yet if learning if learning is truly Hebbian, it begins well before competition completes. This means that multiple word-forms in parallel would be mapped to the referent to varying degrees.

Prior models of word learning ignore the temporal dynamics of lexical activation because this falls outside the purview of investigation; the McMurray et al (2012) study was concerned with how learners deal with referential ambiguity in the visual scene and the capacity for simple mechanisms like competition to account for complex behaviors, so using representations of auditory information that match those in processing was unnecessary. However, these temporal dynamics may prove quite meaningful in how learning proceeds. Adding such dynamics to an unsupervised learning system could alter the way that mappings are formed during word learning.

Whereas supervised learning offers a signal for when to initiate learning, unsupervised systems have no such explicit signal; there is no feedback telling the learner that the learning instance is complete, nor a particular time when the learner can test their prediction. Instead, learning can occur at any time during processing. A pure interpretation of associative learning as a Hebbian process suggests that learning should occur whenever information is available to associate. This would suggest that if a referent is available during periods of lexical competition, learning would nonetheless commence. This would lead to learning occurring while multiple words are active. Thus unsupervised associative word learning might interact strongly with lexical processing dynamics.

1.4 Summary and predictions

Word recognition involves protracted periods of competition, with many words active throughout this process. The listener utilizes acoustic information as soon as it is received, and continuously updates lexical activations to reflect incoming information. These processes occur automatically on the basis of acoustic-phonetic match between the signal and word-forms known by the listener and result in multiple candidate words being activate for brief periods during spoken word recognition. During learning, these same processes of parallel activation and competition are likely to play out (and perhaps be enhanced, as the word-forms are more novel). If learning occurs before this competition is resolved, the learner may form spurious associations between the referent and these parallel-active word-forms.

The form of learning dictates whether such online processing effects on learning should occur. If learning operates through explicit word-learning constraints, such effects are unlikely; these theories suggest that learning operates by activating a single word-form, then using constraints to determine what referents this word-form could represent.

Additionally, these theories argue that learning occurs at a single time per learning event, rather than continuously updating throughout the learning event. However, if learning is more implicit, reliant on unsupervised associative mechanisms, the learner need not necessarily resolve competition to identify a single word-form to perform learning; learning may use repeated updates throughout the course of processing.

Within associative accounts, supervision may eliminate learning during periods of competition. When a supervisory signal is given, this offers a specific time at which to update representations. Typically, this signal occurs after stimulus processing is complete; in order to make a response and garner feedback, the learner has to resolve much of the competition. As such, supervised associative learning predicts little influence of real-time lexical activation processes on learning. Yet unsupervised learning requires no such signal for when to learn; an unsupervised associative word learner could begin forming mappings during periods of lexical competition, when several word-forms are active to varying degrees, and then update these mappings throughout the event. Such learning would lead to associations between referents and competing word-forms.

By creating a situation which encourages unsupervised associative learning of novel words, the experiments in this dissertation aim to test whether unsupervised learning does indeed occur continuously in time. Evidence of the formation of parallel associations during learning would indicate that this form of learning does not rely on processes that signal when competition has resolved in order to begin learning. This would deepen understanding of how learning proceeds in cases when stimulus recognition processes are temporally extended. Simultaneously, this would provide evidence that word learning can occur using unsupervised associative learning. Given that I will be studying word learning in adults, this runs counter to theories that as learners become more experienced, they cease relying on associative mechanisms to acquire word-referent mappings (Golinkoff & Hirsh-Pasek, 2006; Waxman & Gelman, 2009), and it supports implicit, associative theories of word learning (Apfelbaum &

McMurray, 2011; McMurray et al., 2012; Yu & Smith, 2007) rather than hypothesis testing approaches (Medina et al., 2011; Trueswell et al., 2013).

1.5 Relevant studies

Although no studies have directly investigated how real-time processing dynamics interact with word learning, several studies focus on related issues. Many word learning studies are concerned with how lexical competition affects the ability for learners to acquire word-referent mappings, although this competition has not been considered in its real-time form, nor has it been used to investigate the possibility that it gives rise to erroneous associations. Similarly, the speed with which people at different ages can access lexical representations and resolve competition has received considerable attention, although never with respect to how this affects the way that new lexical entries are formed. The literature thus suggests that studies of the interface between lexical processing and word learning are of interest and as yet undone.

1.5.1 Competition and word learning

Competition between similar word-forms is a central issue in studies of word learning and lexical processing. Many studies of word learning focus on the ability of young infants to learn new word-forms that differ minimally (Mani & Plunkett, 2010; Pater, Stager, & Werker, 2004; Rost & McMurray, 2010; Swingley & Aslin, 2000; Thiessen, 2007; Werker, Fennell, Corcoran, & Stager, 2002); although infants are adept at discriminating the sounds of their native language, competition between overlapping word-forms impacts their capacity to learn some words.

The Lexical Restructuring Model (LRM) suggests that such competition between similar word-forms is essential to help children master their native language phonology and become more adept word learners (Charles-Luce & Luce, 1990, 1995; Metsala & Walley, 1998; Walley, Metsala, & Garlock, 2003; though see, Swingley & Aslin, 2002). As learners develop more comprehensive lexica, they also develop a need to represent words more precisely. In this way, phonological competition serves as a tool to guide learning of the phonological structure of a language.

Competition continues to affect the ease with which new words are learned, even into adulthood (Storkel, Armbrüster, & Hogan, 2006). Yurovsky and colleagues conducted a comprehensive analysis on how and when competition affects word learning in unsupervised learning situations for adults learning new words (Yurovsky et al., in press). They showed that at both global and local levels, competition between words makes it more difficult for learners to form appropriate word-referent mappings. When multiple word-forms compete for mappings with the same referent, learners are able to form parallel mappings, but these mappings are less robust than mappings without competition.

Computational models of word learning rely on competition as a key feature to demonstrate empirical patterns of learning. Gliozzi and colleagues' self-organizing map model of word learning relies on competition to exhibit learning after only a few trials (Gliozzi et al., 2009), whereas many models without competition are incapable of showing such fast learning (e.g., Westermann & Mareschal, 2004; see, McMurray et al., 2012 for simulations demonstrating the need for competition to show fast mapping). In a dynamic model that accommodates a range of learning phenomena, McMurray, Horst and Samuelson (2012) used competitive processes to show why learners show different types of knowledge depending on task (see also, Samuelson, Schutte, & Horst, 2009), and why their behavior in the moment often conflicts with their long-term learning. Competition thus plays a central role in models of word learning, suggesting that it is a criterial component of natural word learning systems.

Although these studies clearly demonstrate that lexical competition impacts word learning, they investigate competition across the word-form, in a more global fashion. For example, the neighborhood effects that are central to the LRM rely on competition metrics that compare the complete word-form in a position-invariant way to other words in the lexicon (Charles-Luce & Luce, 1990, 1995; Metsala & Walley, 1998; Walley et al., 2003). This approach disregards the temporal dynamics of competition, as the more global competition metrics are sufficient to answer the relevant research questions. Yet competition processes during lexical access are dynamic, and the timecourse of this competition could affect learning. That is, the way that processing dynamics play out within a learning event could alter the information that is encoded during this event.

1.5.2 Real-time processing in children

Early childhood marks a time of remarkable word learning, with children rapidly increasing the size of the lexica (Ganger & Brent, 2004). During this period of massive word learning, real-time processing abilities are maturing; the second year of life shows impressive changes in the speed of lexical processing (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998), and lexical competition effects are quite strong at this age (Stager & Werker, 1997; Swingley & Aslin, 2007; Swingley, 2009); although children activate words more slowly, they still show incremental processing and parallel activation (Swingley, 2009). Throughout childhood, children exhibit difficulty with competition and inhibition, in both language and non-language tasks (Cepeda & Munakata, 2007; Munakata, Snyder, & Chatham, 2012). As such, real-time processing dynamics are likely quite complex during periods when many words are learned.

However, the ramifications of these real-time processing differences have not been explored in terms of learning. Although a few studies have investigated the degree of lexical competition effects on word learning (Hoover, Storkel, & Hogan, 2010; Storkel, 2001, 2002), these studies have ignored temporal processes within individual trials. Whereas these studies offer important insight into how phonological overlap across the lexicon affects word learning, they offer no headway into understanding whether the processes to cope with this competition during online processing impinge on the representations formed during learning.

Research directly investigating the temporal dynamics of perceptual processing and their interactions with learning are thus necessary to more fully explore the process of learning new words. These studies can provide a means to understand how learning systems cope with ongoing online processing, to determine whether learning is continuous across time or if instead it waits until competition resolves. Such research simultaneously investigates the nature of word learning more specifically, uncovering the forms of learning that allow learners to build word-referent mappings.

1.6 Outline of the dissertation

This dissertation explores how unsupervised word learning interacts with lexical processing dynamics. Unsupervised learning is relied upon as this is the form of learning that offers the clearest predictions of real-time influences of perceptual processes on the representations that are learned. Thus unsupervised word learning offers a potential window into understanding unsupervised learning more generally. The following chapter details a novel paradigm designed to assess how learning occurs within unsupervised word learning trials. This is followed by several experiments investigating how changes to the time at which stimuli are presented affect the way that words are learned. These experiments demonstrate that when unsupervised learning is encouraged, learners begin forming associations during periods of lexical competition. However, such learning is complicated by the complex interactions of auditory and visual processing across time. Other forms of learning are then investigated to demonstrate that spurious associations are more likely to form in instances when unsupervised learning is encouraged. The results of these studies are then summarized to highlight the value of this initial foray into questions of how perceptual processes impact the learning system, and future directions are suggested.

CHAPTER 2

PREDICTIONS, LOGIC AND GENERAL METHODS

2.1 Timing in unsupervised learning

The different approaches to unsupervised associative word learning predict different influences of real-time processing on the learned lexical representations. This section briefly reviews the different forms of unsupervised associative learning, and details how these approaches handle processing dynamics. A paradigm for investigating this experimentally is then described, including the logic behind several implementational details.

How the dynamics of lexical competition interact with learning depends critically on when learning occurs. Inferential approaches (e.g., the constraint approach) to word learning seem to require some signal that competition is complete before learning begins. Such a signal would ensure that the learner forms more appropriate mappings between those items that win competition, without the formation of spurious associations during periods of lexical ambiguity. Similarly, supervised associative learning must wait until a response can be made and feedback received (and in the meantime, allowing competition to resolve) in order to update learned mappings. Such approaches predict that the dynamics of lexical competition have little effect on how words are learned.

Unsupervised learning could accommodate either learning that waits until competition resolves or more learning that is more continuous in time. It may be possible to gate learning using some monitor of ongoing competition in order to signal when learning should begin. Such a monitor may predict little effect of lexical dynamics on learning. In contrast, particularly in unsupervised learning, it may be more parsimonious to eschew such a signal, instead forming associations whenever two representations are active. This simpler learning approach predicts that word learning will initiate as soon as any word-forms are coactive with potential referents. Thus if a referent is present during lexical competition, the parallel-activated word-forms can all be mapped immediately to the referent. Not only would the correct word-referent mapping be formed, but spurious competitor-referent pairings would also surface. Whereas the predictions of inferential and supervised learning are somewhat clear, in the next sections, I discuss the various possibilities raised by unsupervised learning.

2.1.1 Unsupervised learning with a competition monitor

A competition monitor could effectively aid the unsupervised learning system in avoiding spurious associations during parallel activation. Such a monitor needs to track when competition is ongoing in order to signal when an appropriate association can be formed. Although more complex than simply associating any coactive representations, fairly simple processing mechanisms could accomplish such monitoring. For example, dynamic neural field models of cognitive processes often rely on activation exceeding some threshold[1] to gate information between levels of representation (Samuelson, Smith, Perry, & Spencer, 2011; Samuelson et al., 2009; Schutte, Spencer, & Schöner, 2003). With a given threshold, the model does not pass any information to the successive level until enough activation accumulates to surpass this threshold. Although such thresholding creates non-linear dynamics in how the model processes information, it does so without using additional machinery.

A threshold could likewise gate learning to occur only when competition is resolved. Although multiple word-forms are active in parallel throughout the lexical activation process, these activations typically remain fairly weak until sufficient information is available to unambiguously identify the word (Allopenna et al., 1998; McClelland & Elman, 1986). As competition continues, the word-forms that are most consistent with the acoustic signal become more strongly activated while the

[1] These thresholds need not be discrete; more graded thresholds that gate learning in proportion to the degree of competition could produce similar effects.

mismatching forms are suppressed. This typically leads to a “winner” of the competition receiving a high activation value at the expense of the competing forms (McMurray et al., 2012; Samuelson et al., 2011; see also, Zénon, Hamed, Duhamel, & Olivier, 2009). Such competition is consistent with competitive processes between neural activations in the brain (Pauli & O’Reilly, 2008). If word learning relies on an activation threshold, these competitive processes could ensure that no learning proceeds until some representation has a sufficiently high level of activation. This would block learning of weakly-activated competitors early in lexical access, and thereby efficiently cope with the temporal dynamics of word recognition.

An alternative form of learning gate might be a more explicit monitor of the global activation state, such as a measure of entropy in the word recognition system. When many words are active to varying degrees, the system is in a state of high entropy. As competition continues, this entropy decreases until one word dominates activation, and the system is relatively stable. By tracking when the system becomes more stable and then initiating learning, the learner can avoid forming spurious associations during periods of competition. Measures of entropy are a common facet of computational models (e.g., Griffiths, Steyvers, & Tenenbaum, 2007; Harm & Seidenberg, 1999; Kachergis, Yu, & Shiffrin, 2012b; Zhu, Rogers, Qian, & Kalish, 2007). Entropy thus could offer a simple solution to issues of dynamic processes during word learning. More broadly, the learning rule itself could be sensitive to some global measure of the state of the system, and rely on this measure to determine when the system is ready to learn.

2.1.2 Unsupervised learning without a competition monitor

Although simple adjustments to unsupervised learning systems can provide ways to gate learning until competition resolves, there is no empirical evidence detailing whether such gates are used when learning words (or any other type of stimulus). Without such gates, alternative associations are likely to form during unsupervised word learning. No studies have investigated the presence of such associations, leaving the form that unsupervised learning takes an open empirical question.

Without some form of competition monitor, unsupervised learning would occur continuously, whenever representations are active. Indeed, in some computational models utilizing unsupervised learning rules, learning occurs repeatedly throughout a trial, with weights constantly being updated to reflect changes in the activation structure of the system (e.g., McMurray et al., 2012). Such models rely on a pure form of Hebbian associative learning, where learning is an automatic process without any specialized information specifying when to form or augment associations. Such an ungated form of unsupervised word learning would result in learning occurring throughout the lexical activation process, with associations forming while multiple word-forms are active. This would result in associations between a referent and the phonological competitors of that referent's correct word-form. For example, when first learning the name of the small Pacific island nation "Vanuatu," coactivation of words like "vanity" and "vanguard" may lead to small associations between these words and the island.

Research investigating whether learners form associations for referents with both correct and competing word-forms could thus offer insight into the nature of word learning, as well as the timing properties of unsupervised learning more generally. Lack of evidence of spurious associations would argue that learners either do not rely on unsupervised associative learning to learn new words, or that this learning is buttressed with some form of additional monitoring to gate when learning occurs. Evidence for associations formed during periods of lexical competition would support an unsupervised learning account of word learning, and it would argue that unsupervised learning systems are closely coupled to the dynamics present in the domain being learned, with learning occurring continuously in time. This would reinforce the parallels between Hebbian neural learning and unsupervised associative learning, as it would provide evidence that this form of learning is a matter of linking together any coactive representations. It would also provide counterevidence against explicit hypothesis testing theories of word learning (Medina et al., 2011; Trueswell et al., 2013), as these theories posit that only a single word-form is associated with a referent at a time. By determining whether spurious associations form during lexical competition, we can thus develop a deeper understanding of both learning in general and word learning more specifically.

2.2 General methods

In order to gauge whether learners begin forming word-referent mappings during periods of lexical competition, we need a paradigm that manipulates when learning can occur. In cases where referents are available throughout word recognition, learners have the opportunity to begin forming associations immediately, while multiple word-forms are active. However, if the referents are withheld until after competition has resolved, associations can not be formed between the (now-suppressed) competing word-forms and the referent; once the learner identifies that the word being heard is "Vanuatu" and not "vanguard," she can unambiguously map this form onto the referent. Whereas explicit learning and gated unsupervised learning theories predict no difference between these two forms of learning, ungated unsupervised learning predicts that the former case should result in the formation of spurious associations.

The experiments in this dissertation manipulate the relative timing of auditory word-forms and visual referents in an artificial language learning task with adults. The paradigm is designed to determine whether there is evidence of spurious associations in situations that encourage learning during ongoing competition. To this end, adults are taught a new set of words that are designed to elicit coactivation and competition during lexical access (using phonological overlap between the novel words; e.g. *goba* and *gonu*). After performing a phoneme monitoring task to become familiar with the phonological form of these words, the learners complete cross-situational learning trials to determine the word-referent mappings. During these cross-situational trials, referents are presented either concurrently with the word-forms (while competition is unfolding) or after a delay (after competition is resolved). After learning, participants are tested using the visual world paradigm (VWP) to determine whether spurious associations have formed with competing word-forms. In this section, I describe the general methods used by all of the studies (that roughly follow this logic). These choices are justified below to explain how this design allows learning of parallel associations, how the timing manipulation is predicted to affect learning, and how the presence of spurious associations can be measured.

2.2.1 Studying word learning in adults

Although much word learning occurs in early childhood, the experiments detailed in this dissertation focus on novel word learning in adults. These experiments use artificial lexica to train adults on a new set of words (e.g., Magnuson, Tanenhaus, Aslin, & Dahan, 2003; White, Yee, Blumstein, & Morgan, 2013). Word learning in adults may rely on different mechanisms than learning in childhood (Metsala & Walley, 1998; Storkel, 2002; Walley et al., 2003); as the lexicon becomes more densely populated, new processes may take over the learning process. Whereas early word learning appears to rely on very simple associations between acoustic and visual information, more sophisticated word learners may use more complex learning machinery (Golinkoff & Hirsh-Pasek, 2006; Namy, 2012), although this additional machinery may still rely heavily on associative principles. However, such complex machinery predicts little influence of processing dynamics on learning, as described in Chapter 1. The goals of this study are not to investigate the developmental trajectory of word learning, nor to provide a detailed description of natural word learning. Instead, this dissertation aims to uncover more basic characteristics of the word learning system as a method to investigate learning more generally. Adults provide a useful experimental population, as they allow relatively direct investigation of knowledge using more sophisticated tasks that may be required to view the subtle consequences of the spurious associations. Further, this population can more readily handle learning a number of similar word-forms in parallel, as well as completing protracted experimental sessions.

Using novel words allows precise control of stimulus characteristics, which avoids concerns about past experience with the words. Although the question of how parallel activation of known words interacts with learning of new words is quite interesting, it is outside the scope of this dissertation. Instead, these experiments utilize stimuli that are designed to elicit competition primarily with each other, so that differences in learning as a function of timing are more apparent.

2.2.2 Cross-situational learning

In order to train learners on the word-referent mappings, this dissertation relies on cross-situational learning (K. Smith et al., 2011; Yu & Smith, 2007, 2010; Yurovsky et al., in press). Cross-situational learning operates by presenting the learner with a label in the presence of several possible referents. This mimics natural word learning environments, where countless visual stimuli are available upon hearing a referent. Crucially, the learner can not definitively identify the correct referent given this trial alone. However, the auditory-visual pairings are informative across trials; whenever a given label is heard, its corresponding referent is present. Thus, by tracking the co-occurrence statistics of words and referents across trials, the learner can determine which referents are consistently present when hearing specific words. Rather than learning occurring through a single exposure, it spans many instances of word-referent pairings.

This form of learning offers several benefits in studying the temporal dynamics of unsupervised word learning. First and foremost, cross-situational learning most clearly instantiates pure unsupervised learning (though see, Medina et al., 2011; Trueswell et al., 2013), the only form of learning in which may predict the temporal dynamics of processing to matter. Second, cross-situational learning is often conducted in an entirely passive manner (Yu & Smith, 2007); the learner simply looks at the display and listens, without making overt responses. This allows for precise control of trial timing, so that the relative onsets and offsets of the visual and auditory information are precisely known. Adding a response complicates issues of timing and potentially changes the mechanism of learning (learners may use the time of a response as an explicit cue for when to form associations, or it may force them to resolve competition more discretely than they normally would). Most of the experiments in this dissertation use a passive learning regime; however, Experiment 5 investigates the role that responding might play in this form of learning. Third, cross-situational learning is conceptualized as a pure form of associative learning by most researchers (though see, Medina et al., 2011; Trueswell et al., 2013). This provides confidence that this learning task provides a viable environment for continuous unsupervised associative learning. Finally, cross-situational learning entails repeated exposures to word-referent pairings. If learners are forming word-referent mappings on the basis of unsupervised associative learning, additional trials should strengthen these associations, providing a greater likelihood of detecting spurious associations despite more dominant correct mappings.

The nature of the cross-situational learning in this dissertation differed in several ways from many previous cross-situational learning studies. First, in Yu and Smith's (2007) methods, every word on each display is named, although the order of naming is random. This decreases the overall number of trials needed to learn, but also makes control of visual-auditory timing more complex. In this dissertation, only one auditory stimulus is played per trial. Second, the number of alternatives on the display varies widely between different cross-situational studies. Yu and Smith (2007) investigated this and showed that even with six items displayed at once, participants show above-chance performance at test, signaling at least some learning. We present only two pictures at a time, as we are interested in performance for words that are learned quite well, and we rely on a measure of associations that only utilizes trials with correct responses (see section ***2.2.5***). Pilot work showed that including more than two objects per trial resulted in much poorer learning. Third, although the majority of training trials were passive cross-situational learning trials, interim trials were interspersed throughout training, in which participants had to make some form of response. The nature of these interim trials varied between experiments, but their motivations were consistent: to keep participants engaged in the task and to offer some measure of the speed of learning during training. Finally, in our displays, the referent of the competitor word-form was never present in the display. This ensured that any associations between the competitor and the referent arose solely from co-activation, not from actual co-occurrence between these forms and their referents.

2.2.3 Altering the timing of learning

To determine whether learning is occurring continuously throughout real-time processing, we need a task that allows control over *when* associations are formed. By manipulating whether referents are available during periods of temporal ambiguity, we can either allow or disallow the formation of links with coactivated word-forms. Throughout the experiments in this dissertation, we accomplish this by manipulating the relative timing of visual and auditory information in word learning trials. For *synchronous* presentation, the visual and auditory stimuli are presented simultaneously (Figure 2-1A). This presentation format provides referents to map word-forms onto throughout the lexical activation process. For *delay* presentation, the visual stimuli are withheld until the end of the auditory word-form (Figure 2-1B); learners thus have ample time to suppress competitor activation before they can begin forming associations. If learning occurs immediately when visual and auditory stimuli are available, the *synchronous* presentation should elicit interference from competing word-forms, whereas the *delay* presentation should not show such interference. If learning waits until competition resolves, the change in timing should not yield differences in interference

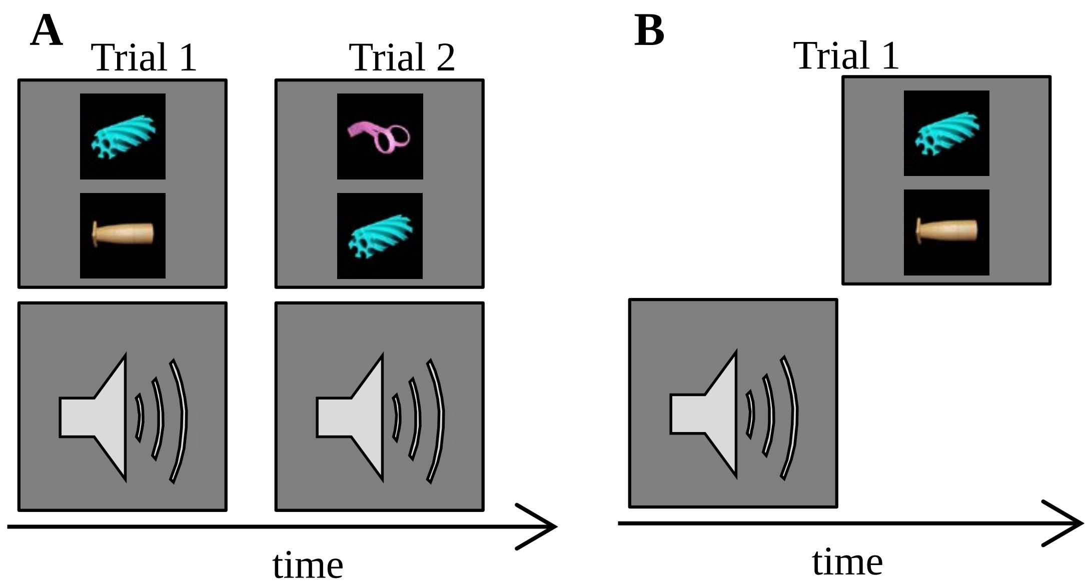

Figure 2-1: Schematic of the trial structure used throughout the dissertation. A) Synchronous presentation of auditory and visual information. B) Delay presentation, with auditory preceding visual information.

from competing word-forms; in both cases, the learner will wait until competitors are suppressed before learning.

2.2.4 Ensuring parallel activation during learning

In order build spurious associations during lexical activation, alternative word-forms must be coactivated when hearing a word. Ample research shows that phonological overlap elicits such parallel activation (Allopenna et al., 1998; Magnuson et al., 2007; Marslen-Wilson, 1987; McMurray et al., 2008; Vitevitch, Luce, Pisoni, & Auer, 1999), even if the overlap is at different positions within the word (Toscano et al., 2013). This coactivation appears to be strongest when the words overlap at onset (Allopenna et al., 1998; Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, 1987; Zwitserlood, 1989); this is likely due to the temporary ambiguity created by onset-overlap: when you have only heard *sand*- there is no way to know whether the word is *sandal* or *sandwich*. Competitor activation is increased when the different phonemes are more featurally similar (Andruski et al., 1994; White & Morgan, 2008; White et al., 2013); the nonword *gat* elicits stronger activation of the word *cat* than does the nonword *wat*.

In order to ensure parallel activation in the experiments in this dissertation, participants in all studies were taught a set of novel words which included substantial overlap. Depending on the experiment and the question being asked, this overlap occurred at different points in the word, and the amount of overlap (both in terms of number of shared phonemes and number of features for unshared phonemes) varied. By using other words that are trained in parallel as the competitor stimuli, concerns about past experience, degree of knowledge and past frequency are eliminated.

However, for words to be activated in parallel, the listener must have access to some representation of the word-forms. In a novel word learning experiment, this means that the learner must know the form of words that are available to be mapped to the referents. To some extent, this can emerge across trials, as learners begin to learn the set of words. However, such across-trial learning leads to minimal coactivation of competing word-forms in the earliest learning trials. As these trials are critical to initial formation of associations, lack of parallel activation in such trials may mask effects of learning during coactivation. Additionally, as some words are learned, mutual exclusivity biases (whether explicit or implicit) serve to limit the amount that related word-forms are considered when learning the names of stimuli (Kachergis, Yu, & Shiffrin, 2012a; Markman, 1990). Although repeated exposure can overcome such biases (Kachergis et al., 2012; Yurovsky et al., in press), it is unclear whether known word-forms that are never encountered in the experiment would ever be considered. Thus, it seemed important to establish enough knowledge of the word forms to create parallel activation of competing words before any meanings are presented, as this maximizes the likelihood that both words are fully activated during learning trials.

To ensure parallel activation even during the earliest word learning trials, participants completed a phoneme monitoring task with all words used in the experiment before beginning the word-referent learning phase of the experiment. Phoneme monitoring does not require any access to word meaning, but it still elicits learning of the phonological word-form, such that this word-form becomes integrated in the lexicon (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; Kapnoula, Gupta, Packard, & McMurray, submitted; Lindsay & Gaskell, 2013). Through repeated exposure in phoneme monitoring, learners acquire representations of the phonological word-forms, allowing parallel activation to emerge from the earliest training trials.

2.2.5 Measuring spurious associations with competing word-forms

In determining whether the timing manipulations affect learning, this dissertation relies on online measures of parallel associations in instances when learners are highly accurate at using word-referent mappings. Although timing may impact the facility with which learners acquire these mappings, identifying parallel associations with competing word-forms provides more comprehensive evidence of learning throughout lexical processing. We focus primarily on cases when the words are learned quite effectively, and investigate the degree to which competitors are considered when the correct referent is identified. This also helps control for changes in target-referent associations as we investigate differences in competitor-referent associations.

Measuring whether the learner has formed associations between referents and phonologically-related word-forms requires a task that demonstrates competition in word recognition performance. If these spurious associations have formed, the learner should show enhanced consideration of competitor words. Eye-tracking in the visual world paradigm (VWP; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) allow detection of quite subtle activation for competitors during lexical processing (e.g., Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Magnuson et al., 2007; Salverda et al., 2006; Toscano & McMurray, 2012). This method is very sensitive to stimulus timing, allowing detection of competition throughout lexical processing.

The learning paradigms used in this dissertation emphasize the formation of spurious associations with word-forms that overlap phonologically. For example, we predict that both *gonu* and *goba* may be associated with the referent of *gonu*, even though *goba* was never heard in the presence of that object. This is because when hearing *go-* in *gonu*, *goba* will be partially active in the presence of that referent.

Online word recognition processes cause words with such overlap to demonstrate competitor effects in VWP experiments (Allopenna et al., 1998; Dahan, Magnuson, & Tanenhaus, 2001; Magnuson et al., 2007; McMurray et al., 2008), even without positing the formation of extraneous associations. This competition arises because of temporal ambiguity in the speech signal, and occurs even if competitor word-forms are not present in the display (Dahan, Magnuson, Tanenhaus, et al., 2001; Magnuson et al., 2007). As such, the presence of competition on its own is not evidence of associations between a referent and its competing word-form. Instead, we look for changes in the *degree* of competitor effects. If continuous encoding leads to the formation of spurious associations, phonological competitors should show greater competition during word recognition. In the experiments in this dissertation, we use competitor effects in the *delay* condition as a baseline measure of online competition. Additional competitor effects beyond this baseline are taken as evidence of associative links between the referent and the competitor word-form.

2.2.6 Multiple exemplar training

Listeners are adept at encoding not just the phonological form of words, but also variable surface forms. They are most accurate at recognizing words as familiar if trained and tested in the same voice and at the same speaking rate (Bradlow, Nygaard, & Pisoni, 1999). Listeners appear to encode sub-categorical information into long-term lexical representations (Goldinger, 1998; Hawkins, 2003; Ju & Luce, 2006; Pierrehumbert, 2003). This form of exemplar learning is also apparent when forming new lexical entries;

Creel and colleagues founds that when taught an artificial lexicon, participants were most accurate when tested on words presented in the same voice as in training (Creel et al., 2008). Rather than processing words solely based on abstract phonological structure, listeners use several indexical and other non-criterial features of the words.

When teaching a learner a novel word, repeating the exact token of the word may lead surface characteristics of the word to be included in its stored representation, as these sources of variability occur consistently with the referent (as in infants; Apfelbaum & McMurray, 2011). Differences in the surface form between phonological competitors may reduce the degree to which learners activate these competitors, as information is present in the signal to disambiguate the words more rapidly. For example, if one member of a pair of cohort competitors has a higher pitch than the other, the learner can identify which word she is hearing well before disambiguating phonological information is encountered, and can thus quite quickly suppress competition. This would lead to depressed competitor effects throughout processing.

In pilot work for this dissertation, we found that single-exemplar training of this sort weakened competitor effects. To combat this concern of exemplar-specific learning, the experiments in this dissertation all included several exemplars of every word. This reduces the ability for learners to memorize the superficial acoustic properties of words so that competitors are more likely to become active during lexical processing. Additionally, the exemplars used at test differed from those used during training, so that learners were unable to rely on surface information to recognize the words, but instead had to rely on more abstract phonological representations.

2.3 Description of included experiments and simulations

The experiments in this dissertation provide a detailed analysis of the conditions under which learners form associations during periods of lexical competition. **Experiment 1** uses the methods above to determine whether displaying visual referents from the onset of auditory presentation leads learners to show larger interference from phonological competitors at test. **Experiment 2** explores the effect of providing additional processing time at the offset of the auditory word. Specifically, this experiment investigates whether enhanced interference from simultaneous auditory-visual presentation is mediated by additional learning that occurs after competition resolves. **Experiment 3** manipulates timing quite differently; in this experiment, visual stimuli are present well before the onset of auditory information. This design controls for overall trial duration without providing additional processing time after lexical competition resolves; it also allows visual processing to affect auditory activations before any word is heard.

Experiment 4 asks whether changes in representations are possible within-participants when only some of the words use a simultaneous presentation format. This strengthens the argument that changes in learning are a result of online processes during specific learning trials affecting the representations that are formed. Finally, **Experiment 5** contrasts alternate types of learning to determine what aspects of training are necessary to see evidence of continuous encoding. Specifically, this experiment contrasts unsupervised learning in which learners make a response on every trial (*active unsupervised learning*) with supervised learning, in which feedback is provided after each response. This experiment provides insight into the role that making explicit responses plays in the timing of learning, as well as into how receiving a feedback signal affects the representations that are formed.

The results across the experiments are then situated in the current literature on learning in general and word learning in particular. This discussion details the contributions that these experiments make for understanding the principles that underlie the learning process and the value in considering domain-specific effects on general learning processes. Finally, limitations of the current experiments and necessary future directions are described.

2.4 Predictions

Throughout the series of experiments in this dissertation, the primary question is whether changes in relative stimulus timing affect the representations that are formed. Forms of learning that argue that word learning relies on mapping post-competition word-forms to their referents predict little effect of stimulus timing; although these approaches might predict that some timing conditions might slow the speed with which learners acquire representations, they would predict that once learning is at ceiling, the associations that have formed should be quite similar regardless of timing during learning. These approaches thus predict that across the experiments, there should be no difference in the degree of interference at test between those training with *synchronous* presentation and those training with *delay* presentation. Similarly, these accounts predict that both the passive learning in the earlier studies and the more active (and supervised) learning of Experiment 5 should elicit similar representations; feedback is not needed to suppress spurious associations, as these forms of learning predict that no such associations form in the first place.

Similarly, accounts of learning that suggest unsupervised learning is supported by competition monitors that gate when learning occurs predict little effect of the timing of auditory-visual presentation. Although *synchronous* presentation provides the opportunity to start learning during periods of lexical competition, the competition monitor delays learning until competition has resolved. This equates learning in cases with *synchronous* presentation to that of *delay* presentation, when the referents are not available until after competition has resolved. As in the explicit or supervised learning cases, these accounts of unsupervised learning also predict little change in interference at test from different forms of learning. Although they predict that feedback-driven learning is a fundamentally different process than that done in the passive learning cases, in both scenarios participants do not form spurious associations.

Finally, if word learning relies on unsupervised mechanisms that learn immediately, regardless of competition structure, learning in the *synchronous* case is predicted to exhibit enhanced competition over that in the *delay* case. When learners begin forming association, multiple word-forms are active, leading to associations between competing word-forms and the referent. At test, this would culminate in increased consideration of the referents of competing word-forms, as the target word-form has become associated with these objects. This account of pure unsupervised word learning also predicts major differences when feedback is given; in this scenario, learning is no longer continuous in time, but instead relies on an error signal at a specified time in the trial to correct learned mappings. However, the case of *active unsupervised learning* is more nuanced; if this condition shows evidence of increased competition, it suggests that learning is not time-locked to response generation. If instead this condition shows no evidence of great interference during *synchronous* training, it suggests that learners use the representations formed during response generation as the basis of word-referent learning.

CHAPTER 3

EXPERIMENT 1: EVIDENCE FOR CONTINUOUS LEARNING

The framework I have presented thus far suggests that if an unsupervised learning system is used for learning novel words, the temporal dynamics of lexical processing may alter the course of learning. As the to-be-learned words are heard, multiple words consistent with the acoustic input are activated in parallel. If learners are building associations between words and meaning while these multiple word-forms are competing for recognition, associations could form between the appropriate word-form and its referent as well as between the competing words and the referent. These parallel associations would exacerbate online competition effects after the words are learned.

Enhanced competitor effects in cases that encourage learning during parallel activation would thus be evidence of temporally-continuous learning during word recognition. Experiment 1 tests this prediction. It utilizes a design that allows some participants to start learning while lexical activation processes are ongoing, while others must wait to initiate learning (during which time lexical competition is predicted to resolve). Thus, this experiment provides a method for determining whether learners show evidence of forming spurious associations during parallel activation.

3.1 Background

This experiment set out to establish whether the representations formed during word learning are affected by temporal dynamics during word recognition. To this end, participants were taught a novel set of words with phonological overlap that encourages parallel activation during word recognition. By manipulating the relative timing of acoustic information and the visual referents during training, we provide some participants the opportunity to begin forming associations between words and objects before word recognition processes complete, whereas others are not afforded the chance to learn these mappings until competition has had time to resolve. Increased competition during the VWP trials in the former group after learning would indicate that having the potential to learn words immediately leads learners to form both correct and spurious word-referent mappings.

3.2 Methods

3.2.1 Participants

Fifty-seven participants from the University of Iowa community completed the study. Participants were paid $15 or received partial course credit for their participation. All participants self-reported normal hearing and normal or corrected-to-normal vision. Data from 46 participants were included in analyses: 23 participants from the *synchronous* condition, and 23 from the *delay* condition. An additional 11 participants were run but excluded from analysis due to low accuracy at test (below 75% correct on either the *onset competitor* or *offset competitor* words; nine participants total: six in the *synchronous* condition, three in the *delay* condition) or poor eye-tracking (2 participants).

3.2.2 Design

Participants were taught word-referent pairings for eight novel words; four of these words had phonological competitors in the set that overlapped at word onset, and four had phonological competitors that overlapped later in the word. The *onset competitors* should elicit early and strong coactivation, whereas the *offset competitors* should have weaker coactivation that occurs later during word recognition.

Participants were first familiarized with the auditory word-forms through the phoneme monitoring task. This ensured that listeners would coactivate the overlapping word-forms during the referent learning portion of the task. During word-referent training, half of the participants learned the pairings with *synchronous* timing, in which auditory and visual stimuli were presented at the same time; the other participants learned with *delay* timing, in which visual stimuli appeared after auditory stimulus offset.

These word-referent training trials utilized cross-situational learning to teach participants the word-referent mappings. On each trial two visual stimuli were presented along with one auditory stimulus labeling one of the visual stimuli. The auditory stimulus always corresponded to one of the two referents. The alternative referent was never the phonological competitor of the target, to ensure that there was no basis in co-occurrence to form a spurious competitor-referent association. The alternative referent was drawn randomly from the other six items, giving an approximate co-occurrence rate of 1/6 with each foil referent. Participants made no responses during these learning trials. Intermittently during training, participants completed trials where a single visual stimulus was presented along with an auditory word, and they indicated whether this word was the appropriate label for the visual stimulus. These trials were included to ensure that participants maintained attention throughout the experiment.

After training, participants were tested using the VWP. Four objects were presented on the monitor, including a target and its phonological competitor. An auditory stimulus was played, and participants clicked on the picture they thought matched this stimulus. Eye-movements were tracked to determine how much participants looked to both the target item and its phonological competitor, relative to the other, phonologically-unrelated items on the screen.

We included word pairs that overlapped at onset as well as those that overlapped at offset. This was done for two reasons. First, if all words have onset competitors but none have offset competitors, the word onsets become less informative about word identity; learners could attend solely to the second syllable of the words and learn all the words perfectly. Second, the differences in the timing of competition offer interesting predictions for how learning might change in a continuous-learning account. The different points of overlap lead to parallel activation at different points in the recognition process. This may lead to changes in the degree of spurious associations that are formed. Specifically, *onset competitors* may induce more pronounced spurious associations, as these words exhibit very large competition effects, and they are likely quite well suppressed by word offset.

3.2.3 Stimuli

All participants learned a set of eight novel words mapped to images of novel objects. The word-object pairings were randomized for every participant. The visual stimuli were color photographs of objects that are hard to identify, excised from background context, and presented on a black background (See Figure 3.1)

The auditory stimuli were two-syllable consonant-vowel-consonant-vowel (CVCV) words that adhered to the phonotactic rules of English, but do not have referents (Table 3-1). These words included two pairs of *onset competitors*, which shared the first CV, as well as two pairs of *offset competitors*, which shared the second CV. The same set of words was learned by participants in both timing conditions.

Figure 3-1: Visual stimuli used in Experiments 1-3.

Table 3-1: Phonetic transcriptions of words taught to all participants in Experiments 1-4.

Onset competitors		**Offset competitors**	
busɑ	bureɪ	dʒɑfɑ	meɪfɑ
goʊnu	goʊbɑ	pɑtʃoʊ	lutʃoʊ

Auditory stimuli were recorded from a male native speaker of English in a sound-treated room. The speaker produced approximately 35 exemplars of each word. These exemplars were then isolated in Praat (Boersma & Weenik, 2013), and the 26 clearest exemplars of each word were selected. These selected stimuli were adjusted so that they had consistent peak amplitudes, and 100 ms of silence was added to the beginning and end of each.

3.2.4 Procedure

The procedure was identical for all participants except for the timing of stimulus presentation during the training trials. All participants learned the same set of words and had identical trial structure during pre-exposure, interim testing trials and VWP trials. This ensured that differences between groups are due to changes that occurred during word-referent learning, as all other factors are the same between groups. An Eyelink II head-mounted eye-tracker was used to monitor fixations during the final portion of the experiment; subjects were calibrated before beginning the experiment, and completed a drift correct before the VWP trials to ensure that the track was still accurate. No eye-movement data were collected during the other segments of the experiment.

3.2.4.1 Pre-exposure

Before beginning word-referent training, all participants were pre-exposed to the auditory word-forms several times. This pre-exposure was used to teach the participants the auditory word-forms, which encourages parallel activation of competing word-forms from the earliest word-referent training trials. Before beginning the pre-exposure, participants were advised that the words heard during this portion of the experiment would later be used in the word-referent learning part of the experiment.

During pre-exposure, participants completed a phoneme monitoring task. This task has proven effective for teaching the phonological form of words (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; Kapnoula, et al., submitted; Lindsay & Gaskell, 2013). Participants heard words played over headphones and were instructed to press the spacebar if the word contained an "O" sound, and to do nothing if the word did not. Real-word examples with this sound in various lexical positions were provided. Half of the words used in the study contained this sound, so participants had to make regular responses. If the participant pressed the spacebar after hearing the word, the trial ended immediately. If no spacebar press was registered within 2000 ms of the onset of the auditory stimulus, the trial ended (the longest word had an offset at 740 ms, ensuring plenty of time for participants to process the stimuli). There was a 500 ms inter-trial interval. Each word was played eight times during the pre-exposure block, in random order for a total of 64 trials. Each token heard during this phase was randomly chosen from the 26 exemplars for that stimulus. Overall, those participants whose data were included in the analysis were quite good at this task; they correctly identified 93% of tokens with "O" sounds, and had a false alarm rate just below 10%.

3.2.4.2 Word-referent training

After completing the pre-exposure phase, participants were told that the words they had been hearing during the phoneme monitoring task would now be paired with visual objects, and that their job was to determine which words went with which objects. These word-referent training trials used cross-situational learning to teach the word-referent pairings (Yu & Smith, 2007). Although most prior cross-situational learning studies included 4-6 referents on the screen at a time, we only presented two items. This was done to ensure that participants learned the words to near ceiling performance during a fairly short training session; experiments utilizing many referents often show learning that is modest (though above chance; e.g. Yu & Smith, 2007 never found performance above 80% correct with more than two items displayed per trial), and pilot work for this dissertation showed that the available participant population was unable to learn effectively with more than two objects displayed at a time.

At the beginning of each trial, a blue dot appeared on the screen. The participant clicked this dot to start the trial. The auditory stimulus began 100 ms after the dot was clicked. For participants in the *synchronous* condition, two visual objects appeared simultaneously with auditory onset, on the left and right of where the dot had appeared. In the *delay* condition, these objects were displayed 1000 ms after the onset of the auditory stimulus; this was well after the offset of all recorded tokens used in the study (mean duration[1]: 614 ms; max: 740 ms). For both conditions, the visual stimuli remained on the screen for 800 ms, after which the screen was blank for 550 ms before the next trial began. No responses were required during these training trials.

Every word was presented with its correct referent 32 times during training, for a total of 256 such trials. Items were blocked, such that all eight words were heard in a random order before any word was repeated. The location of the correct referent of the auditory stimulus was randomly selected as the left or right object on each trial. The alternative referent on the screen was never the phonological competitor of the target (e.g. when the word *goba* was played, the *gonu* object was never in the display), but was instead randomly selected from the remaining six words in the training set. Never presenting the referent of the phonological competitor ensured that bottom-up statistics did not endorse forming competitor-referent associations; in fact, the competitor is the

[1] Mean and max durations are presented including the 100 ms of silence at onset, but not including silence at offset.

least likely association to form given the presentation, as no other word-referent pairs had a co-occurrence probability of 0. This manipulation ensures that learned spurious associations with phonological competitors could only arise from parallel activation during lexical access, and never result from trained associations.

Auditory stimuli during training (and interim trials) included 16 of the 26 exemplars of each word. The specific 16 tokens used were randomly selected for each participant, and each token was hear twice during word-referent training (interim auditory stimuli were randomly selected from the 16 training tokens for that participant).

3.2.4.3 Interim testing trials

Throughout training, participants were sometimes asked to complete trials where a response was required. During these trials, a single picture was displayed and an auditory word was played, and participants responded whether they thought this word-referent pairing was correct or incorrect by clicking on “match” or “mismatch” on the screen. These trials were included to ensure that participants maintained attention throughout training; pilot work using only passive-learning training showed that some participants simply clicked through the trials without attending to information. Adding these intermittent test trials improved learning. These trials also offer a coarse measure of the trajectory of learning throughout the training phase.

Participants completed four of these interim trials after every 32 training trials (every four blocks), for a total of 32 interim trials. Before each set of interim trials, participants were reminded that they would need to make a response for these trials. Each word was used on four interim testing trials over the course of the experiment: twice with the correct visual referent, and twice with a mismatching referent. When the referent was mismatching, it was never the referent of the word’s phonological competitor. For both groups of participants, the visual and auditory stimulus initiated simultaneously, and the visual stimulus remained on the screen until a response was made. After a response was given, the screen was blank for 550 ms before the next trial began.

3.2.4.4 VWP testing trials

Finally, the presence of spurious associations was tested using the VWP. Participants were presented with four visual stimuli while hearing a word, and were instructed to click the referent of the word. Fixations were tracked throughout these trials to determine the degree that both the target and its competitor were considered during the trial. Each trial contained two pairs of phonological competitors: one *onset competitor* pair and one *offset competitor* pair that was phonologically distinct from the *onset competitor* pair; pairings were randomly chosen for each participant and were consistent throughout the test trials. Item location was randomized for each trial.

Before the first test trial, a drift correction was performed on the eye-tracker. Participants then completed 320 test trials, with a drift correction every 32 trials. Each block of eight trials included one repetition of each target word. The first 20 blocks (160 trials) were identical to the last 20 blocks. The trial structure was identical for both groups of participants. The four visual referents appeared in the corners of the screen, along with a blue dot in the center of the screen. After 500 ms, the dot turned red, at which point participants clicked on the dot to start the trial. An auditory stimulus identifying one of the referents was played, and the display remained on the screen until a response was registered. After the participant responded, the screen went blank for 300 ms before the next trial or drift correct.

Auditory stimuli during the VWP trials consisted of the 10 exemplars of each word that were not used during training for that participant. Each of these repetitions was used four times during testing. Because these tokens were never encountered during training, participants could not identify the referent using surface form information of the auditory stimuli.

3.3 Results

3.3.1 Interim testing trials

No data were recorded during the cross-situational training trials, as no responses were required. However, the results gathered during the interim trials offer some insight into the learning process (Figure 3-2). The results from this analysis should be considered quite cautiously, as the testing task is quite a bit easier than that used later in the experiment. As Samuelson and colleagues have shown, changes in test format can exaggerate the degree of knowledge exhibited by the learner (McMurray et al., 2012; Samuelson et al., 2009). However, these results do offer a coarse measure of whether the timecourse of learning differs between the two groups.

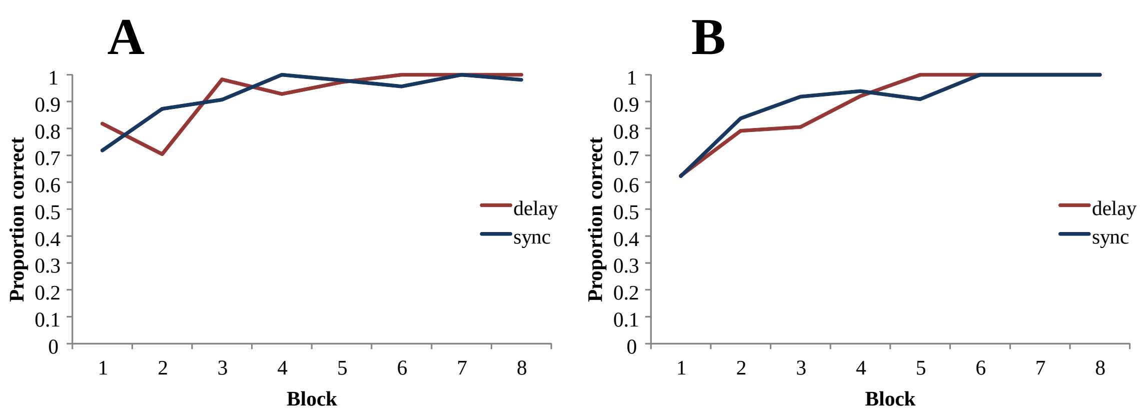

Figure 3-2: Experiment 1 – Accuracy of responses during yes/no interim trials, by block and timing condition. A) Onset competitor trials. B) Offset competitor trials.

Participants proved quite adept at this task from very early; even by the first set of these interim trials, participants were accurate on 69% of trials (chance was 50%). To determine whether the word-type or training condition affected the speed with which participants learned the referents, we analyzed the accuracy data from these trials using a mixed effects model (Baayen, Davidson, & Bates, 2008; Jaeger, 2008) using the lme4 package in R (Bates & Sarkar, 2011). This model used a binomial linking function. The model included group (contrast coded: *synchronous*: -.5; *delay*: +.5, between subject), word-type (contrast coded: *onset*: -.5; *offset*: +.5, within-subject) and block (eight blocks, coded as true block number, within-subject) as fixed factors. Participant and auditory word were random intercepts (adding random slopes word-type within participants did not improve fit of the model for the VWP data by χ^2 test; p>.1; all subsequent analyses thus did not included random slopes to maintain consistent model structure). The correlations between fixed effects were all R<.07.

There was a significant main effect of block (B=.80, SE=.077, Z=10.3, p<.0005), as participants were more accurate for later blocks. There was a marginal effect of word-type (B=-.82, SE=.44, Z=-1.9, p=.061), as participants were slightly more accurate for *onset competitor* words than for *offset competitors*. No other effects or interactions approached significance (training condition: B=-.37, SE=.52, Z=-.71, p=.48; condition × word-type: B=.11, SE=.84, Z=.13, p=.90; condition × block: B=.14, SE=.15, Z=.90, p=.37; word-type × block: B=.14, SE=.15, Z=.93, p=.35; condition × word-type × block: B=-.091, SE=.31, Z=-.29, p=.77). Although there was a trend toward faster improvement in the *synchronous* group, this difference was not reliable. Both groups thus showed a comparable course of learning using this coarse measure, and the *onset competitor* items may have been slightly easier for participants to learn.

3.3.2 VWP testing trials

The VWP task required subjects to select the correct referent out of four possible choices, including the referent of the phonological competitor. Due to online competition processes, both groups should show competitors effects, such that they look more to the referent of the phonological competitor than to the unrelated objects. Evidence of false associations between phonological competitors and referents is defined operationally as increased looks to competitors beyond that seen from online competition when participants identify the correct target. That is, if the *synchronous* presentation led to spurious associations, these participants should show greater consideration of competitor objects than the *delay* participants. Although changes in accuracy could also signal increased competition from false associations, such accuracy effects could emerge if the change in timing simply decreased overall learning. Thus these analyses concern those cases in which learning was quite effective, so effects are likely not a result of overall poorer learning in one group.

3.3.2.1 Accuracy

Participants whose performance was below 75% on either *onset* or *offset competitor* words were excluded from the analysis of the VWP data; most of these participants were at or near chance (25%) performance. The vast majority of remaining participants were near ceiling performance on the VWP task. Overall, the participants included in analyses chose the correct target on 98% of trials. This performance was high across training groups and word types (Figure 3-3). However, a mixed effects model on accuracy for these participants (condition and word-type, contrast coded, as fixed effects, participant and word as random intercepts; binomial linking function) showed a significant effect of word-type (B=.49, SE=.15, Z=3.2, p=.0012), as *offset competitor* words (M=98.8% correct) were identified more accurately than *onset competitors* (M=98.1% correct; in contrast to what was found in the interim testing trials). There was no main effect of timing condition (B=-.28, SE=.39, Z=-.7, p=.47), however the interaction was marginally significant (B=-.51, SE=.29, Z=-1.8, p=.074), as the *synchronous* group showed a larger effect of word-type than did the *delay* group.

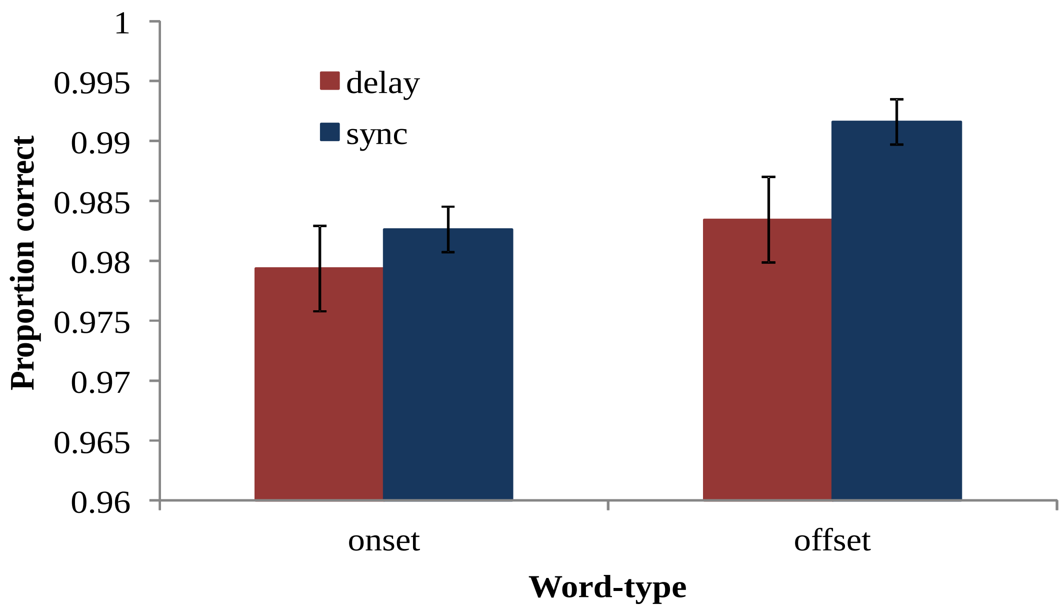

Figure 3-3: Experiment 1 – Accuracy for included participants in the VWP trials by word-type and training condition. Error bars represent standard error for that condition.

Overall, these data showed no evidence of worse learning for the *synchronous* group; if anything, this group trended toward slightly better performance. As such, increased competition effects in the eye-tracking analysis for the *synchronous* group can not be attributed to poorer overall learning. This suggests that comparison of competitor effects is viable, as all participants appeared to have learned the correct word-form-referent mappings quite well. Increased competition in the *synchronous* condition thus indicates that the learners also formed additional, spurious mappings while learning.

3.3.2.2 Eye movements

The eye-tracking analysis only considered those trials in which participants selected the correct target. On these trials, we were primarily concerned with fixations to the referents of the phonological competitor of the target (henceforth: competitor). However, fixations to items in a display are not independent of one another; if the participant is fixating the competitor, she can not simultaneously be fixating the target or one of the unrelated objects. Additionally, changes in timing during training may affect both the learned associations and more general dynamics of looking behavior (for example, increasing looking to everything on the screen regardless of its phonological match to the input). Thus rather than considering only the fixations to the competitor, we need to consider these relative to fixations to other objects in the display to determine whether changes in competitor fixation are greater than predicted by overall changes in looking behavior.

Figure 3-4 shows the timecourse of fixating the competitor and unrelated objects across condition and word type. It suggests that timing condition appears to affect looks to the visual competitor, but it also changes how participants look to the average unrelated item (because there are two unrelated items and only one visual competitor, analyses use the mean of fixations to both of the unrelated items). Because there are changes in how much participants consider objects that are completely unrelated to the auditory stimulus, we need a measure of how strongly participants are considering the competitor that considers the looks to the competitor relative to those to unrelated items.

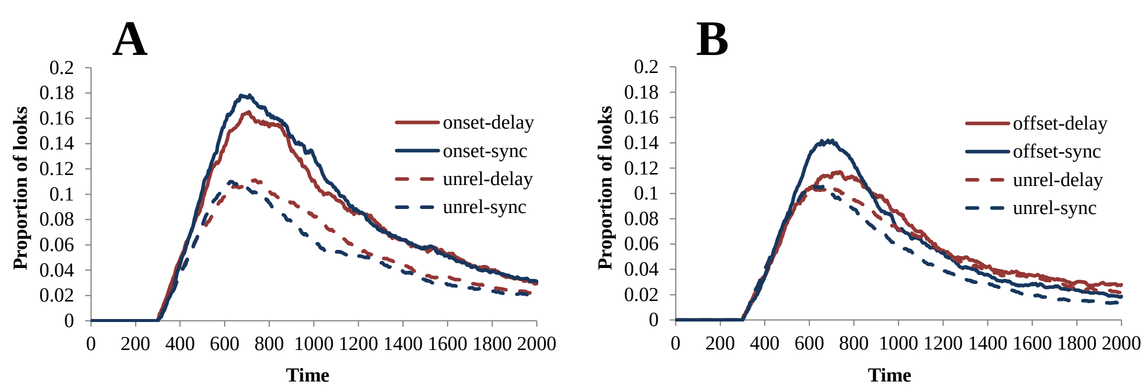

Figure 3-4: Experiment 1 – Proportion of looks to competitor items and average unrelated items across time by training condition. A) Onset competitor trials. B) Offset competitor trials.

That is, we want to examine whether participants look to the competitor object more than they look to any object in the display, and whether the disparity between competitor and unrelated is greater for *synchronous* than for *delay* participants.

The traditional approach to such a comparison is to use an ANOVA with object type (competitor or unrelated) and training condition (*synchronous* and *delay*) as factors. An interaction between these factors would indicate an effect of object type (and presumably a competitor effect). However, this approach violates assumptions of independence necessary for statistical analysis. Looks to the unrelated items are not independent of those to the competitor; a participant can only fixate one object at a time, so increased competitor looks necessitate decreased unrelated looks. Thus including the looks to the two object types separately in the analysis is inappropriate. Instead, looks to the two object types should be collapsed into a single measure that is more appropriate for such proportional data.

For instance, looks to unrelated items could be subtracted from looks to competitor items. For visualization purposes, such an approach is presented in Figure 3-5,

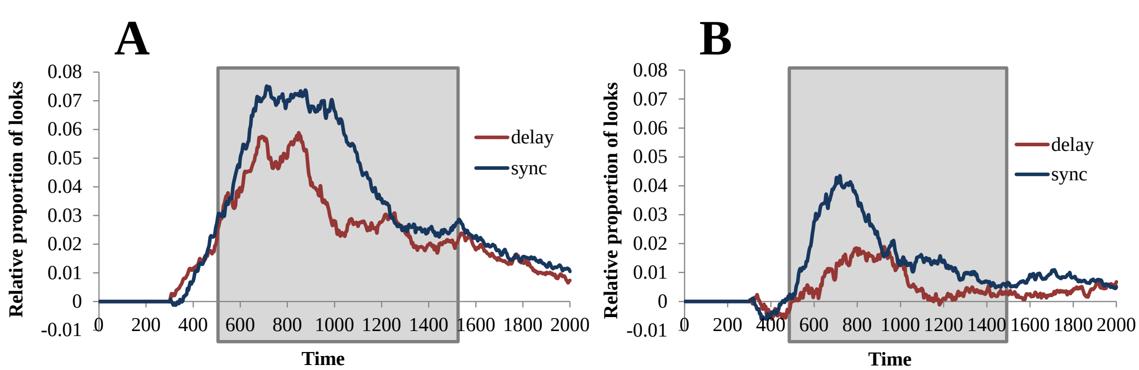

Figure 3-5: Experiment 1 – Relative proportion of looks to competitor objects (competitor – average unrelated). A) Onset competitors. B) Offset competitors.

which plots the difference at each time point between the competitor and the average unrelated item. This represents the raw increase in proportion of looks to the competitor relative to unrelated visual stimuli (i.e. competitor looks – unrelated looks). However, while visually informative, this subtraction analysis is not the most appropriate way to consider competitor effects, as these are proportions rather than linearly-scaled variables. When comparing proportions, a ratio may be more appropriate. For example, if we consider the difference between looks to different types of visual items, a 5% difference in the proportion of looks between items is more notable if one is at 5% and the other is at 10%, compared to if one is at 75% and the other is at 80%. A ratio captures this distinction well, as the increase from 5% to 10% doubles the odds-ratio, whereas an increase from 75% to 80% only increases by a factor of 1.06. For fixations to items other than the target, this is particularly important; as the target is fixated more effectively later in the trial, looks to other items get squashed. This suppresses the difference between looks to competitors and unrelated items. A ratio more effectively compares the proportion of fixations relative to how many fixations are made to those two objects overall. Thus, a form of odds-ratios was used to analyze the competitor fixations relative to the unrelated fixations.

Further, because we wish to use linear statistical approaches (mixed effects models, which are powerful, capable of capturing multiple random effects and nested fixed effects, and are well-developed for such data), raw odds-ratios may be inappropriate, because they do not scale linearly and do distribute in a Gaussian manner. Thus, all of the analyses use log-odds-ratios, which take the log of the ratio of looks to the competitor and the unrelated items (Roembke & McMurray, submitted). This style of analysis is similar to the empirical logit transformation commonly used to scale proportions for a linear analysis, but rather than looking probability vs. chance, they allow us to compare two probabilities[2].

Odds-ratios were computed for each stimulus for each subject. These ratios were computed using the average number of looks to a given item over the time window of 500 ms to 1500 ms (Figure 3-6). The onset of this window was selected as this was when competitor fixations became consistently greater than unrelated fixations for both groups of subjects for both *onset* and *offset competitors* (the initial 300 ms can not elicit looks related to the auditory signal, as all auditory files started with 100 ms of silence, and it takes approximately 200 ms to program an eye-movement in response to auditory information: Matin, Shao, & Boff, 1993). The offset was chosen based on when target looks neared asymptote.

For each participant and each stimulus, we computed the average proportion of looks to the competitor objects and the average of the looks to the two unrelated objects

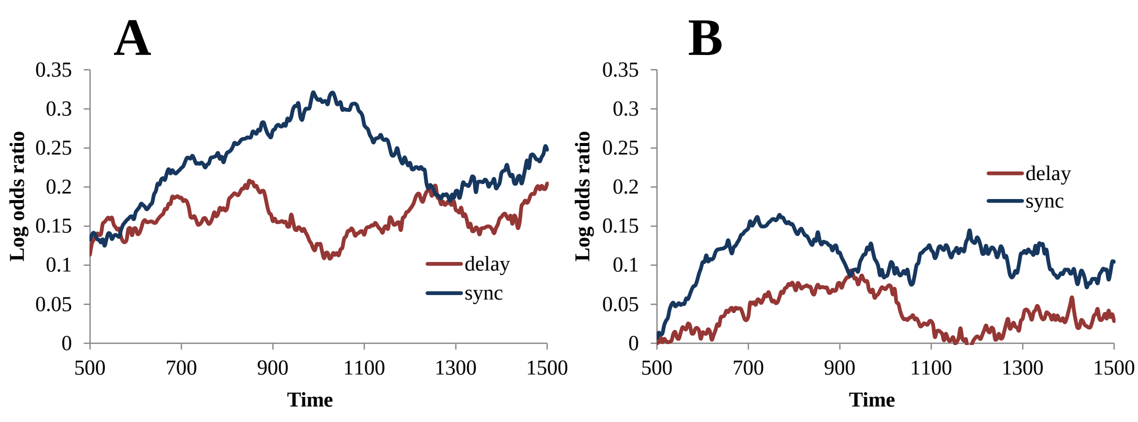

Figure 3-6: Experiment 1 – Log-odds-ratio of proportion of competitor looks to proportion of looks to average unrelated item across time, by training conditions. A) Onset competitors. B) Offset competitors.

---

2 Thanks to Toby Mordkoff for suggesting this analysis technique.

during this time window. We divided the looks to the competitor by the looks to the unrelated, and then took the log of this value to get the log odds ratio for the subject and stimulus. In a very few cases (3 cases out of 368), the average proportion of looks to either the competitor or the unrelated items over this time window was 0, which leads to an odds ratio of 0 or infinity, neither of which has a real-valued logarithm. Twice this occurred because of no looks to the unrelated objects, and once because of no looks to the competitor. Two of these cases were for *offset competitors* in the *delay* condition, while the third was an *onset competitor* in the *delay* condition. These cases were excluded from analysis (replacing these cases with the mean for that participant in that word-type yielded highly similar results). Log-odds-ratios greater than 0 indicate more looks to the

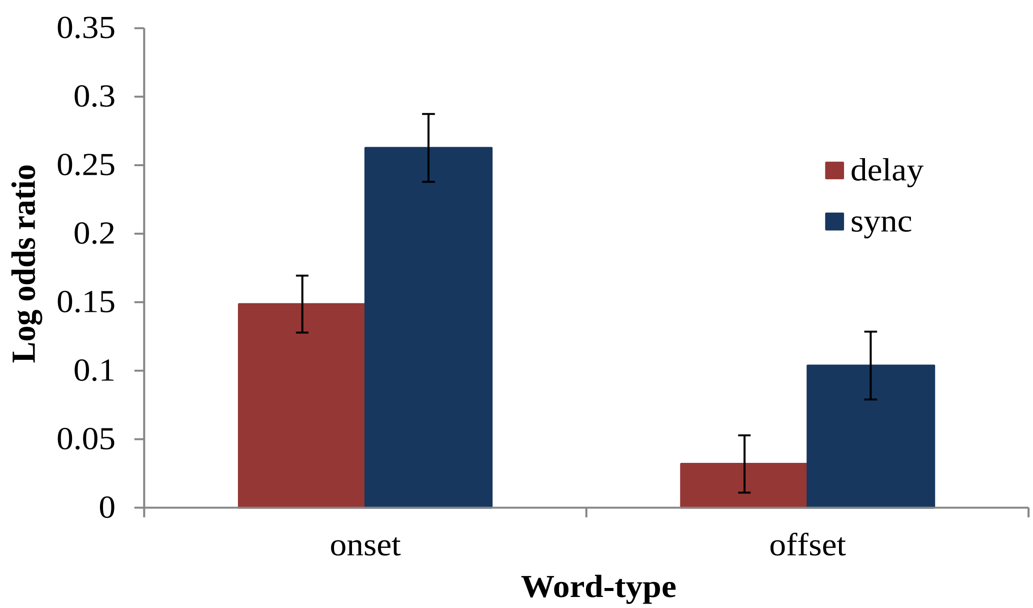

Figure 3-7: Experiment 1 – Log-odds-ratio within the 500-1500 ms analysis window by training condition and word-type. Error bars represent standard error for that condition.

competitor object than to the average unrelated object during the time window. We can then compare the log-odds-ratios between subject groups and between word types to determine whether the competitor effects differ between conditions.

Figure 3-7 shows the mean log-odds-ratio for both groups and word-types. Overall, the *synchronous* group had higher log-odds-ratios than the *delay* group, and the *onset competitors* yielded higher log-odds-ratios than the *offset competitors*. These data were analyzed using a mixed effects model with a linear linking function. As in other models, training condition and word-type were contrast coded and included as fixed effects, while participant and auditory stimulus were included as random intercepts. There was no correlation between the fixed effects (R=.002). MCMC simulations (20,000 iterations[3]) of the results were used to get significance values from the linear model.

This analysis revealed a significant main effect of training condition (B=-.091, SE=.046, p_{mcmc}=.049), as the *synchronous* group showed greater competitor effects than did the *delay* group. There was also a main effect of word-type (B=-.13, SE=.039, p_{mcmc} =.0006), with greater fixations to *onset competitors* (relative to unrelated items) than *offset competitors*. The interaction was not significant (B=.046, SE=.060, p_{mcmc} =.45). Participants in the *synchronous* group showed increased competition for both *onset* and *offset competitor* words. This suggests that during word learning, associations between words and objects occur continuously in time and do not wait for real-time processing to complete: those participants who could form associations during periods of competition showed increased consideration of competitors at test.

The repeated testing trials provide a way to test the magnitude of the competitor effect throughout testing. As both groups of participants complete testing trials with the same timing, the difference between groups may have changed during testing. These trials provide the opportunity to continue updating the representations formed during

[3] For all MCMC simulations used throughout this dissertation, 20,000 iterations were used

learning, as it serves as another form of cross-situational learning (although the statistics are constant across trials, as the foils were consistent across VWP trials). As both groups completed these trials identically, the learned representations are likely to converge somewhat during testing. The data for the two testing blocks are presented in Figure 3-8. Participants in the *synchronous* group appeared to show consistently larger competitor effects for *onset competitors* for both the first and second half of trials. The effect for *offset competitors* appeared to increase with additional testing trials, showing numerically larger effects in the second half of testing trials. We analyzed these data including trial block (first vs. second, contrast coded) as a factor. Due to the decreased number of trials

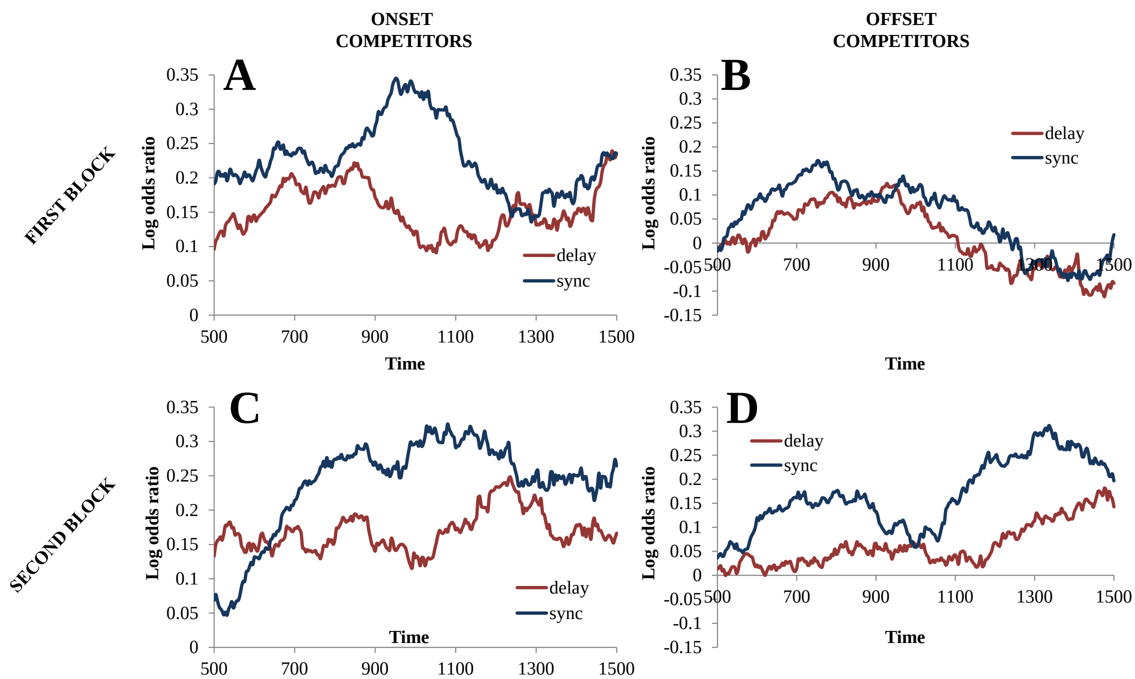

Figure 3-8: Experiment 1 – Log-odds-ratio of proportion of competitor looks to proportion of looks to average unrelated item across time by training condition. A) First block onset competitor trials. B) First block offset competitor trials. C) Second block onset competitor trials. D) Second block offset competitor trials.

per cell, a larger number of stimuli had to be excluded from this analysis (28 stimuli out of 736 total) because of zeros in either the numerator or denominator. Of these excluded trials, eight were in the *onset competitor* condition, while 20 were *offset competitors* trials. The majority of cases were for the *delay* condition (21 of the 28) and they were evenly split between trial blocks (14 in each). There were slightly more cases with no looks to the unrelated objects (10 cases) than with no looks to the competitor (15 cases); in the remaining three cases, there were no looks to both the unrelated objects and the competitor. The structure of the statistical model was identical to the initial model for this data with the addition of trial block as a factor. No contrasts were correlated (all $R<.003$).

This model showed the expected main effects of training condition (B=-.11, SE=.048, p_{mcmc}=.019) and word-type (B=-.15, SE=.033, p_{mcmc}<.00005), and no interaction between training condition and word-type (B=.022, SE=.050, p_{mcmc} =.66). The effect of trial block was not significant (B=.015, SE=.025, p_{mcmc}=.55), nor were any interactions with trial block (block × training condition: B=-.027, SE=.050, p_{mcmc} =.59; block × word-type B=.050, SE=.050, p_{mcmc}=.33; block × training condition × word-type: B=-.12, SE=.10, p_{mcmc}=.22). Throughout testing, the enhanced interference seen in the *synchronous* group was consistent for both *onset* and *offset competitors*.

Fixations to the target are also informative about how participants are processing the words. Whereas the enhanced competition in the *synchronous* group is predicted to be indicative of increased associations arising from learning, this effect could also arise from poorer overall learning; if participants are simply scanning the display more, they may look more to competitors and less to the correct targets. If this was the case, participants should display fewer looks to targets in the *synchronous* condition. However, as seen in Figure 3-9, quite the opposite pattern emerged: participants in the *synchronous* group showed more fixations to the target throughout the trial. This trend was consistent across the timecourse of the trial, and occurred for both *onset competitors* (Figure 3-9A) and *offset competitors* (Figure 3-9B).

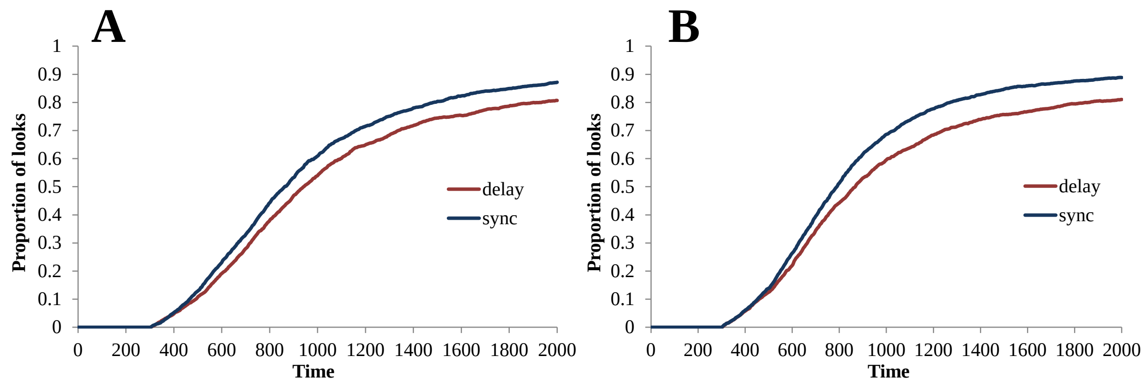

Figure 3-9: Experiment 1 – Proportion of looks to target items across time, by training condition. A) Onset competitors. B) Offset competitors.

To analyze the proportion of looks to the target items, the empirical logit transformation was applied to the data, as looks to the target often do not fit a Gaussian distribution. This transformation was used instead of a log-odds-ratio, as proportions are being compared between trials, so independence is maintained when comparing fixations between conditions. The average proportion of looks across the same analysis window used for competitor analyses (500-1500 ms) was entered into a mixed effects model with a linear linking function, with training condition and word-type (contrast coded) as fixed factors and participant and auditory stimulus as random intercepts. The fixed effects were not correlated (R<.0001). Significance was again determined using MCMC simulations. The main effect of training condition was significant (B=-.13, SE=.067, p_{mcmc}=.0026), with more fixations to targets by the *synchronous* group than by the *delay* group. This runs counter to predictions that competitor effects are indicative of poorer learning; the *synchronous* participants were in fact more adept at fixating the target. The effect of word-type was also significant (B=.094, SE=.017, p_{mcmc}=.0018), with more target fixations for *offset competitor* trials than for *onset competitor* trials. This result is expected, as looks to competitors were greater for *onset competitor* trials; these looks likely reduced the looks to the target in these trials. The interaction between training condition and word-type was not significant (*B*=-.037, *SE*=.033, p_{mcmc}=.32).

3.4 Discussion

Experiment 1 established that changes in timing during learning lead to increased competitor effects for *synchronous* participants. When identifying a target word, learners in the *synchronous* condition fixated the referents of phonological competitors more than those in the *delay* condition, despite both conditions having identical online processing demands. This effect persisted throughout testing, suggesting that the effect is relatively strong; although both groups received the same form of trials at test, the increased competition for the *synchronous* group did not deteriorate. There was no evidence of overall poorer learning in the synchronous group; interim trials gauging learning throughout training showed equal speed of word-referent acquisition, and at test the *synchronous* group was more likely to fixate the target.

This suggests that the increased competitor fixations in the *synchronous* group arise from differences in associations with competitors, rather than from poorer learning of the correct word-referent mappings. The likely source for such increased competition is learned associations between referents and phonological competitors that formed during periods of lexical competition. Multiple word-forms are active in parallel as they compete for access during the learning trials. For the participants in the *synchronous* condition, the visual referents are present as these competitive processes are occurring, allowing mapping of parallel-activated word-forms. For the *delay* group, the referents are not present until the end of the auditory stimulus, providing ample time for competition to resolve (especially for the *onset competitors*). In this condition, participants don't begin learning until many of the competing word-forms activated in parallel have been suppressed.

These associations between referents and phonological competitors are particularly remarkable as co-occurrence statistics between competitors and referents suggest that these are the least likely associations that participants should form. When hearing a given word during training, one of the two objects displayed was always the correct referent. The alternative object was never the referent of the phonological competitor, giving a co-occurrence rate of 0 linking this word-object pair. The alternative object was instead drawn randomly from the other six available referents, giving a co-occurrence rate of approximately 1/6 for all the other word-referent pairings. During the test trials, participants thus had good reason to have formed associations between what will be the *unrelated* items and the target word, but no reason based on actual co-occurrence to have associations with the competitor items. These associations instead must have formed from parallel activation during word recognition.

More generally, this experiment establishes a viable method for investigating interactions between perceptual processing and learning. By manipulating the timing of the auditory and visual presentation, this experiment augmented the mappings that were formed during learning. This demonstrates that the paradigm is sensitive to changes in learned representations, making it a viable basis for further studies on how learning is affected by processing dynamics.

Although this experiment is indicative of changes in learning as a result of manipulating the timing of word-referent pairings, it leaves many additional questions to answer. The timing manipulation between participants in this experiment not only differed in terms of the relative timing between the visual and auditory stimuli, but also in terms of overall trial duration. This leaves two possible provenances for the weaker competitor effects in the *delay* condition: they could emerge because no learning occurs during lexical activation, or they could emerge because additional learning occurs after competition is resolved. That is, even if learning initiates during online lexical activation processes, continued updating of this learning after competition has resolved might quash the associations learned in parallel. In order to develop a more comprehensive understanding of how timing interacts with word learning, additional timing manipulations are necessary.

CHAPTER 4

ALTERNATIVE TIMING MANIPULATIONS

Experiment 1 provided evidence that manipulating the timing of the *visual* stimuli affects when during *auditory* processing learners can form word-referent mappings and this in turn alters the form that the mappings between word-forms and objects take. Specifically, *synchronous* auditory-visual presentation leads learners to show increased competition after learning the words, despite the fact that in both conditions the word-object linkages are learned quite well. This suggests that when learning occurs during periods of lexical competition, partially active competing word-forms are also mapped to the referent. However, these results do not rule out the possibility that with additional processing time after the offset of the referent, learners can continue updating their learned mappings to overcome these earlier spurious associations; in Experiment 1, the *synchronous* group continued to the next trial immediately upon offset of the visual information, giving them no chance to update their learning after competition resolves.

The experiments in Chapter 4 provide a more detailed analysis of the learning process during word learning trials. Specifically, Experiment 2 investigates whether additional processing time at the offset of the visual stimulus for the *synchronous* group eliminates false associations. If learners are indeed learning continuously, then additional processing time may help learners overcome associations formed during periods of ambiguity. If this additional processing time reduces spurious associations, this suggests that learning can operate on the basis of memory representations of stimuli in addition to raw co-occurrence; as objects are held in working memory, these representations can be used to update the learned mappings between stimuli. This form of learning is more complex than standard approaches to associative learning; rather than associative learning operating solely on the basis of perceptual representations (i.e. behaviorism), this would provide evidence of learning operating on the basis of internal representations. Such results would demonstrate the capacity for associative learning to develop and operate on abstract representations.

Experiment 3 provides visual information before auditory information, such that trials are equal in duration to those of the *delay* condition in Experiment 1, but participants still learn during periods of lexical competition. This design provides participants with the same trial duration as those in the *delay* condition, but still allows spurious associations to form during periods of lexical competition (without additional processing time after auditory stimuli). Equating trial duration controls for time pressure on participants that may affect how they learn; *synchronous* presentation as in Experiment 1 has rapid transitions between trials, which may lead participants to change their threshold for when to learn. Although Experiment 3 was designed to control overall trial duration in the *synchronous* condition without changing the learning, this experiment led to quite interesting findings regarding the role of prediction in word learning. Providing participants with visual referents before the onset of auditory stimuli may bias them to consider certain word-forms before even hearing the auditory stimulus.

4.1 Experiment 2: Synchronous training with additional processing time

Experiment 2 builds on the results of Experiment 1 by investigating how additional processing time after auditory offset affects learning that occurs during lexical competition. Experiment 1 showed that training with *synchronous* timing between auditory and visual stimuli resulted in increased competition. However, additional learning time after competition resolves may lessen this effect, as learners can continue to update their representations of the stimuli, and thereby continue altering the learned mappings. As competition resolves, the learning operates to strengthen connections between the word-form that won competition (likely the correct word-form) and the referent, while also weakening the spurious connections formed earlier in lexical processing. Although both the visual and auditory stimuli are no longer being presented, ongoing processing of the auditory and visual inputs in working memory may be used to continue updating the word-referent mappings. If additional time is available at the end of trials, such updating may alleviate some of the increased competition seen for the *synchronous* training condition.

To this end, Experiment 2 mirrors the *synchronous* group of Experiment 1, but adds a pause with a blank visual display between trials to equate total trial duration to that of the *delay* group. This allows participants the same amount of time post-auditory-stimulus to learn as in the *delay* group, but also allows the formation of false associations with parallel-activated word-forms early in the learning trial. The results of Experiment 2 will then be compared to both the *synchronous* and *delay* conditions of Experiment 1; these comparisons allow analysis of whether the additional processing yields continued evidence of spurious association (i.e. if Experiment 2 looks like the *synchronous* condition), or if such associations are eliminated (i.e. if Experiment 2 looks like the *delay* condition).

4.1.1 Methods

4.1.1.1 Participants

Thirty participants from the University of Iowa community completed the study. Participants were paid $15 or received partial course credit for their participation. All participants self-reported normal hearing and normal or corrected-to-normal vision. Data from 22 of the participants were included in analyses; an additional seven participants completed the study but were excluded for low accuracy at test, and one participant was excluded as a result of poor eye-tracker calibration.

4.1.1.2 Design

The design was similar to that of Experiment 1. Participants first performed a phoneme-monitoring task to become familiar with the words, and then learned the words in a cross-situational learning task. All participants received the same timing of word-referent pairings, such that visual stimuli and auditory stimuli began at the same time, and a blank screen was displayed for 1,550 ms at the end of the auditory stimuli, giving additional processing time. These results are then treated as a third experimental condition in comparisons with the two training conditions of Experiment 1. The same form of interim test trials and the same VWP test as in Experiment 1 were administered.

4.1.1.3 Stimuli

The same stimuli used in Experiment 1 were also used in Experiment 2.

4.1.1.4 Procedure

Most aspects of the procedure were identical to Experiment 1. The pre-exposure, interim test trials and VWP test trials were identical between the two experiments. However, the timing manipulation in the word-referent training trials was slightly different. Auditory and visual stimuli onset simultaneously after the trial was initiated by the participant, as in the *synchronous* condition of Experiment 1. Also like the *synchronous* condition of Experiment 1, the visual stimuli remained on the screen for 800 ms, after which the screen went blank. Unlike in Experiment 1, this blank screen remained up for 1,550 ms (Figure 4-1C). This 1,550 ms includes a 550 ms ITI as used in Experiment 1, although in this experiment that ITI was indistinguishable from the 1,000 ms of blank screen presented after the trial. This delay ensured that the trial duration for this experiment matched the trial duration of the *delay* condition in the previous experiment (which included 1000 ms before visual stimulus display, 800 ms of stimulus display, and 550 ms of inter-trial interval). After the blank screen, the next trial commenced. Each word was a target on 32 trials, for a total of 256 total training trials.

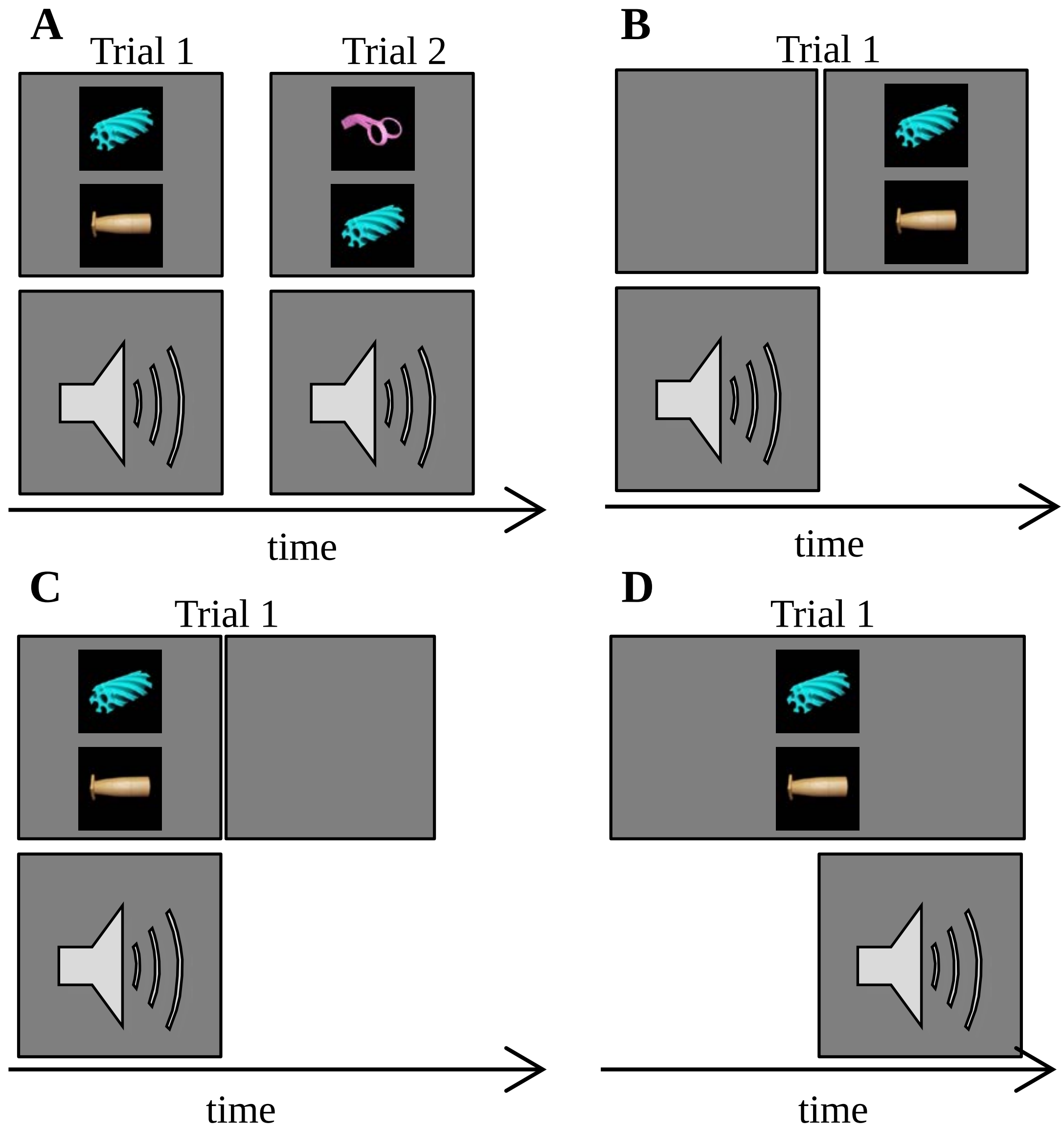

Figure 4-1: Schematic of the timing conditions comparing both conditions of Experiment 1 to the timing in Experiments 2 and 3 (not to scale). A) Synchronous condition of Experiment 1. B) Delay condition of Experiment 1. C) Experiment 2. D) Experiment 3.

4.1.2 Results

4.1.2.1 Pre-exposure trials

Participants overall were quite good at the phoneme-monitoring task during the pre-exposure phase of the experiment. They correctly identified 92% of tokens with “O” sounds, and had false alarm rates of approximately 12%. These values are quite comparable to those of Experiment 1 (93% correct positive rate, 10% false alarm rate).

4.1.2.2 Interim testing trials

The interim test trials of Experiment 2 exhibited performance that was fairly similar to that of Experiment 1 (Figure 4-2), with slightly faster learning for the *onset competitor* items. These learning data were compared to both conditions of Experiment 1 in a mixed effects model with a binomial linking function. For this model, two separate contrast codes were included to compare Experiment 2 independently to the *synchronous*

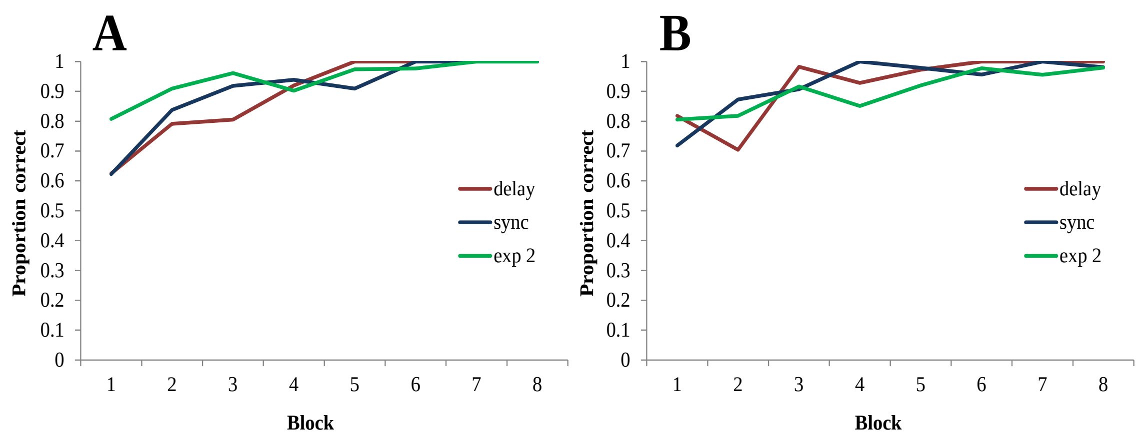

Figure 4-2: Experiment 2 – Accuracy of responses during yes/no interim trials for Experiment 2 compared against both conditions from Experiment 1. A) Onset competitor items. B) Offset competitor items.

and *delay* conditions of Experiment 1 (to compare *synchronous* to Experiment 2 (E2vsSync): *synchronous*=-.5, *delay*=0, Experiment 2=+.5; to compare *delay* to Experiment 2 (E2vsDel): *synchronous*=0, *delay*=-.5, Experiment 2=+.5). The model included each of these contrast codes as factors, along with word-type (contrast coded) and block (eight blocks, coded as true block number). The contrast codes to contrast Experiment 2 to the two conditions of Experiment 1 were entered into the analysis as separate interactions with word-type and block. Participant and auditory word were included as random intercepts[1]. The fixed factors were not correlated (all $R<.08$).

Table 4-1 presents the results of this analysis. The main effect of block shows that participants perform better on later blocks, and the marginally-significant effect of word-

Table 4-1: Results of the statistical analysis of interim test trials comparing Experiment 1 to Experiment 2.

Factor	*B*	*SE*	*Z*	*p*
Block	.69	.060	11.54	<.0001
Word-type	-.60	.36	-1.69	.091
E2vsSync	.39	.63	.62	.53
E2vsDel	1.14	.63	1.79	.074
Block × type	.18	.12	1.53	.13
Block × E2vsSync	-.080	.17	-.48	.63
Block × E2vsDel	-.36	.18	-2.00	.046
Type × E2vsSync	.56	.96	.58	.56
Type × E2vsDel	.33	.99	.33	.74
Block × type × E2vsSync	-.013	.33	-.039	.97
Block × type × E2vsDel	.18	.36	.49	.63

[1] In all analyses in Experiments 2 and 3, no random slopes were included. This was done to maintain the same model structure as that used in Experiment 1, as these experiments are compared directly to Experiment 1.

type emerged due to slightly better performance overall for *onset competitors*. The comparison of Experiment 2 to the *delay* condition of Experiment 1 was marginally significant, and was moderated by an interaction with block; this emerged because Experiment 2 participants reached peak performance slightly faster than the *delay* group. No other effects or interactions were significant. This set of analyses shows that the Experiment 2 participants learned the words as quickly as the *synchronous* participants of Experiment 1, and slightly faster than the *delay* participants.

4.1.2.3 VWP testing trials

4.1.2.3.1 Accuracy

As in Experiment 1, participants whose accuracy was below 75% on either *onset* or *offset competitor* trials were excluded. This comprised seven participants. Many of these participants had accuracy levels near chance, indicating that they did not learn at all during training. The remaining 22 participants were near ceiling, with an average accuracy near 99%. As seen in Figure 4-3, these accuracy levels were quite comparable to those in Experiment 1. These data were entered into a mixed effects model with a binomial linking function. This model used the same contrast codes of the previous analysis to compare Experiment 2 separately to the *synchronous* and *delay* conditions of Experiment 1 (as in the previous analyses, comparisons are coded as E2vsSync for the comparison with the *synchronous* group and E2vsDel for the comparison with *delay*). Word-type (contrast coded) was included as well as its interaction with these contrast codes. Training condition and word-type were fixed factors while participant and auditory stimulus as random intercepts. The fixed effects were not correlated (all R<.07).

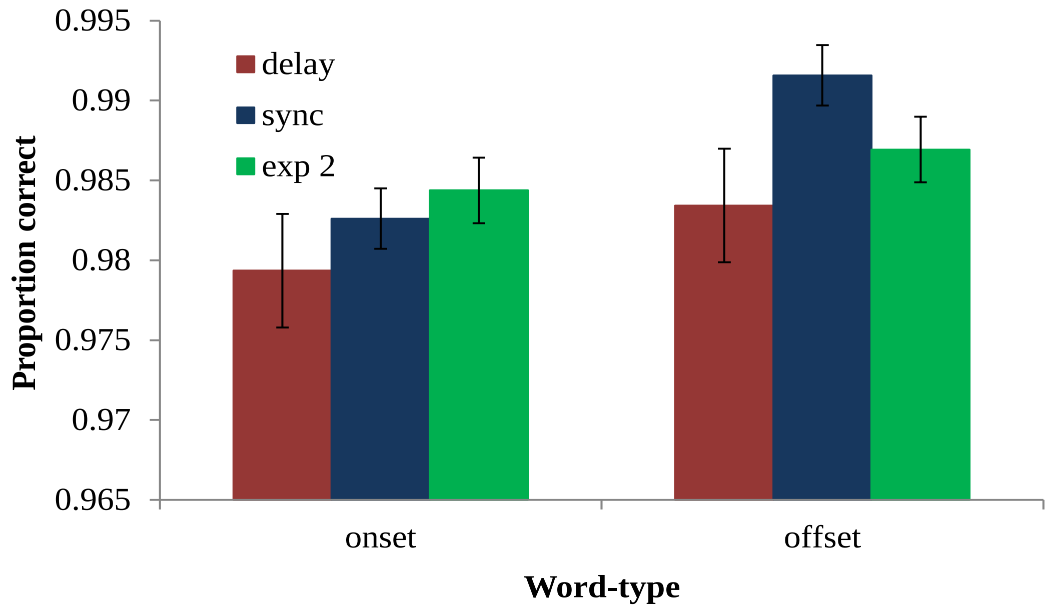

Figure 4-3: Experiment 2 – Accuracy for included participants in the VWP trials of Experiment 2 compared to the conditions of Experiment 1, by word type. Error bars represent standard error for that condition.

This model revealed a significant main effect of word-type (*B*=.39, *SE*=.12, *Z*=3.31, *p*=.0009), as participants were more accurate for *offset competitors* than for *onset competitors*. Neither effect comparing Experiment 2 to the timing conditions of Experiment 1 was significant (E2vsSync: *B*=-.45, *SE*=.40, *Z*=-1.13, *p*=.26; E2vsDel: *B*=.12, *SE*=.39, *Z*=.30, *p*=.76). The interaction of word-type and E2vsDel was not significant (*B*=.31, *SE*=.31, *Z*=.98, *p*=.33). However, E2vsSync interacted significantly with word-type (*B*=-.72, *SE*=.35, *Z*=-2.05, *p*=.040). Simple effects analyses comparing these two conditions independently within the two word-types were conducted using models with the same structure as the omnibus model, but including only E2vsSync as a factor. These models showed no difference between groups for *onset competitors* (*B*=.063, *SE*=.35, *Z*=.18, *p*=.86), however the difference for *offset competitors* was marginally significant (*B*=-.97, *SE*=.55, *Z*=-1.76, *p*=.079); the *synchronous* participants in

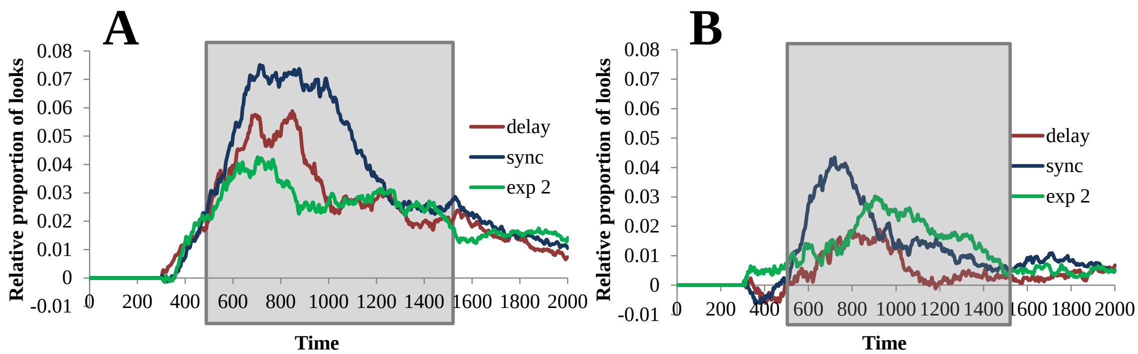

Figure 4-4: Experiment 2 – Relative proportion of looks to competitor objects (competitor – average unrelated). Analysis window highlighted in gray. A) Onset competitors. B) Offset competitors.

Experiment 1 were slightly more accurate for *offset competitors* than were the Experiment 2 participants (although both were highly accurate).

4.1.2.3.2 Eye movements

As in Experiment 1, the analysis of eye-tracking data only included trials in which participants selected the correct target item (which comprised the bulk of the trials as accuracy exceeded 98%). Analysis was conducted as in Experiment 1, and results are compared between the two experiments to determine whether participants trained with the timing condition of Experiment 2 show competition patterns more akin to the *synchronous* or the *delay* condition of Experiment 1 (or are distinct from both). Competitor looks were thus considered relative to unrelated looks similarly to Experiment 1. Figure 4-4 shows the relative proportion computed by subtracting the looks to the average unrelated item from the competitor item and Figure 4-5 shows the same results presented as log-odds-ratios over time. The two comparisons are quite similar. For *onset competitor* items, participants in Experiment 2 showed approximately the same amount of competition as those in the *delay* condition of Experiment 1. For *offset competitor* trials, Experiment 2 participants showed a more nuanced pattern with respect to Experiment 1: early in the trial, these participants showed only weak

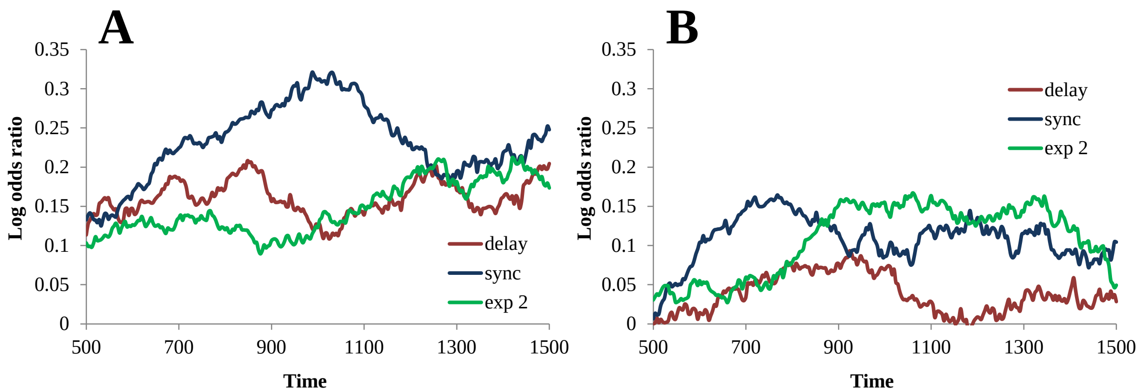

Figure 4-5: Experiment 2 – Log-odds-ratio of proportion of competitor looks to proportion of looks to average unrelated item across time, by training condition. A) Onset competitor trials. B) Offset competitor trials.

interference from the competitor, but late in the trial, they showed even more interference than the *synchronous* participants from Experiment 1.

These data were analyzed using log-odds-ratios (Figure 4-5). As seen in the figure, transforming the data this way showed similar patterns to those seen with the subtraction method (Figure 4-4): no evidence of increased interference for *onset competitor* trials, and evidence for late interference for *offset competitor* trials. To analyze this data, two separate mixed effects models were used. These models allowed independent comparison of Experiment 2 to each of the conditions of Experiment 1 without including data from the other condition to affect the error term. Including both factors simultaneously in a single model produced quite similar results. Both models used linear linking functions (the log-odds-ratio transformation makes this appropriate for this DV). The DV for each model was the log-odds-ratio of the mean proportion of looks to the competitor and the mean of the two unrelated items across the analysis window (500-1500 ms; Figure 4-6). One cell was excluded from the Experiment 2 data because there were no looks to the unrelated objects; this cell was an *offset competitor*. Fixed factors included a contrast code for the given comparison between training conditions (i.e. -.5 for

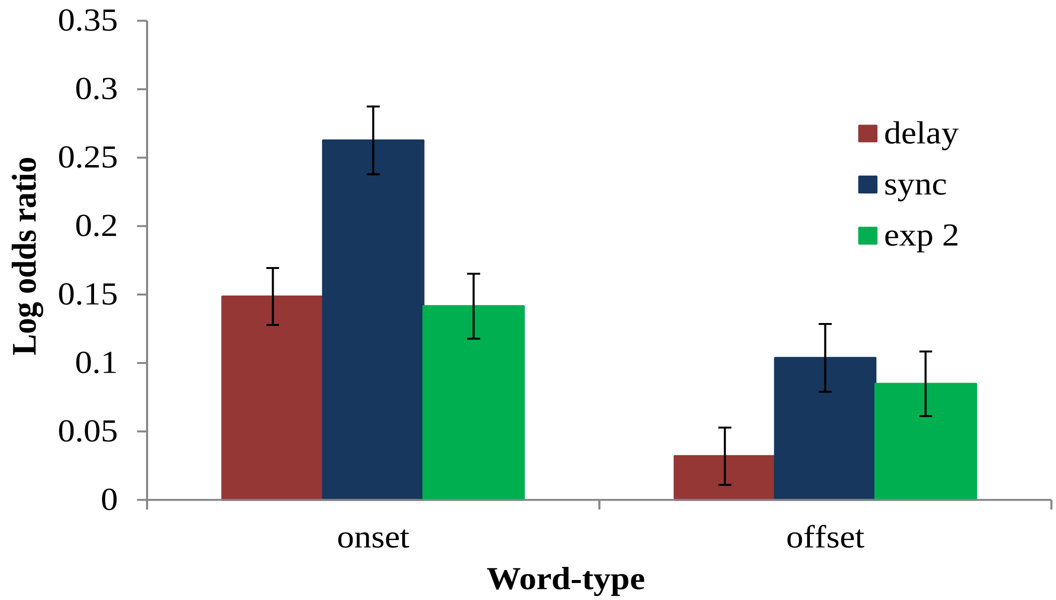

Figure 4-6: Experiment 2 – Log-odds-ratio within the 500-1500 ms analysis window, by training condition and word-type. Error bars represent standard error for that condition.

either *synchronous* or *delay*, +.5 for Experiment 2) and trial type. The fixed factors were not correlated (all R<.003). This model included training condition and word-type as fixed factors, and participant and auditory stimulus as random intercepts. Significance values were estimated through MCMC stimulations.

The analysis comparing Experiment 2 to the *delay* condition of Experiment 1 did not show an effect of training condition (B=.023, SE=.029, p_{mcmc}=.51). There was a marginally-significant main effect of word-type (B=-.087, SE=.038, p_{mcmc}=.059), with greater competition for *onset competitors* than for *offset competitors*. There was no interaction (B=.060, SE=.059, p_{mcmc}=.31).

The analysis comparing Experiment 2 to the *synchronous* condition of Experiment 1 showed a marginally-significant effect of training condition (B=-.069, SE=.044, p_{mcmc}=.083), with overall greater competition in the *synchronous* group, and a significant main effect of word-type (B=-.11, SE=.037, p_{mcmc}=.026), with greater overall competition effects for *onset competitor* words. There was also a marginally-significant interaction between word-type and training condition (*B*=.10, *SE*=.061, p_{mcmc}=.095). To investigate this interaction, simple effects models were used to compare the training conditions independently for each word-type. These models maintained the same structure as the primary model (though without the effect of word-type). These sub-analyses showed a significant effect of training condition for *onset competitors* (*B*=-.12, *SE*=.052, p_{mcmc}=.030), with greater competition effects for the *synchronous* participants in Experiment 1. There was no difference between the training conditions for *offset competitors* (*B*=-.017, *SE*=.047, p_{mcmc}=.66).

Thus, these analyses suggest that synchronous exposure with additional processing time did not elicit increased competitor effects for *onset competitors* (that is, results appeared closer to the delay condition in Experiment 1); however, *offset competitors* looked more like the *synchronous* condition of Experiment 1, showing greater competition effects given the training of Experiment 2 relative to the *delay* condition of Experiment 1.

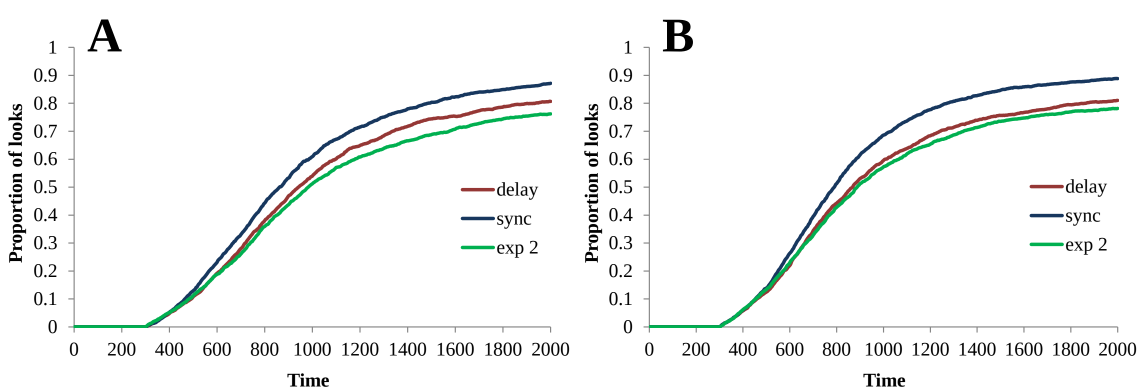

Figure 4-7: Experiment 2 – Proportion of looks to target items across time, by training condition. A) Onset competitor trials. B) Offset competitor trials.

Target looks were analyzed similarly, using parallel models comparing Experiment 2 to the two training conditions of Experiment 1 (Figure 4-7). These models included the same structure as in the analysis of competitor looks, but used empirical-logit-transformed proportions of looks to the target as the DV, as in the analysis of Experiment 1. The fixed factors were not correlated in either model (all R<.01). The comparison with the *synchronous* condition showed a significant effect of training condition was (B=-.17, SE=.071, p_{mcmc}=.0003), with more looks to the target for the *synchronous* condition than for Experiment 2. There was also a significant effect of word-type (B=.11, SE=.019, p_{mcmc}=.0017), with more looks to the target for *offset competitor* trials. The interaction was not significant (B=-.0063, SE=.034, p_{mcmc}=.87). The comparison with the *delay* condition also showed no effect of training condition (B=-.086, SE=.14, p_{mcmc}=.32). There was a main effect of word-type (B=.076, SE=.024, p_{mcmc}=.014). However, there was no interaction between word-type and training condition (B=.062, SE=.066, p_{mcmc}=.41). Looks to the target item in Experiment 2 were thus quite comparable to the *delay* condition of Experiment 1 for both word-types (despite the differences in competitor looks), with decreased looks to the target relative to the *synchronous* condition of Experiment 1.

4.1.3 Discussion

The results of Experiment 2 show that with additional processing time after the visual and auditory stimuli are no longer present, participants who learned words even with *synchronous* timing show a different pattern of spurious associations with phonological competitors from what was observed in Experiment 1. These participants showed smaller interference effects for *onset competitors* than did the participants in the *synchronous* condition of Experiment 1, whose timing was identical except for the additional blank screen at the end of the trial. However, this reduction in interference was not apparent for *offset competitors*, suggesting that the additional processing time did not eliminate the spurious associations learned during co-activation late in words.

The effects of Experiment 1 were interpreted as evidence that learners initiate learning as soon as auditory and visual information are both available. In the *synchronous* condition of Experiment 1, visual referents were present both during the beginnings of words, when *onset competitors* are highly co-active, as well as at the end of the words, when *offset competitors* have become active. Thus learners in this timing condition had the opportunity to form associations with both forms of competitors. Meanwhile, those in the *delay* condition did not see the visual referents until after word offset, providing no opportunity to build associations with parallel active competitors. The results of Experiment 2 complicate this interpretation somewhat. In this experiment, participants have the opportunity to form associations with both *onset* and *offset competitors*. However, evidence of these associations is only seen for *offset competitors*, suggesting that the spurious associations for *onset competitors* are ameliorated with additional processing time.

Although the results of Experiment 1 show evidence of learning during parallel lexical activation, they do not signal that learning stops while activation is ongoing. Instead, a true form of temporally-continuous associative learning suggests that learning initiates immediately and continues to update throughout the learning event, as more information arrives. In the *synchronous* condition of Experiment 1, the trial ended quite shortly after word onset, providing little opportunity to continue updating; this forces learners to rely on the associations formed during the auditory presentation. Experiment 2 provides additional time to continue updating the mappings learned during the auditory stimuli. Learners can thus adjust the learned associations to better represent the complete word-form if they continue to update learned mappings on the basis of working memory representations of the stimuli.

Why does additional processing time affect linkages formed with *onset competitors* but not *offset competitors*? One possibility is that the late overlap for *offset competitors* leads to lasting co-activation after the word. This may require even more processing time to suppress the competitors completely and eliminate learned spurious associations. Also, as no further auditory information is forthcoming to disconfirm these coactive *offset competitors,* their activation may persist for some time. Learners can thus start to form associations with *offset competitors* late in auditory presentation, and then continue to update these representations during the beginning of the blank screen period in Experiment 2. Participants may need additional time to elapse after word offset to completely suppress these associations. In contrast, *onset competitors* are activated quite early, allowing suppression well before the blank screen period in Experiment 2. Additionally, later auditory information in the word disconfirms these competitors, perhaps providing an easier path to suppression of the competitors (and thus more time update mappings after the competitors are suppressed).

The *delay* condition of Experiment 1 also provides a chance for participants to learn during the post-offset parallel activation of *offset competitors*; however, this learning only catches the tail of competitor activation, and so is not as strong as that seen in Experiment 2. Future studies including longer periods of delay after auditory offset can determine whether even more processing time eliminates associations to *offset competitors*.

The delay after word offset in Experiment 2 did not include additional visual presentation, yet it led to continued learning about the word-referent pairing. A basic associative explanation of word learning suggests that learning occurs only when both word-form and referent stimuli are available. However, in this case, working memory representations of the visual objects may allow continued learning during this period. As the learning trials included only two objects per display, learners could easily hold these referents in memory during the interim period. Indeed, in the *delay* condition of

Experiment 1, learning must occur on the basis of some working memory representation of the auditory stimulus, as the word has completed before visual presentation onset. This interpretation is only possible if learners have the capacity to store the stimuli in working memory. If a larger number of visual stimuli were presented during training, learners may have to rely more on active perceptual representations rather than maintained working memory representations. This would be predicted to lead to the maintenance of spurious associations despite additional processing time.

The participants in this experiment also showed fewer looks to the target item than participants in the *synchronous* condition of Experiment 1, looking instead quite like the *delay* participants. The increased processing time afforded by this experiment thus led participants to form associations between the correct target and the referent to the same extent as participants who only learn after word offset. Despite the persistence of increased competitor effects for *offset competitors*, these participants still identify the target much the same as the *delay* group. It is unclear what leads these target fixations to decrease; perhaps the longer training trial durations led to overall decreased looking for participants in these two conditions.

Experiment 2 thus reinforces the findings from Experiment 1 that learners form word-referent mappings during periods of lexical competition. However, the learning during these periods is malleable, as later-occurring information provides the chance to update these representations to reflect further processing. This updating can help learners avoid spurious word-referent associations during learning, as learning situations likely rarely have the same time pressure applied in laboratory experiments. Instead, learners often have a chance to continue updating learned representations after competition resolves.

Differences between Experiment 2 and the *synchronous* condition of Experiment 1 seem to arise from additional processing that occurs after word offset, which was denied the *synchronous* participants as the next trial began immediately. However, the trials in Experiment 2 also had overall longer duration than those of the *synchronous* condition of Experiment 1; this presents a potential confound that could explain the results. Whenever trials are longer in duration, competitor effects appear weaker (although the significant difference between Experiment 1's *delay* condition and Experiment 2 cannot be explained by such a confound). As such, finding a way to control processing time while also equating trial duration offers a more thorough analysis of these effects.

4.2 Experiment 3: Visual presentation preceding auditory presentation

This experiment aimed to control for overall trial duration while manipulating the timing of stimulus presentation. Whereas Experiment 2 maintained the stimulus timing of the *synchronous* condition of Experiment 1, it also increased the time each trial took. Participants may change their word-referent learning strategy when placed under more stringent trial timing; this would suggest that some of the increased interference effects seen in Experiment 1 arise from strategic changes in encoding rather than from temporally-continuous learning emerging naturally given auditory-visual synchrony.

Experiment 3 aimed to equate timing between a synchronous and a delay presentation without adding additional processing time after word offset for the *synchronous* group. To this end, visual stimuli were presented before the onset of the auditory stimuli. This provided additional time within a trial without allowing word-referent learning. This contrasts with the presentation format of Experiment 2, which lengthened trials by adding blank space at the end of the trial; that blank delay provided additional time to update the word-referent mappings.

4.2.1 Methods

4.2.1.1 Participants

Twenty-six participants from the University of Iowa community completed the study. Participants were paid $15 or received partial course credit for their participation. All participants self-reported normal hearing and normal or corrected-to-normal vision. Data from 20 participants were included in analyses; an additional six participants completed the study but were excluded for low accuracy at test.

4.2.1.2 Design

The design was similar to that of Experiments 1 and 2. Participants performed a phoneme-monitoring task to become familiar with the words, and then learned the words in a cross-situational learning task. All participants received the same timing of word-referent pairings, with visual stimuli beginning before the onset of auditory stimuli, and remaining on the screen throughout auditory presentation. This timing condition was compared to both the *synchronous* and *delay* conditions of Experiment 1 to determine whether auditory-visual synchrony produces spurious associations even when trial duration is equated; similarity to the *synchronous* group suggests the formation of spurious associations, whereas similarity to the *delay* group suggests that such associations do not form. After auditory presentation, the trial ended and the next one began shortly, providing no opportunity for additional processing time. The same form of interim test trials and the same VWP test as in Experiment 1 were administered.

4.2.1.3 Stimuli

The stimuli used in this experiment were the same as those used in Experiments 1 and 2.

4.2.1.4 Procedure

The procedure closely mirrored that of the previous experiments. The pre-exposure, interim test trials and VWP testing trials were identical in every respect to Experiments 1 and 2. However, the cross-situational word-referent learning trials differed from previous experiments. On each of these trials, 100 ms after the participant clicked the central dot to begin the trial, the visual stimuli appeared. After 1000 ms, the auditory stimulus was played, with the visual stimuli still on the screen. The visual stimuli remained on the screen for 800 ms after the onset of the auditory stimulus, and then a blank screen was presented for 550 ms before the next trial. These trials (schematized in Figure 4-1D) were identical in duration to the *delay* condition of Experiment 1 and the trials of Experiment 2.There were an equal number of training trials as in the previous experiments (32 repetitions of the eight words, for 256 total training trials).

4.2.2 Results

4.2.2.1 Pre-exposure

Participants in this experiment were surprisingly poor at the pre-exposure portion of the experiment. Correct identification of words with "O" sounds was 89.5%, however the false alarm rate when no "O" sound was present was around 20%. These error rates are a good deal higher than those in Experiment 1 (93% correct positive rate, 10% false alarm rate) using the same stimuli, suggesting that participants in this experiment may not have been focused on the task. To assess this, the error rates were compared between experiments using a mixed effects model with experiment (Experiment 1[2] vs. Experiment 3, contrast coded: -.5/+.5) and word-type (contrast coded) as fixed factors, and participant

[2] Performance was compared between Experiments 1 and 3, not including Experiment 2, as the Experiment 3 data are never directly compared to those of Experiment 2. Additionally, because the pre-exposure phase precedes all between-participants manipulations, the participants from Experiment 1 are all included as a single group.

and auditory word as random intercepts. This model used a binomial linking function. The fixed effects were not correlated (R=.005). This model confirmed that participants in Experiment 3 were less accurate on the pre-exposure trials than were participants in Experiment 1 (B=-1.09, SE=.15, Z=-7.17, p<.0001). The model also showed a marginal effect of word-type (B=1.00, SE=.57, Z=1.75, p=.080), with better performance for *offset competitor* words. The interaction was not significant (B=.29, SE=.24, Z=1.21, p=.23).

Inspection of individual error rates showed that the inflated overall error rates arose from a handful of participants with extremely high error rates (five participants had overall error rates greater than 25%). The remaining participants showed error rates in line with those of the previous experiments (95% correct identification rate, 12% false alarm rate). Because the participants with high error rates learned the words to a high enough criterion to merit inclusion (greater than 75% accuracy for both *onset competitors* and *offset competitors* at test), they were included in further analyses[3]. However, it is possible that their poor performance during the pre-exposure phase may have led to poor encoding of the auditory word-form (and therefore poor parallel activation of the word-forms during the early training trials).

4.2.2.2 Interim testing trials

Participants in Experiment 3 showed a similar trajectory of learning during the interim test trials as participants in both groups of Experiment 1 (Figure 4-8). These data were analyzed as in the comparison of Experiment 2 to Experiment 1: two separate contrast codes were used to compare Experiment 3 to the separate conditions in Experiment 1. These are referred to throughout all analyses of Experiment 3 as E3vsSync for the comparison with *synchronous* and E3vsDel for the comparison with *delay*. These were entered into a single model with word-type (contrast coded) and block (eight blocks,

[3] These participants also showed quite similar learning profiles in the interim testing trials.

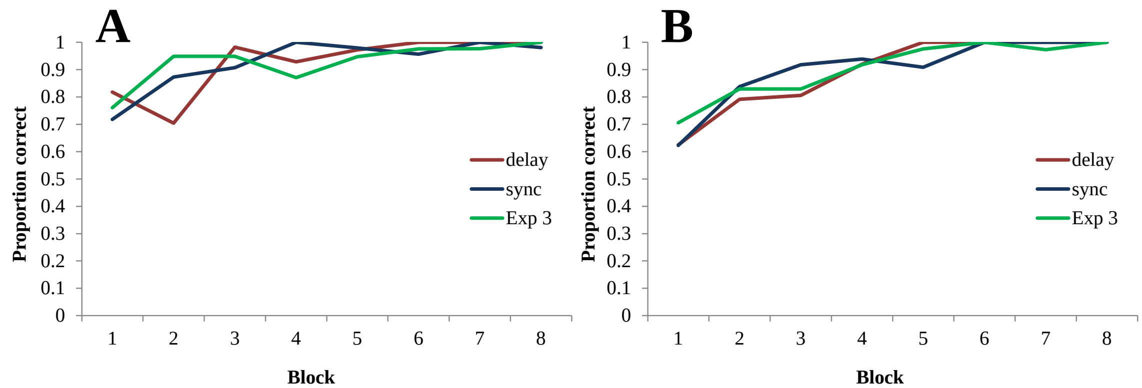

Figure 4-8: Experiment 3 – Accuracy of responses during yes/no interim trials for Experiment 3 compared against both conditions from Experiment 1, by block. A) Onset-competitor items. B) Offset-competitor items.

entered as raw block number) as fixed effects along with each of their interactions with the contrast codes comparing across experiments. Participant and auditory word were included as random intercepts. The model used a binomial linking function, and the fixed effects were not correlated (all R<.08).

This analysis revealed a main effect of block (B=.73, SE=.060, Z=12.06, p<.0001), with higher accuracy in later blocks. There was also a main effect of word-type (B=-.88, SE=.39, Z=-2.2, p=.028), with overall higher accuracy for *onset competitors*. However, neither effect comparing learning in Experiment 3 to the conditions in Experiment 1 was significant, nor were any of the interactions (Table 4-2); participants in Experiment 3 thus showed very similar learning profiles to those in Experiment 1.

4.2.2.3 VWP testing trials

4.2.2.3.1 Accuracy

Accuracy in the VWP testing trials was compared between Experiment 3 and both training conditions of Experiment 1 (Figure 4-9). Overall, the included participants were quite accurate (mean: 97.9% correct). As in the previous experiments, participants were slightly more accurate for *offset competitors* (98.5%) than for *onset competitors* (97.4%).

Table 4-2: Results of statistical analysis of interim trials comparing Experiment 1 to Experiment 3.

Factor	*B*	*SE*	*Z*	*p*
Block	.72	.060	12.06	<.0001
Word-type	-.88	.39	-2.25	.024
E3vsSync	-.29	.53	-.55	.58
E3vsDel	.91	.55	1.65	.10
Block × type	.18	.12	1.48	.14
Block × E3vsSync	.0022	.17	.013	.99
Block × E3vsDel	-.29	.18	-1.60	.11
Type × E3vsSync	.30	.95	.31	.76
Type × E3vsDel	-.12	.98	-.13	.89
Block × type × E3vsSync	-.13	.33	-.39	.70
Block × type × E3vsDel	.12	.36	.32	.75

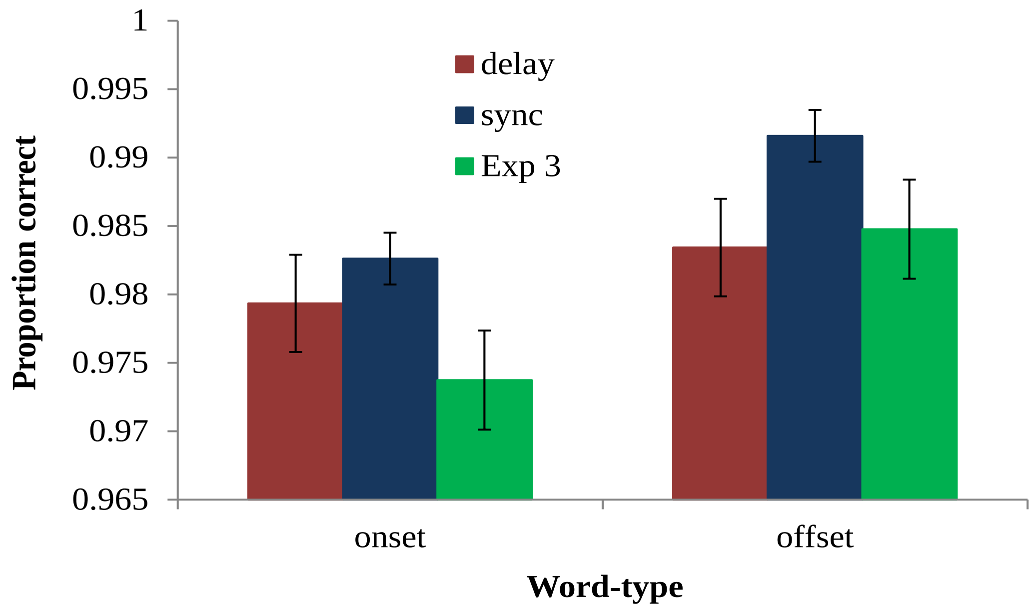

Figure 4-9: Experiment 3 – Accuracy for included participants in the VWP trials of Experiment 3 compared to both conditions of Experiment 1, by word type. Error bars represent standard error for that condition.

These data were analyzed in the same manner as the VWP accuracy data for Experiment 2: independent contrast codes were used to separately compare Experiment 3 to each training condition of Experiment 1. These contrast codes were included in an analysis along with separate interactions of each contrast code with word-type (contrast coded) in a mixed effects model with a binomial linking function. Participant and auditory word were random intercepts. The fixed factors were not correlated (all $R<.05$). This analysis revealed a significant main effect of word-type ($B=.52$, $SE=.13$, $Z=4.10$, $p<.0001$), with higher accuracy for *offset competitor* trials. Neither of the comparisons between Experiment 3 and the training conditions of Experiment 1 were significant (E3vsSync: $B=-.61$, $SE=.42$, $Z=-1.44$, $p=.15$; E3vsDel: $B=-.044$, $SE=.42$, $Z=-.11$, $p=.92$), nor was the interaction between word-type and E3vsSync ($B=-.46$, $SE=.35$, $Z=-1.33$, $p=.18$). However, the interaction between word-type and E3vsDel was marginally significant ($B=.56$, $SE=.31$, $Z=1.83$, $p=.068$). This interaction was explored with simple models comparing Experiment 3 to the *delay* condition separately for each word-type. These models maintained the same effects structure as the larger model, but only included the contrast codes for E3vsDel as a factor. These sub-analyses did not reveal significant effects of E3vsDel for either word-type (*onset competitors*: B=-.36, SE=.35, Z=-1.05, p=.30; *offset competitors*: B=.077, SE=.64, Z=.12, p=.90). Overall, these results indicate that participants in Experiment 3 were approximately as accurate during the VWP trials as those in Experiment 1.

4.2.2.3.2 Eye-tracking

In our analysis of the eye-movements on the VWP trials, only trials in which participants selected the correct target were analyzed (which included 97.9% of VWP trials). On these trials, as in the analysis of previous experiments, we investigated looks to the competitor item relative to looks to the unrelated distractors on the display. This is visualized in Figure 4-10 as the difference score across time, and in Figure 4-11 as the log-odds-ratio across time. For *onset competitors*, Experiment 3 participants appear to

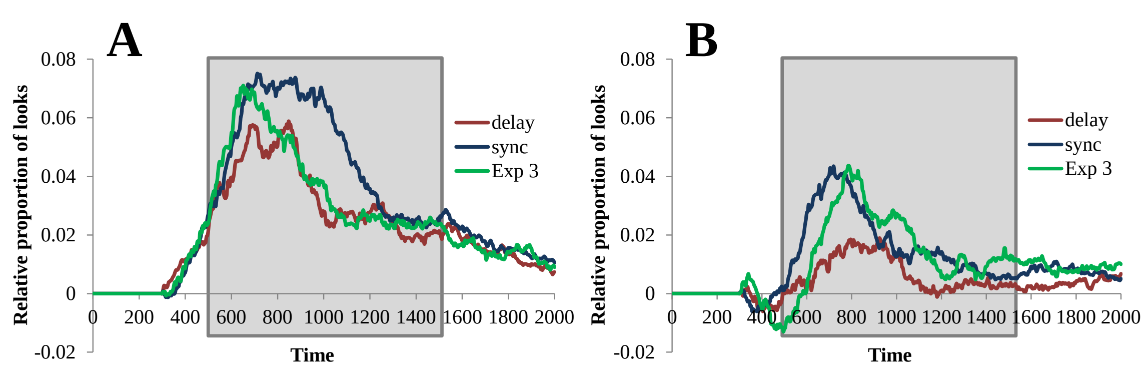

Figure 4-10: Experiment 3 – Relative proportion of looks to competitor objects (competitor – average unrelated). Analysis window highlighted in gray. A) Onset competitors. B) Offset competitors.

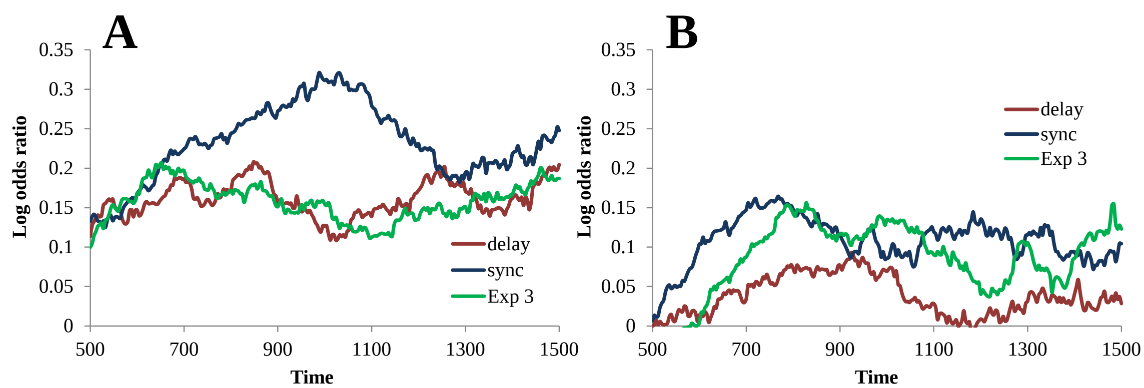

Figure 4-11: Experiment 3 – Log-odds-ratio of proportion of competitor looks to looks to average unrelated item across time, by training condition. A) Onset competitor trials. B) Offset competitor trials.

pattern quite similarly to the *delay* group in Experiment 1 (although there is some evidence of larger interference effects early, especially in the difference scores). For *offset competitors*, however, the Experiment 3 participants appear to pattern much more closely with the *synchronous* group, showing increased interference, despite a trend toward increased competition for *offset competitors*; recall that for Experiment 2, *offset competitors* showed more enduring evidence of increased competition given the additional processing time.

These data were analyzed using a mixed effects model with a linear linking function. The mean log-odds-ratio across the analysis window (500-1500 ms) was used as the DV (Figure 4-12). Because this model used a linear linking function, MCMC simulations were used to determine significance levels. As in the analysis for Experiment 2, two separate models were used to independently compare the results of Experiment 3 to each condition in Experiment 1.In each of these models, a contrast code was used to compare the training conditions (-.5 for either *synchronous* or *delay*, +.5 for Experiment 3). Word-type (contrast coded) was also a fixed factor, while participant and auditory word were random intercepts. The fixed factors were not correlated (all R<.002).

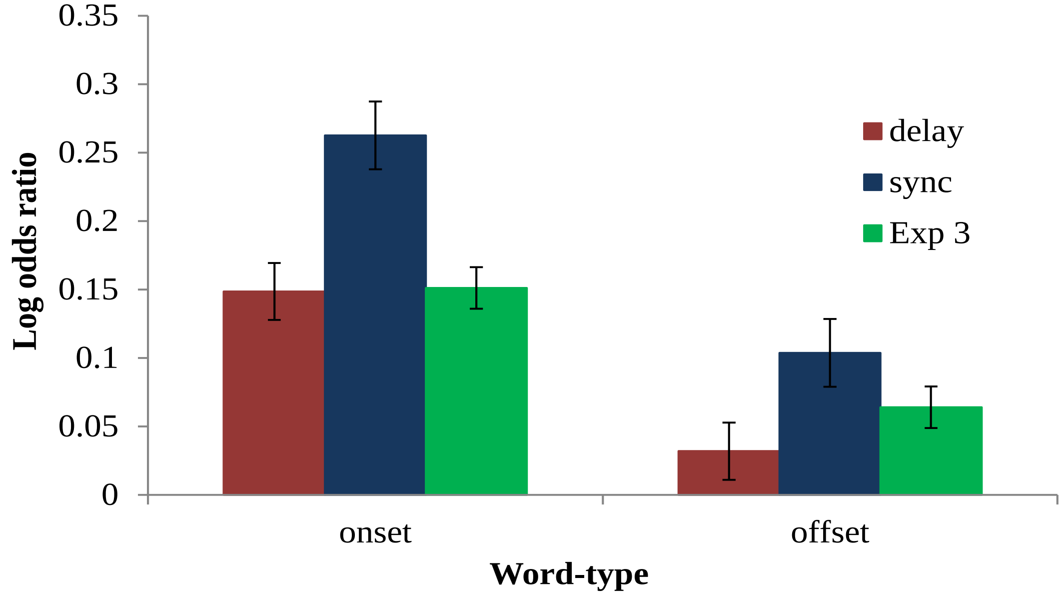

Figure 4-12: Experiment 3 – Log-odds-ratio within the 500-1500 ms analysis window comparing Experiment 3 to both conditions of Experiment 2, by word type. Error bars represent standard error for that condition.

The model comparing Experiment 3 to the *synchronous* condition of Experiment 1 showed a significant effect of training condition (*B*=-.082, *SE*=.035, p_{mcmc}=.022), with greater interference for the *synchronous* participants than for the Experiment 3 participants. There was also a significant main effect of word-type (*B*=-.12, *SE*=.040, p_{mcmc}=.021), with greater interference effects for *onset competitors*. The interaction was not significant (*B*=.072, *SE*=.059, p_{mcmc}=.23). The model comparing Experiment 3 to the *delay* condition of Experiment 1 did not show an effect of training condition (*B*=.022, *SE*=.034, p_{mcmc}=.52). The main effect of word-type was significant (*B*=-.10, *SE*=.033, p_{mcmc}=.017), but the interaction of training condition and word-type was not (*B*=.026, *SE*=.053, p_{mcmc}=.62). That is, the participants in this experiment showed interference effects quite similar to those of the *delay* group in the first experiment, with reliably less interference than the *synchronous* group. However, the trend toward greater competition for *offset competitors* with *synchronous* timing suggests that with additional power, the

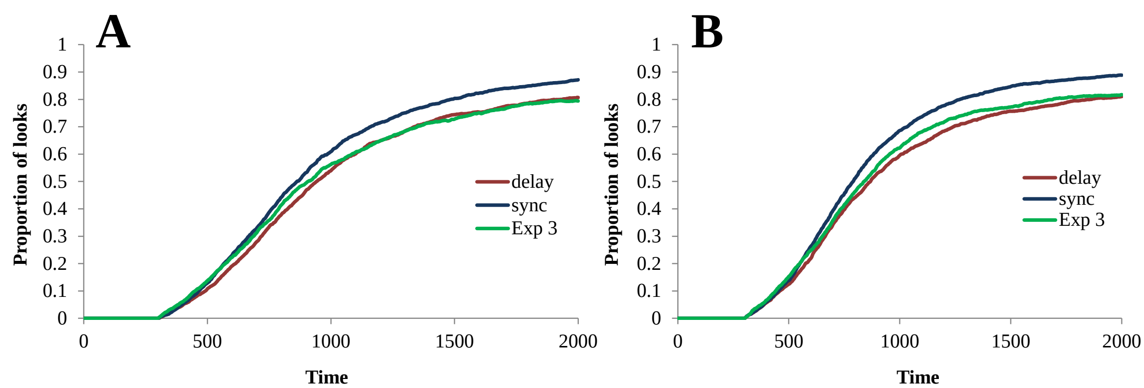

Figure 4-13: Experiment 3 – Proportion of looks to target items across time, by training condition. A) Onset competitor trials. B) Offset competitor trials.

effect may have emerged (though exploratory simple effects analyses did not approach significance: *p*=.50).

Target looks were analyzed using the same analysis window (500-1500 ms), and using the same form of parallel models used to analyze competitor looks. As in the previous experiments, data were entered using the empirical logit transformation of the raw proportion of target fixations during the analysis window. Overall, the looks to the target in Experiment 3 looked quite similar to those by the *delay* group in Experiment 1 (Figure 4-13). Mixed effects models were used to analyzed these data (linear linking function; no correlation between fixed effects: all R<.001). The model comparing against the *synchronous* group revealed an effect of training condition (B=-.15, SE=.029, p_{mcmc}<.0001), as participants in the *synchronous* group of Experiment 1 looked more to the target than did participants in Experiment 3. There was also a main effect of word-type (B=.11, SE=.021, p_{mcmc}=.0009), with more looks to the target on *onset competitor* trials. The interaction of word-type and training condition was not significant (B=-.012, SE=.043, p_{mcmc}=.78). The model comparing against the *delay* group again showed no effect of training condition (B=-.0056, SE=.031, p_{mcmc}=.86). There was a main effect of word-type (B=.088, SE=.019, p_{mcmc}=.0028), with more target looks for *onset competitor* trials. The interaction was not significant (*B*=.025, *SE*=.038, p_{mcmc}=.52). These analyses show that participants in Experiment 3 quite closely mirror those from the *delay* group in Experiment 1 in terms of looks to the target items in the display.

4.2.3 Discussion

In Experiment 3, all participants heard the auditory stimuli while the referents were displayed, as in the *synchronous* condition of Experiment 1. They were not provided with additional processing time at the end of trials, but did receive a preview period in which they saw the referents before the auditory stimulus was presented. This was provided as a way to equate the total trial duration of the *delay* trials in Experiment 1 while providing the opportunity for participants to continuously form word-referent mappings during the auditory stimulus. This design was predicted to elicit spurious associations with competitor word-forms, as seen in the *synchronous* group of Experiment 1. However, participants in Experiment 3 showed no evidence of spurious associations; instead, their performance quite closely mirrored that of the participants in the *delay* group from Experiment 1.

These results were unexpected. Although the Experiment 3 participants had longer trial durations, the trial ended shortly after word offset, as in the *synchronous* condition of Experiment 1, so the learners could not continue updating their representations after competition resolved. However, they seem not to form the spurious associations found when no preview period is given. Three possible explanations for why this preview period blocks the formation of spurious associations are explored. First, the poor performance during pre-exposure may have weakened coactivation during word-referent learning. Second, the longer trial duration may have changed the learning because participants had less time pressure during the learning phase; the rapid sequence of training trials in the *synchronous* training of Experiment 1 may have encouraged the formation of word-referent mappings before participants were ready, in order to prepare for the following trial. Finally, the preview period may have served to help learners form expectations for the upcoming auditory stimulus; by biasing the interpretation of the information, competitors could be suppressed from quite early in the auditory stimulus. I discuss each in turn.

Participants in this study showed much poorer performance in the pre-exposure phase than did participants in Experiment 1. This may indicate that the participants in Experiment 3 were not attending to the stimuli closely enough to learn the auditory word-forms during pre-exposure. This would limit the degree of coactivation during early word-referent learning trials, which would in turn weaken any spurious associations. However, it is unlikely that the lack of spurious associations arise solely from this result. Many participants performed quite well on this task, and there was very little evidence of spurious associations for *onset competitors*. It is possible that with greater attention during pre-exposure, the *offset competitor* effect may have reached significance. Thus the poor pre-exposure performance may have reduced some evidence of spurious associations, but is unlikely to explain the entire pattern.

Shorter trial durations could affect participants' encoding strategies, as they must prepare for the successive trial. Participants in the *synchronous* condition of Experiment 1 cycle rapidly through trials, whereas the trials in the *delay* are a good deal longer. When the trials cycle quickly, participants may adjust their encoding accordingly, by setting a specific time to learn or lowering the threshold for when learning should occur. The preview before the trials in this experiment may serve as a way to decrease time pressures, as the trials have longer duration than the *synchronous* condition of Experiment 1.

However, the time pressures on participants in the *synchronous* condition of Experiment 1 can be quite easily alleviated by the participants to allow more natural learning times. In all three experiments, the start of trials is self-paced, with participants clicking on the blue center dot to begin the trial. If the speed of trials was impinging on learning, participants could wait to click this dot (and thereby create a blank preview period before auditory and visual presentation). Time pressure thus may not be as severe as the *synchronous* trials suggest. More telling, if time impaired overall learning, the *synchronous* participants would be expected to show poorer accuracy and slower learning during the interim testing trials; no such evidence was found, suggesting that time pressures did not impair learning in this group. Finally, the overall trial duration was identical in Experiment 2, and timing pressures may have been even weaker (as no information was available during the delay); nonetheless, participants in Experiment 2 showed evidence of spurious associations with *offset competitors*.

Further, it is unclear what form of adjustment in learning as a result of time pressures could elicit both the results of Experiment 2 and those of Experiment 3. Participants in these experiments have equal trial durations, suggesting similar time pressure. However, those in Experiment 2 showed evidence on spurious associations for *offset competitors*; the time pressures appear insufficient to eliminate learning for late-occurring competitors in this condition. Experiment 2 should have particularly weak time pressures, as no information is provided in the delay between trials; in Experiment 3, visual information is provided, which offers at least some opportunity for stimulus processing. There is thus no clear explanation for why the even weaker time pressures of Experiment 2 would fail to alter the learning strategy as in the *onset competitors* of that experiment and both sets in Experiment 3.

An alternative explanation for these results is that the preview period served to help learners make predictions about the upcoming auditory stimulus, and thereby changed the lexical processing when the auditory stimulus was received. Learners' processing of the visual information before auditory presentation may have impacted the course of auditory processing. The visual presentation included only two referents, which would allow participants to activate the names for each of the objects during the 1000 ms delay before auditory stimulus onset. Such pre-naming of the objects could lead to more efficient processing of the word-forms once they are encountered. If the participants accurately covertly name both objects, they can suppress the competitor before the word even begins; if they know that a *gonu* is on the screen and no *goba* is present, they can activate the correct word-form upon only hearing the word-initial /g/. Such predictions would greatly reduce parallel associations, as parallel activation in general would be decreased. However, such effects could only emerge once participants can accurately name the visual objects – in early trials, the participants would be unable to activate the correct name in order to begin suppressing competitors. This would predict that earlier in training, participants in this condition should show increased competition, but that additional training eliminates this effect, as participants learn the word-referent pairings[4]. The data in the current experiment can not address this, as associations were only gauged at a single point, after the words were learned quite well. Such concerns of pre-naming are common in eye-tracking studies, as there is worry that participants are considering only the words of the displayed objects rather than using a more natural, unconstrained set of referents (Huettig & McQueen, 2007; Magnuson et al., 2007). However, these debates are as yet unresolved, leaving the extent to which pre-naming affects lexical processing an open question.

The *synchronous* condition of Experiment 1 does not provide an opportunity to rely on such predictions to limit the considered acoustic candidates. During the *synchronous* presentation, auditory and visual information start simultaneously, so lexical processing begins before the participant has a chance to name the visual items. This allows unsupervised associative learning to initiate based on ongoing processing of both the visual and auditory stimuli.

---

[4] This would also help in natural word-learning situations. As the learner becomes more confident in word-referent pairings, they can begin to form predictions about what words will be said, and thereby decrease competitor effects.

Further experiments are thus needed in order to identify why participants with a preview period differ from those with the *synchronous* trial structure from Experiment 1. Although these results are suggestive that visual processing is playing a major role in the formation of word-referent associations, deeper investigation of pre-naming and visual activations could greatly clarify how learning is proceeding. Such experiments could test at different points throughout learning; participants in the *synchronous* condition should strengthen the spurious word-referent connections with additional word-referent training, whereas those in the preview condition of Experiment 3 should show a decrease in the effect after the word-referent mappings have been learned effectively. Alternatively, the inclusion of some form of auditory mask to limit pre-naming could eliminate learners' ability to begin activating auditory word-forms before hearing the stimulus. Finally, the use of a within-participants manipulation can gauge whether time pressures lead to global changes in learning strategy. Experiment 4 utilizes a within-participants design, and so can speak to this.

4.3 General discussion

The first experiment in this dissertation demonstrated that manipulating the relative timing of auditory word-forms and their referents alters what associations are formed between word-forms and referents. When participants saw the potential referents while resolving lexical competition, they showed evidence of forming associations with both competing word-forms in parallel. The results of Experiments 2 and 3 offer a more nuanced picture of this process. These findings build on the evidence from Experiment 1, but they suggest that the interactions between stimulus timing, perceptual processing and word learning are quite complex. Associative learning appears to go beyond forming links between whatever stimuli are perceptually present; instead, learning also occurs on the basis of ongoing processing for working memory representations of stimuli.

Experiment 2 showed that learning does not end when visual and auditory stimuli end, but instead continues as participants continue processing these stimuli. Providing a delay between trials after the end of both auditory and visual presentation led to a decrease in competition from *onset competitors*. Participants were able to update their learned representations during this delay to reflect the resolution of competition from these competitors. However, such a decrease in competition was not apparent for *offset competitors*. Instead, participants continued to show evidence of spurious associations with these competitors. The delay between trials was thus not sufficient to overwrite the associations formed during lexical competition. *Offset competitors* are activated late in lexical processing, and their activation may extend until after word offset. Additionally, no forthcoming information after word offset contrasts with the *offset competitors*, so there is no additional pressure to suppress these competitors. As such, they may remain active longer, and continue to associate with the referents[5].

The results of Experiment 2 thus point to a complex interplay between stimulus timing, lexical processing dynamics and learning. Spurious associations do not form simply on the basis of playing words and showing pictures simultaneously; instead, the critical component is that the learner is maintaining active *representations* of a word-form and referent simultaneously. The learning can occur on the basis of memory representations of the stimuli or actual physical co-occurrence.

Experiment 3 presented a very different stimulus timing, in which visual referents were available before auditory stimuli. This provided participants with an opportunity to process the visual referents and predict their auditory labels (if the word-referent pairings were sufficiently learned). However, if the formation of word-referent mappings occurs

---

[5] In order to eliminate concerns of working memory representations of the stimuli, we attempted an experiment including visual masks after visual presentation. The results turned out quite similarly to those in Experiment 3. However, the visual masks likely don't flush working memory, as there was no need for participants to encode these masks.

only on the basis of associating objects during perceptual co-occurrence, this design should continue to show the formation of spurious associations (as should the design in Experiment 2); when the auditory word was played, the visual referents were on the screen. In this case, participants showed no evidence of learned spurious associations between referents and phonologically-competing word-forms. This may suggest that the dynamics of visual object processing affect auditory activations (and thus learning of auditory-visual associations) – the capacity to name the visual stimuli before the auditory stimuli reduces competition during lexical activation (e.g., Chen & Mirman, 2012). Although the auditory signal and the visual referent are co-present during the trial, the processing of the auditory signal changes as a result of the preceding visual information. These results point to a failure to consider the processing dynamics of the *visual* stimuli in this experiment; including a visual preview period does more than lengthen trial, as participants are also actively processing these stimuli. Further research is needed on this point, but it offers a compelling explanation for the results.

These results may thus also support a complex relationship between stimulus timing, on-line perceptual processing and learning. Interaction between representations can affect the activation profiles for stimuli before they have occurred, which in turn affect how the processing of these stimuli impacts learning. Unsupervised associative learning does not occur in a vacuum of raw stimulus-stimulus associations, but instead reflects specific activation patterns for representations of stimuli that emerge from a given task structure.

CHAPTER 5

EXPERIMENT 4: TIMING EFFECTS WITHIN-PARTICIPANTS

The preceding chapters detail how the relative timing of auditory and visual information affect the representations formed during word learning. Specifically, these experiments suggested that when the formation of associations can only occur during periods of lexical competition, learners form mappings between the correct word-form and its referent, as well as between competing word-forms and the same referent. Parallel activation during lexical access leads to the formation of parallel word-referent mappings. These results point to a critical interplay between real-time stimulus processing and learning in terms of which representations are formed.

Although these preceding manipulations of timing are suggestive, it remains possible that the changes in the representations formed by participants were a result of more global changes in learning, rather than local differences in stimulus processing. Experiment 3 showed that preceding visual information eliminated the formation of spurious associations. Although this was interpreted as evidence of visual processing altering auditory activation, it may also have been evidence of more global changes in learning strategy when time pressures are relaxed. In each of the preceding experiments, all word-referent training trials for each participant used the same stimulus timing. Thus it is possible that the timing condition altered the way that the participant formed word-referent mappings across all trials. For example, participants with *synchronous* presentation may have opted to always encode information early in the trial, whereas those in the *delay* presentation and those in Experiment 3 always waited until after competition resolved. Although such strategic changes appear unlikely, particularly given the results of Experiment 2, evidence that changes in learned representations emerge as a result of more local stimulus characteristics would offer a more thorough demonstration of real-time interactions between lexical activation processes and word learning.

To this end, Experiment 4 utilizes a within-participants design where timing is manipulated between word pairs learned by all participants. If the timing within an individual learning event affects the word-referent associations that a learner forms, words trained with *synchronous* timing should show increased competitor effects relative to those with *delay* timing.

5.1 Introduction

To investigate whether the increased competitor effects with *synchronous* stimulus timing emerged from continuous learning during periods of lexical competition, or if instead they resulted from altered overall processing strategies caused by changes in stimulus timing, Experiment 4 used a within-participants design to measure how timing manipulations affect word learning. This experiment mirrors many aspects of Experiment 1, with the same *delay* timing and a *synchronous* condition without additional processing time following visual offset, to determine whether participants show word-specific increases in competitor effects when only some of the trained words are presented with auditory-visual synchrony. In this experiment, primarily *onset competitor* words are used; the within-participants design necessitated increasing the number of words to increase power, and this was best accomplished by focusing on a single type of competitor word.

Evidence of word-specific changes to the representations formed in this study would reinforce the theory that the differences in learning in previous experiments occur because of online processing interacting with the learning process. However, if participants show a similar degree of competition between both *synchronous* and *delay* words, this suggests that the earlier effects arise from more global changes in the method of learning.

5.2 Methods

5.2.1 Participants

Forty-six participants from the University of Iowa community completed the study. Participants were paid $15 or received partial course credit for their participation. All participants self-reported normal hearing and normal or corrected-to-normal vision. Data from thirty-seven participants were included in analyses; eight additional participants were excluded for low accuracy at test (below 75% on all four *onset competitor* pairs), and one participant was excluded as a result of a poor eye-tracker calibration.

5.2.2 Design

The design mirrored that of Experiment 1 in many ways. Participants completed a phoneme-monitoring task to become familiarized with the words used in the study, and then learned the words in a cross-situational word-referent training task. Intermittent active trials were used throughout training to maintain participant focus and gauge learning. Participants were then tested using the VWP to determine the degree of interference from the referents of phonological competitors of the target word (and thus to gauge whether false associations were formed during learning).

Several aspects of the design differed from Experiment 1. First, participants in this experiment learned eight *onset competitors* and four *offset competitors*. The additional *onset competitors* were necessary to increase power given the within-participants design. The *offset competitors* were included to encourage participants to attend to word onsets[1]. These *offset competitors* were not used in analysis, as there was only one pair per condition per participant, leading to quite noisy data; inclusion of more

[1] An earlier version of this experiment did not include *offset competitors*. This resulted in decreased overall competitor effects. This likely occurred because listeners could ignore word onsets and still learn words quite effectively.

words was deemed out of reach for a single learning session, as pilot work showed poor learning when 16 words were trained.

Interim test trials also differed from Experiment 1. In Experiment 4, these trials looked much like the word-referent training trials, but participants were asked to choose the correct referent. This allowed us to maintain the same appearance of all trials, continue timing manipulations during these test trials, and gauge learning in a manner that was more akin to the actual learning trials.

Most importantly, the timing manipulation was done within-participants, rather than between-participants. Two of the four pairs of *onset competitors* and one of the two pairs of *offset competitors* were assigned to the *synchronous* condition; the other words were assigned to the *delay* condition. This assignment was chosen randomly for each participant from the pairs in the word list assigned to that subject. The two words in the competitor pair were always assigned to the same timing condition. During the training and interim trials, *synchronous* and *delay* trials were randomly intermixed, such that the participant received both types of trials throughout training but not in a predictable order.

5.2.3 Stimuli

Participants learned 12 novel words mapped to images of novel objects. The word-object pairings were randomized for every participant. The visual stimuli included the eight images used in the previous experiments, as well as four additional images that came from the same original set that the first eight were chosen from (Figure 5-1).

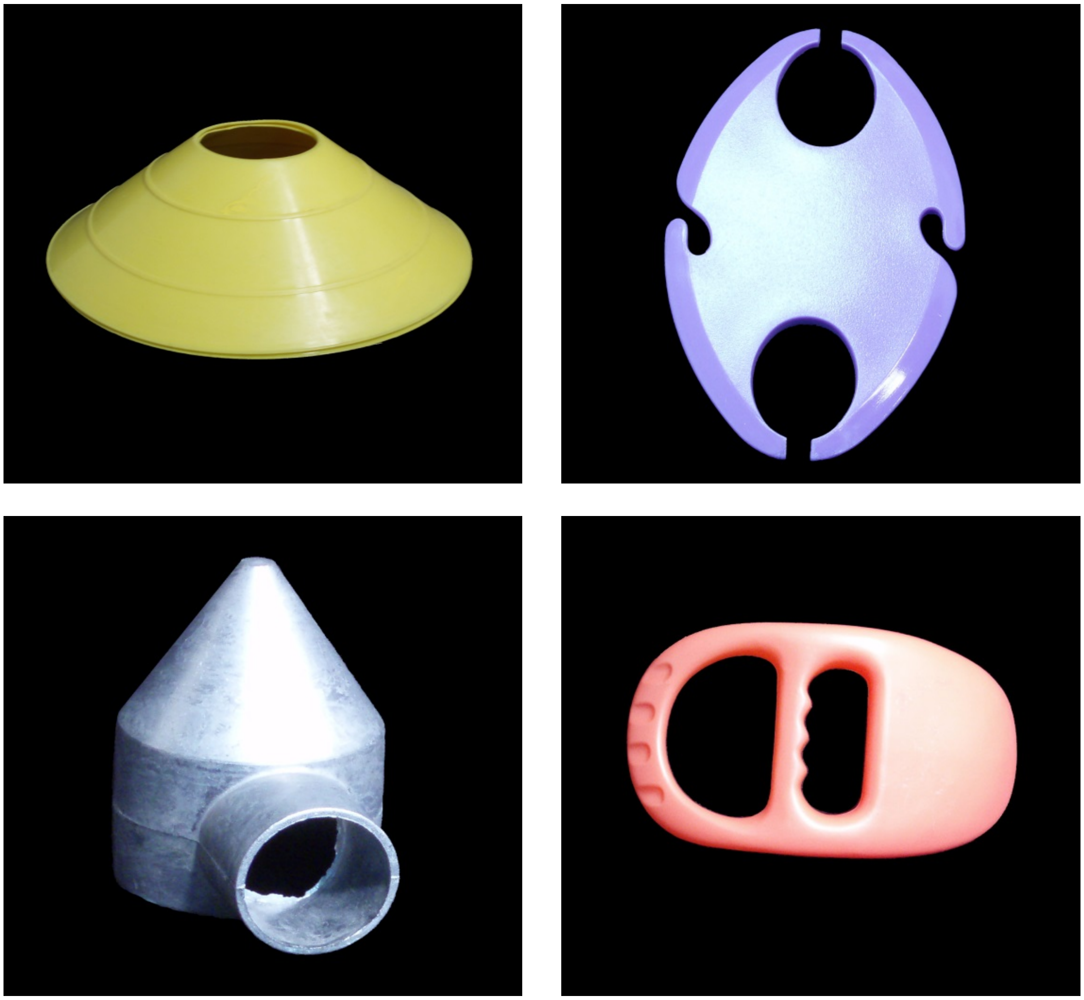

Figure 5-1: Additional visual referents included in Experiments 4 and 5.

Increasing the number of words required altering the set of words learned. There is notable variability between words in terms of the degree of competitor effects; in an earlier version of this experiment, several of the words used produced minimal competitor effects in either condition. These poor competitor effects indicate little online coactivation between the words, and thus little chance to observe learning during (weak) parallel activation. To combat this issue, two distinct sets of words were chosen, with half of the participants learning each set (Table 5-1). This increases the likelihood of choosing words that exhibit strong online competition effects, and thus that offer the capacity to support learning of spurious associations during lexical competition.

Table 5-1: Phonetic transcription of words taught to participants in Experiments 4 and 5.

WORD SET 1				WORD SET 2			
Onset competitors		**Offset competitors**		**Onset competitors**		**Offset competitors**	
ɵupsɛd	ɵupsiv	zɑimpæf	jempæf	rubsɪf	rubsʌp	dævlik	hɛvlik
bivdup	bivdʌf			vimbɑim	vimbæl		
kælboʊm	kælboit	roikzɪb	boʊkzɪb	zeɪftad	zeɪftug	mɑgrʌs	dʌgrʌs
vɪsmeɪv	vɪsmɑk			lɪdzoʊv	lɪdoil		

The auditory stimuli were two-syllable CVC'CVC words that adhered to the phonotactic rules of English and were syllabified between the two medial consonants. The switch to the more complex, six-phoneme words was chosen to increase the amount of overlap between competitors (in Experiment 1, competitors shared two of four phonemes; in this experiment, they shared four of six phonemes) and to limit the chance that some words could be learned by analogy to similar-sounding real words (e.g. in Experiment 1, *bure* sounds similar to *BlueRay*). Within each word set, words were chosen to minimize overlap across competitor pairs; there was no overlap between pairs at any phoneme position. The word sets included four pairs of *onset competitors,* which overlapped on the first four phonemes, and two pairs of *offset competitors,* which overlapped on the final four phonemes.

Auditory stimuli were recorded from a male native speaker of English in a sound-treated room. This speaker produced approximately 25 exemplars of each word. These were isolated in Praat and the 15 clearest exemplars of each word were selected. These were adjusted to have the same peak amplitudes, and 100 ms of silence was added to the beginning and end of each token.

5.2.4 Procedure

This experiment followed the same basic procedure as in the previous experiments, but with minor variations in each task. The order of tasks remained the same, and all participants completed the same series of tasks.

5.2.4.1 Pre-exposure

Pre-exposure used phoneme-monitoring to familiarize participants with the words used in the study. Participants were advised to press the spacebar if the word they heard contained a "V" sound, and do nothing if it did not. Real-word examples of words with this sound in various lexical positions were provided in the instructions for the task. In each word set, five of the 12 words contained a "V" sound somewhere in the word. The timing for these trials was identical to that in previous experiments (if no space bar press was registered within 2000 ms of stimulus onset, the trial ended). All *onset competitor* words were presented eight times during this pre-exposure phase. Due to a programming error, half of the *offset competitor* words were presented eight times; the other half did not appear during pre-exposure[2]. Throughout pre-exposure, the auditory exemplar played was randomly selected from the 15 recordings of each word.

5.2.4.2 Word-referent training

Training trials used the cross-situational learning paradigm used in earlier studies to teach participants the word-referent pairings. For *synchronous* words, the visual stimulus presentation began simultaneously with auditory onset, and the images remained on the display for 800 ms. After this period, the trial completed, and the screen was blank for 550 ms before the next trial began. For *delay* words, visual stimuli onset 1000 ms

---

2 Although this oversight was unfortunate, it is not predicted to affect the results; the *offset competitor* trials are not analyzed throughout the study, as they are included primarily to encourage participants to attend to word onsets. Of the stimuli played, four words in word set 1 had "V" sounds and five in word set 2 did.

after the onset of auditory stimuli. They also remained on the screen for 800 ms, followed by 550 ms of blank screen before the succeeding trial.

During the cross-situational word-referent training trials, participants formed associations based on the degree of co-occurrence between a word-form and a referent. In all the experiments in this dissertation, the randomly selected foil referent on any given trial was never the referent of the phonological competitor, to ensure that the learners had no reason to form an association between phonological competitors other than through parallel activation. However, in Experiments 1-3, all other referents were potential competitors. This should lead to weak associations between each referent and every other word except the phonological competitor; on 1/6 of trials, that pairing was accurate. These weak associations may mask the predicted competitor effects in this study; at test, the participant should have weak associations to the two unrelated items, making them only partially unrelated to the target. While this was clearly not an issue overall in Experiment 1 (since we nonetheless observed these effects), there was concern that the (possibly subtler) within-participant effects might be masked by such effects. Thus, to counteract these associations, we yoked the competitor pairs in Experiment 4 such that two pairs of items never co-occurred during training. These pairs were then used together at test to ensure that neither the competitor nor the unrelated items had encouraged the formation of associations with the referent during training.

Two of the four *onset competitor* pairs and one of the two *offset competitor* pairs were assigned to each timing condition. Assignments were randomly selected for each participant from the word list assigned to that participant. Each pair was yoked to a pair from the opposite timing condition (i.e. each *synchronous* pair was yoked to a *delay* pair) to never co-occur during word-referent training. The *offset competitors* were always yoked to each other. Throughout training, the competitor for each trial was selected such that it was not the referent of the phonological competitor of the target, nor was it the referent of either of the words in the yoked pair; the competitor was randomly selected from the remaining eight referents. The placement on the display of the target and distractor object was randomly selected for each trial.

Each word was presented as the correct target 32 times, for a total of 384 training trials. These trials were blocked such that each word was a target once before any repetitions were encountered (interspersed with interim trials, as described below). Auditory stimuli during training included 10 of the 15 available exemplars for each word. Which 10 exemplars were used during training was randomly selected for each participant. During word-referent training and interim trials, a random choice from among these 10 exemplars was used each trial.

5.2.4.3 Interim testing trials

Interspersed throughout testing, participants completed active test trials to gauge how well they had learned the words, and to maintain their attention on the task. These trials differed substantially from the interim trials used in the previous experiments in order to more comprehensively control timing throughout training, to offer a more compelling measure of the trajectory of learning, and because parallel work (Roembke and McMurray, submitted) showed much better learning performance when participants made an active response during the learning trials.

The interim trials used the same presentation format as the training trials, with two referents displayed on the monitor. Participants clicked the blue dot to begin the trial, then heard the word and saw the referents with their assigned timing manipulation from the training trials. After the visual referents were removed from the display, two boxes appeared in their place. Participants then clicked on the box they believed the referent of the played word had appeared in. The boxes remained on the screen until the participant clicked in one of them. No feedback was given. After the response, the boxes disappeared and 550 ms elapsed before the start of the next trial. To alert the participants that a response was required on these trials, the background of the screen was green (instead of the white background of the regular word-referent training trials) throughout the trial.

For these interim trials, foils were chosen as in the training trials: the foil was never the referent of the target's phonological competitor, nor of either of the members of the yoked pair. Three interim testing trials were presented during each block of training. The target for these trials was selected randomly.

5.2.4.4 VWP testing trials

For the VWP testing trials, each competitor pair was always presented along with its yoked pair. Thus none of the four items on the display had ever appeared together during training. The location on the screen of each of these items was randomly selected for each trial.

Before the first VWP trial, a drift correction was performed on the eye-tracker. Participants then completed 240 test trials, with a drift correction every 24 trials. Each block of 12 trials included one repetition of each target word. Participants completed 20 such blocks. All VWP trials included the same timing as in all previous experiments, and there was no difference between trials for words in the different timing conditions. Auditory stimuli during the VWP trials were randomly selected from the five recorded exemplars of each word not used during training.

5.3 Results

5.3.1 Exclusions

Due to the more difficult learning in this experiment (the increased number of words, the longer words, and the greater overlap between words), it was necessary to adopt different exclusion criteria. Whereas in the previous experiments all participants with below 75% accuracy on any set of words were excluded, in this experiment exclusions were done on a word-pair basis. Any pair for which a participant correctly selected the target at least 75% of the time at test was included. For eight participants, accuracy was below this threshold for all four pairs of *onset competitors*; all data from these participants were excluded. These participants were evenly split between the two word lists (four from each). One additional participant was excluded from all analyses because of poor eye-tracking. The remaining 37 participants each contributed at least one word pair to analyses. The majority of these participants contributed data to both training conditions; three participants did not have an included *delay* pair, and another three did not have an included *synchronous* pairs. Five of these participants learned *word list 2*. The remaining 31 participants had at least one included word pair in both training conditions.

For these 37 participants, there were 148 possible pairs of *onset competitor* words (37 participants × four pairs). Of these 148 pairs, 114 had accuracy high enough at test to include in the analysis. Figure 5-2 displays the distribution of those pairs that were excluded. The 34 excluded pairs were exactly evenly split between the *synchronous* and *delay* conditions (17 excluded from each). However, the majority of excluded pairs came from *word list 2* (8 from *word list 1*, 26 from *word list 2*), suggesting that learners struggled more with this set of words (although approximately the same number of participants in the two groups failed to learn completely).

5.3.2 Pre-exposure

Overall, participants were relatively good at the pre-exposure task. Participants accurately identified words that contained a "V" sound in 91% of trials and accurately rejected words that did not have a "V" sound in 82% of trials. These values were quite consistent if word pairs that were excluded from analysis due to poor VWP performance were not included (92% correct positives; 82% correct rejections). Participants responding to *word set 2* were slightly more accurate than those responding to *word set 1*, especially for trials where no "V" sound was present (Figure 5-3). These data were

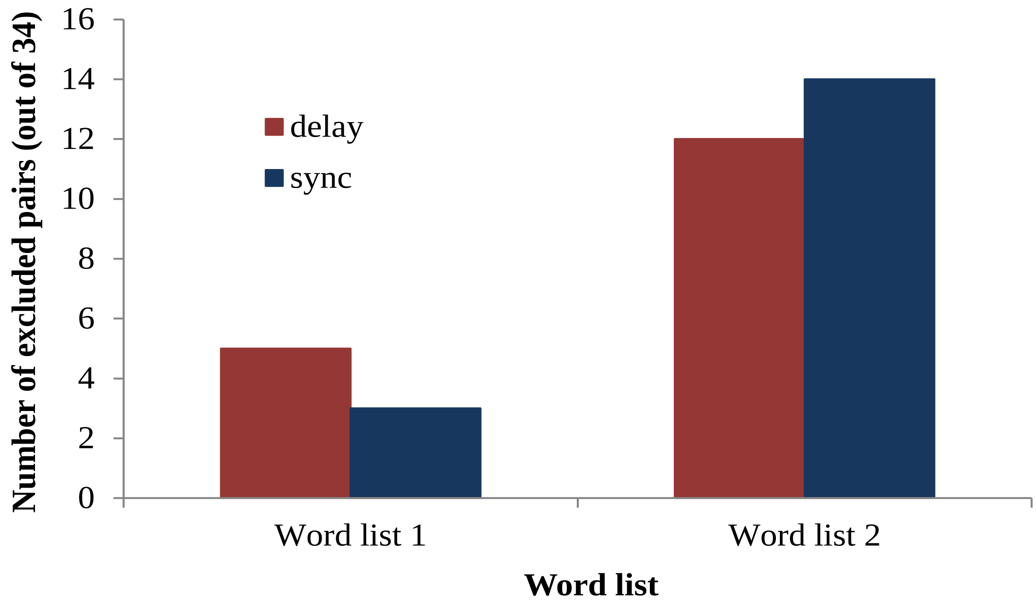

Figure 5-2: Experiment 4 – Number of word pairs excluded due to low accuracy at test, by word list and condition. Overall, 34 pairs were excluded.

analyzed with a mixed effects model using a binomial linking function. Word-set (contrast coded) and trial-type (without or with a “V” sound, contrast coded as -.5/+.5) were fixed factors, while participant and auditory stimulus[3] were random intercepts. No random slopes were included as they did not improve model fit for the VWP analyses (by χ^2 test, p=.41); all models within this experiment used the same random effects structure for consistency. The fixed effects were not correlated (R=-.088). This analysis revealed no significant main effect of list (B=.19, SE=.38, Z=.51, p=.61). Although participants were numerically more accurate for *word set 2*, this difference was not reliable. However, there was a main effect of trial-type (B=.94, SE=.24, Z=3.9, p=.0001), as participants were much more accurate for “V” present trials. There was no interaction (B=-.0042, SE=.48, Z=-.009, p=.99).

[3] In other cases, the identity of the stimulus has been used as a random effect. Here, I used the actual stimulus as the random effect because stimulus identity is perfectly correlated with trial-type. Including word rather than stimulus would thus mask effects of trial-type.

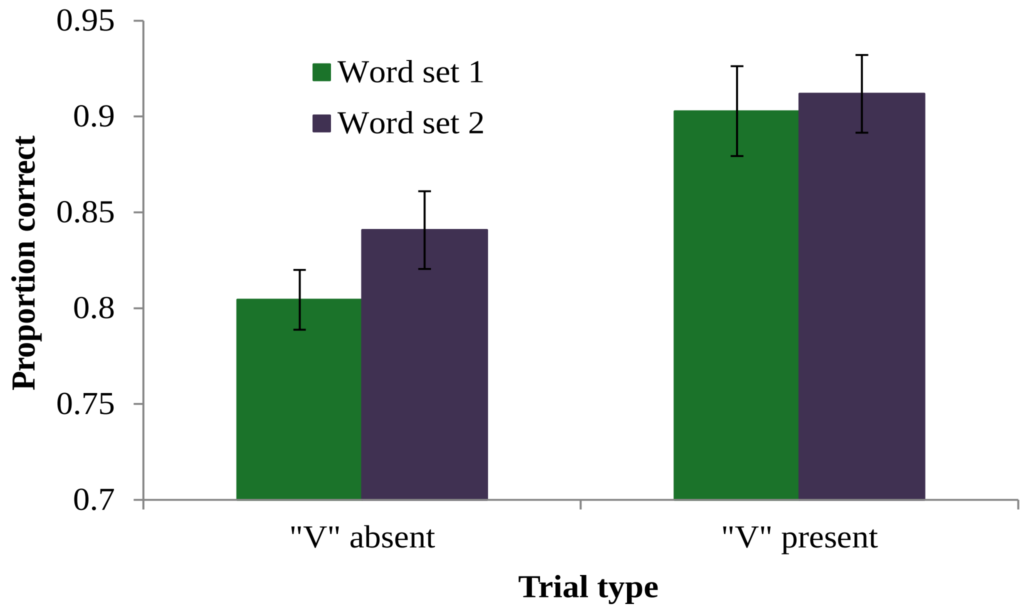

Figure 5-3: Experiment 4 – Accuracy of identifying whether a word had a "V" sound during pre-exposure, by word set and trial type. Error bars represent standard error for that condition.

5.3.3 Interim testing results

The interim testing trials provide insight into the progress of learning throughout training. This is particularly interesting in the current experiment, as many of the words were not learned effectively. These trials thus offer a measure of whether participants learned anything about these words, or whether they showed no improvement throughout training for the words. Figure 5-4 contrasts the course of learning for words whose accuracy at test was sufficient for inclusion with the learning for words which showed poor performance at test. Interestingly, both types of words reached near-ceiling performance by the end of training, though the words that were inaccurately identified at test did so more slowly[4]. Breaking down performance by training condition (Figure 5-5)

[4] The impressive performance for these trials likely emerges in part from the easier task than that used in the VWP: these trials only include two response possibilities, and phonological competitors are never present as foils.

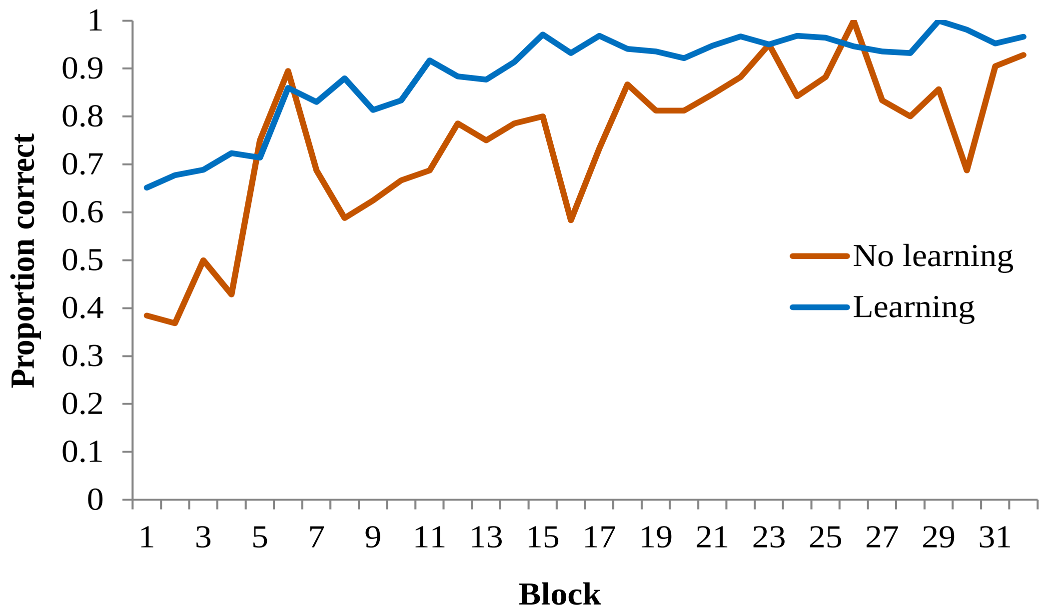

Figure 5-4: Experiment 4 – Accuracy of responses during interim testing trials for Experiment 4 words that were accurately identified during the VWP test and those that were poorly identified during the VWP test.

shows little effect of training condition for words that are learned effectively; words that were not learned well are more difficult to gauge, as the low number of observations per cell (only 17 word pairs of each type were not learned, and each block included only three total trials per participant) leads to noisy data[5]. Because there were so few observations of these trials, only words that were accurately learned were included in the statistical analysis[6].

The data for the words with accuracy above criterion in the VWP trials were entered into a mixed effects model with a binomial linking function. Accuracy during the

[5] Because of the very small number of word pairs with poor learning performance in *word list 1*, these data are not broken down by word list and accuracy at test. Figure 5-6 displays word list effects for words that were accurately identified at test.

[6] Including the words that were poorly identified at test exhibited qualitatively similar statistical results, except that the interaction between block and word list dropped out of significance.

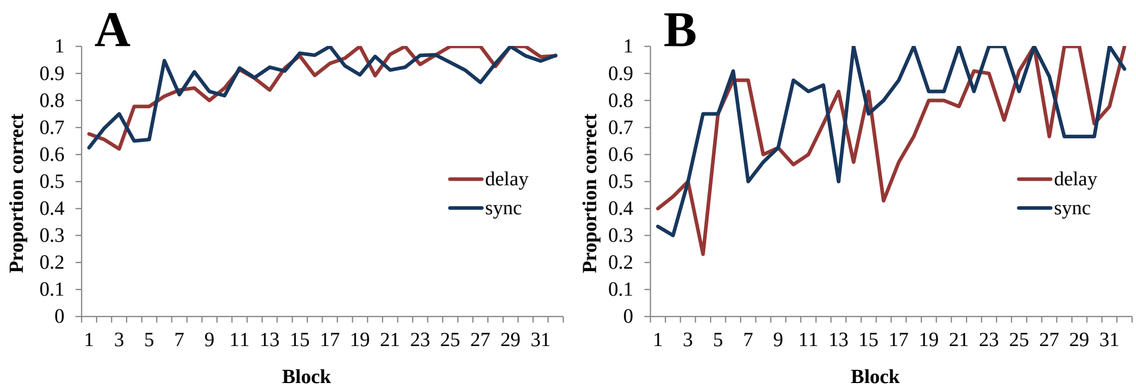

Figure 5-5: Experiment 4 – Accuracy of responses during interim testing trials for Experiment 4 across training blocks, by training condition. A) Words accurately identified at test. B) Words poorly identified at test.

interim trials was the DV, while training block (32 blocks[7], entered as raw block number), word list (contrast coded) and training condition (contrast coded) were fixed factors. Participant and auditory word were random intercepts. Word list and block were mildly correlated (R=-.22); no other factors were correlated (all R<.1). The effect of timing condition was not significant (B=-.29, SE=.28, Z=-1.03, p=.30), signaling similar learning between words with different stimulus timing. There was also no main effect of word list (B=-.19, SE=.38, Z=-.50, p=.62). The main effect of block was highly significant (B=.13, SE=.013, Z=10.25, p<.00001), as participants were more accurate on later blocks than on earlier ones. The effect of block interacted with word list (B=.073, SE=.025, Z=2.89, p=.0039), as the words in *word list 2* were learned slightly faster than those in *word list 1* (Figure 5-6). No other interactions reached significance (word list × training condition: B=.21, SE=.57, Z=.36, p=.72; training condition × block: B=.034, SE=.025, Z=1.35, p=.18; word list × training condition × block: B=.0068, SE=.050, Z=.14, p=.89).

[7] In this experiment, there were interim trials in every block, so learning was gauged more often. In the previous experiments, learning only occurred once per four blocks, hence the larger number of blocks in this experiment.

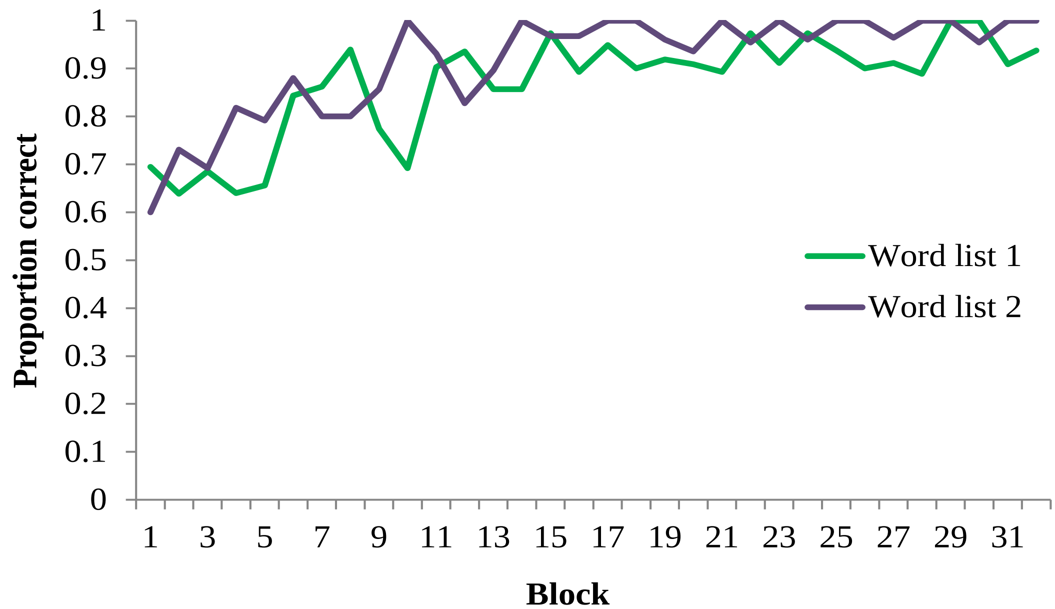

Figure 5-6: Experiment 4 – Accuracy of responses during interim testing trials for Experiment 4 across training blocks, by word list.

5.3.4 VWP testing trials

All VWP analyses for Experiment 4 focus exclusively on *onset competitor* trials. Given the design, there were not enough *offset competitor* trials to conduct an appropriate analysis of performance. These analyses were conducted only on those word pairs that passed the exclusion criteria detailed in section ***5.3.1***.

5.3.4.1 Accuracy

Accuracy for the word pairs included in analysis was quite high, and this was consistent across word lists and training conditions (Figure 5-7). A mixed effects model with a binomial linking function was used to interpret these data. Fixed effects were training condition and word list (contrast coded), while participant and auditory word were random intercepts. The fixed effects were not correlated (R=-.036). This analysis revealed no reliable differences in accuracy between any factors (training condition: B=-.

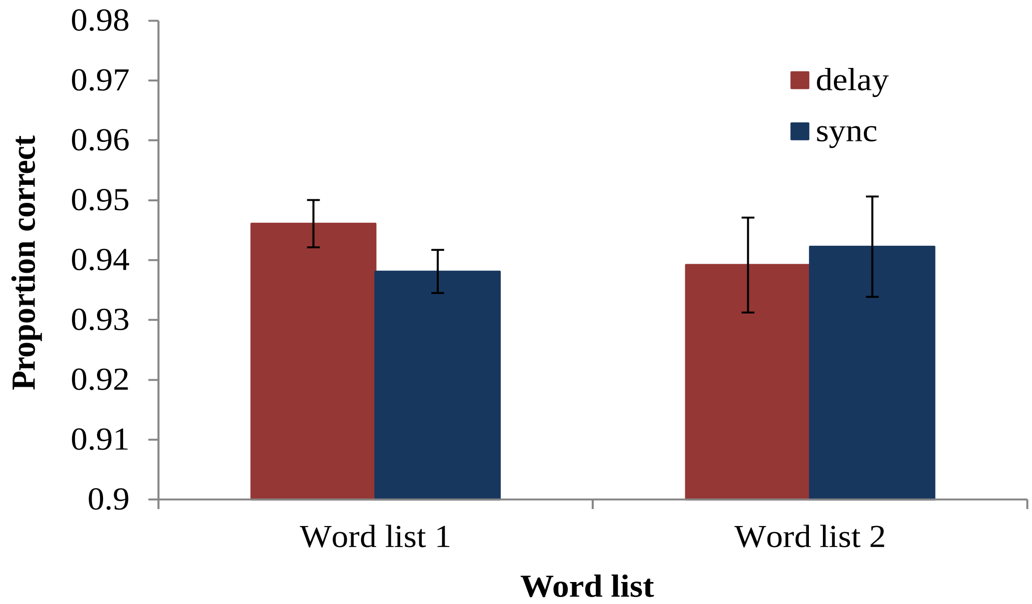

Figure 5-7: Experiment 4 – Accuracy for included word pairs in the VWP trials of Experiment 4, by training condition and word list. Error bars represent standard error for that condition.

× word list: *B*=-.19, *SE*=.25, *Z*=-.78, *p*=.44). As in the previous studies, there was no evidence that words training under *synchronous* timing were learned more poorly than those with the *delay* training.

5.3.4.2 Eye movements

Eye movements to visual competitors were analyzed relative to looks to the unrelated objects using log-odds-ratios, as in previous experiments (Figure5-8). We used the same analysis window as in previous experiments (500 ms to 1500 ms), and computed the log-odds-ratio based on the mean number of looks to the competitor object during this window relative to looks to the mean of the two unrelated items during the same window (Figure 5-9). This log-odds-ratio served as the DV for a mixed effects model of this data. One stimulus for one participant was excluded (out of 224) because no looks were made to either unrelated object during the time window; this stimulus came from the *delay* condition, and had no looks to the unrelated object. The model used a

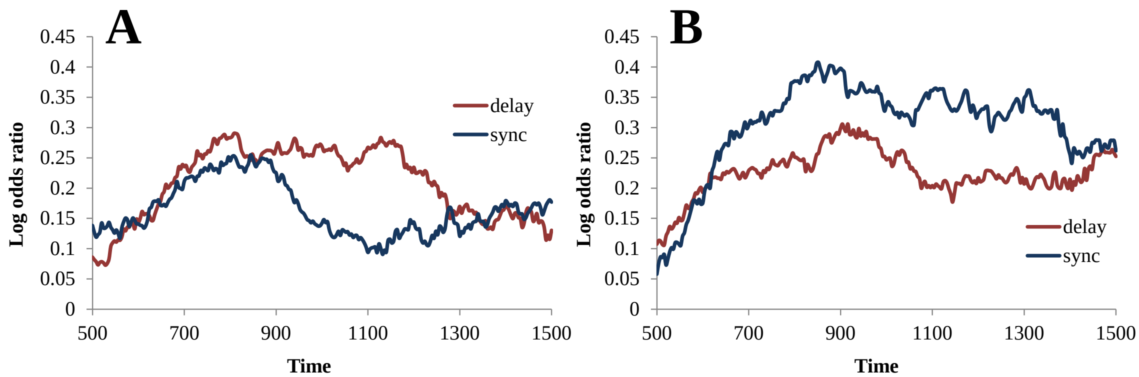

Figure 5-8: Experiment 4 – Log-odds-ratio of proportion of competitor looks to proportion of looks to average unrelated item across time, by training condition. A) Word list 1. B) Word list 2.

linear linking function, and included training condition and word list (contrast coded) as fixed factors, as well as participant and auditory word as random intercepts. The fixed effects were not correlated (*R*=-.018). To determine significance values, MCMC simulations were conducted.

Somewhat surprisingly, this analysis did not reveal a main effect of training condition (*B*=-.038, *SE*=.038, p_{mcmc}=.32). Word list was also not significant (*B*=.057, *SE*=.059, p_{mcmc}=.34). However, the interaction of training condition and word list was marginally significant (*B*=-.14, *SE*=.076, p_{mcmc}=.057). To examine this interaction, simple effects analyses were conducted on each word list. These models used the same structure as above, but included training condition as the only fixed factor. For *word list 1*, the effect of training condition was not significant (*B*=.035, *SE*=.049, p_{mcmc}=.47). However, for *word list 2*, the effect was marginally significant (*B*=-.11, *SE*=.059, p_{mcmc}=.083), with a larger competitor effect for the *synchronous* words than for the *delay* words.

As in Experiment 1, data from the first and second half of the VWP trials were analyzed to determine whether the effect was consistent throughout testing. Note that this experiment included fewer VWP trials (320 in earlier experiments; 240 in current experiment), so these conditions comprise somewhat less data. The log-odds-ratios across

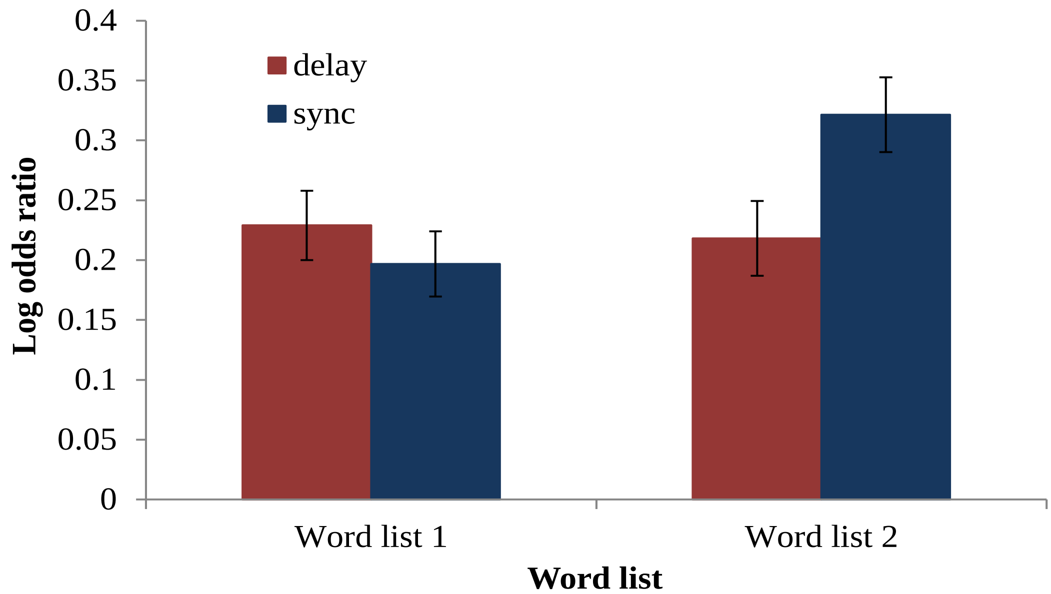

Figure 5-9: Experiment 4 – Log-odds-ratio within the 500-1500 ms analysis window, by training condition and word list. Error bars represent standard error for that condition.

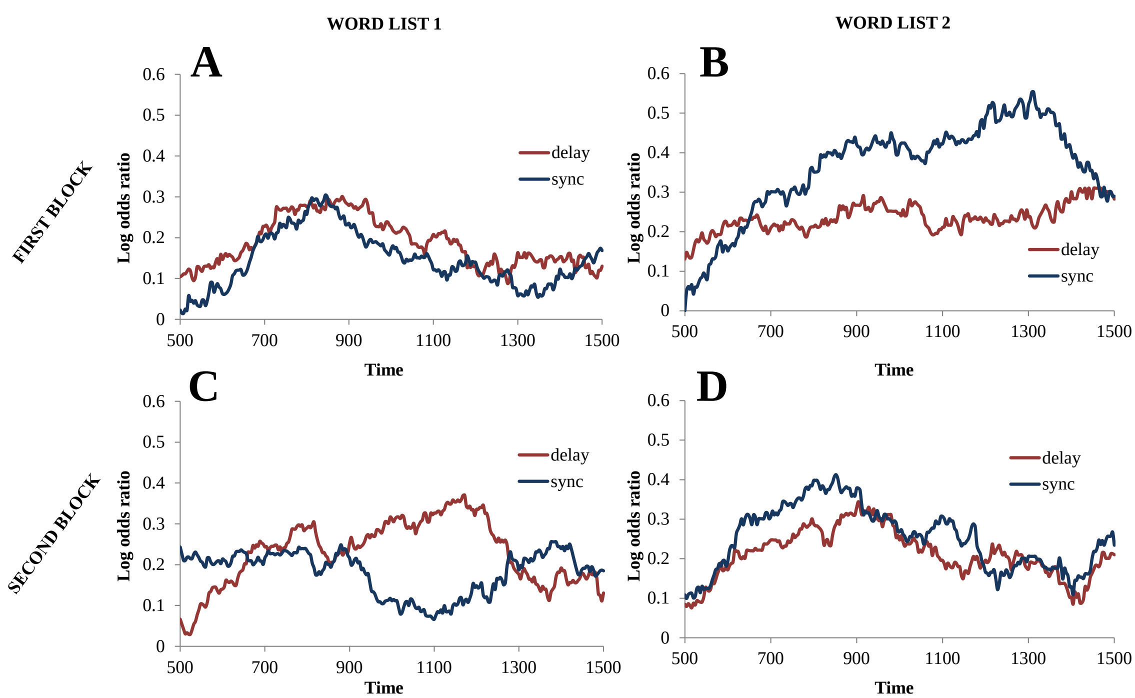

Figure 5-10: Experiment 4 – Log-odds-ratio of proportion of competitor looks to proportion of looks to average unrelated item across time, by training condition. A) First block for word list 1. B) First block for word list 2. C) Second block for word list 1. D) Second block for word list 2.

the analysis window are displayed in Figure 5-10; the mean log-odds-ratios used in the analysis are shown in Figure 5-11. Whereas *word list 1* appears to show little differentiation between training conditions for either block of testing trials (and perhaps a small reversal of the effect for later trials), *word list 2* shows a strong effect in the first block of testing trials that diminishes somewhat later in testing. These results were analyzed by adding testing block (contrast coded: -.5 for block 1, +.5 for block 2) to the mixed effects model used above. This model thus included training condition, word list and testing block as fixed factors; these factors were not correlated (all $R<.015$). Because there were fewer trials per cell, a slightly higher (though still small) number of stimuli were excluded from analysis (16 of 448 excluded). The *delay* condition contributed 10 of these cases. No looks were made to the unrelated objects for 10 cases, and none were

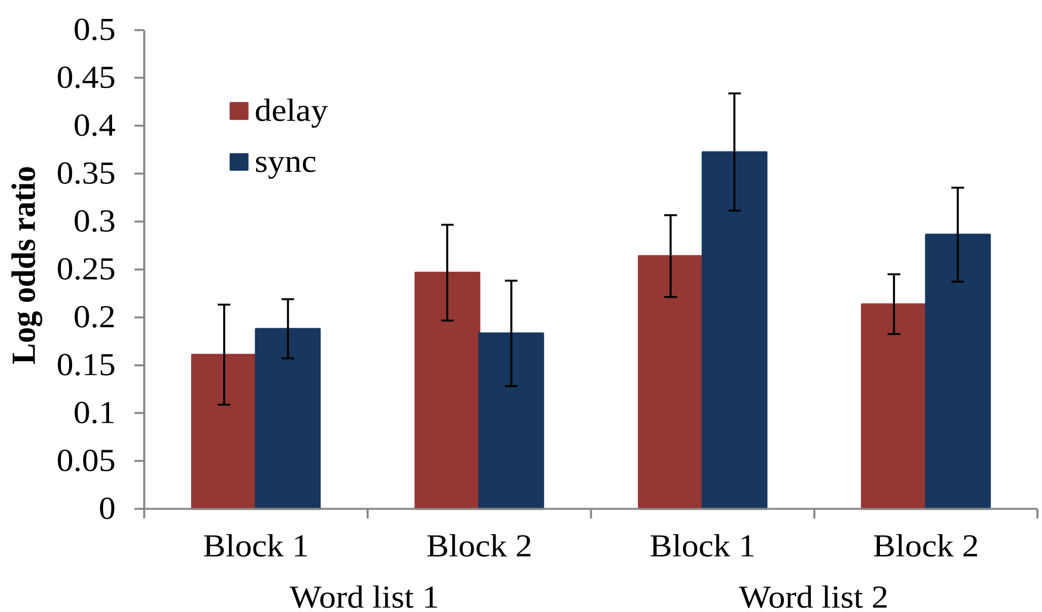

Figure 5-11: Experiment 4 – Log-odds-ratio within the 500-1500 ms analysis window, by training condition, word list and test block. Error bars represent standard error for that condition.

made to the competitor for the remaining six cases. No main effects and no interactions were significant in this analysis (Table 5-2).

Table 5-2: Results of statistical analysis of VWP trials in Experiment 4 by testing block.

Factor	*B*	*SE*	p_{mcmc}
Training condition	-.032	.038	.39
Word list	.085	.070	.23
Test block	-.016	.037	.67
Condition × list	-.11	.075	.13
Condition × block	.059	.073	.42
List × block	-.11	.073	.15
Condition × list × block	-.055	.15	.71

We next analyzed fixations to the target object (Figure 5-12). There appeared to be very little difference in target fixation between conditions. A mixed effects model was used to confirm this. We maintained the 500-1500 ms analysis window and used the empirical logit transform of the raw proportion of looks to the target during this window as the dependent variable. This model used a linear linking function, and it included training condition and word list as fixed factors, and participant and auditory word as random effects. The fixed factors were not correlated (R=-.011). This analysis confirmed that there were no differences in target fixation as a function of training condition (B=-.011, SE=.026, p_{mcmc}=.78) nor of word list (B=.012, SE=.077, p_{mcmc}=.90), and the interaction was not significant (B=-.027, SE=.053, p_{mcmc}=.61). Looks to the target were extremely consistent across training conditions and across the two word lists.

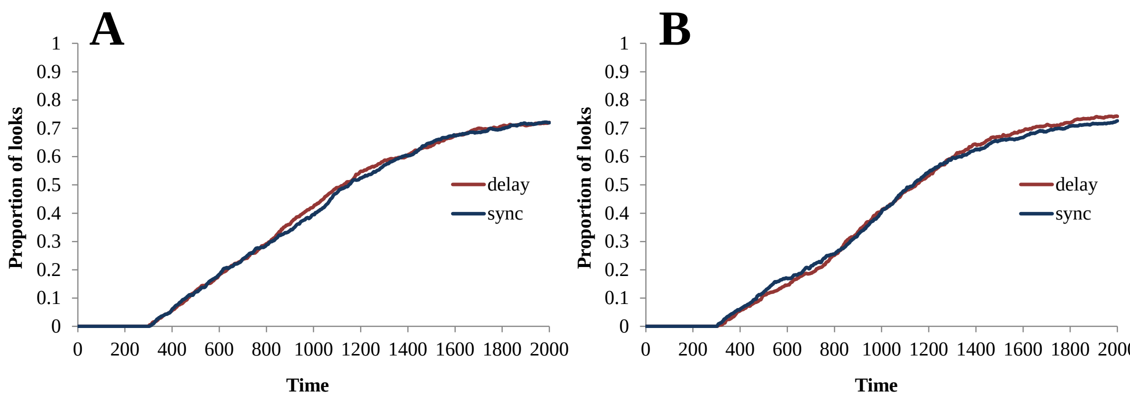

Figure 5-12: Experiment 4 – Proportion of looks to target items across time, by training condition. A) Word list 1. B) Word list 2.

5.4 Discussion

The use of a within-participants design in this experiment offered only weak support for lexical activation dynamics altering the course of word learning. For a subset of words used, auditory-visual synchrony led to increased interference from phonological competitors at test, as predicted by theories of continuous associative learning, although this increase was modest. However, the other word set showed no effect in either direction, with virtually no difference in the degree of competition seen for the two timing conditions.

The presence of increased competition in one word list, however, is encouraging. The separate word lists were created in part because some words produce weak competitor effects. *Word list 2* showed a trend toward greater overall competitor effects (although the main effect of word-list was not reliable), suggesting that perhaps one reason for the lack of effect in *word list 1* is weaker co-activation during learning.

It is unclear why there was so much variability in competitor effects between word sets. Both sets of words included substantial phonological overlap, with the initial four phonemes shared by the *onset competitor* pairs. There was no apparent difference in the speed of learning the two word sets, and they showed similar patterns of target fixation at test. *Word list 2* showed slightly increased incidence of failed learning of specific word pairs, although participants in the two training conditions were equally likely to show complete failure to learn all words in the study. One potential source is durational differences. Both word sets used in this experiment were longer than those used in Experiments 1-3 (*word set 1* mean: 837 ms; *word set 2* mean: 858). Additionally, the increased overlap between the words may have led to longer-lasting competitor effects. These factors may have led to ongoing coactivation into the visual presentation in the *delay* trials of this experiment. This would reduce the difference between conditions, masking spurious associations in the *synchronous* condition.

Another alternative is that yoking competitor pairs to never co-occur during training affected performance at test. Perhaps the overall zero association strength between all foils and the target word led participants to approach the VWP task differently; the lack of ambiguity may have caused them to be more certain in their target choice, suppressing competitor activations. Future work is needed to test these hypotheses.

Creating additional lists which mix together the words from the two word sets would allow a more thorough investigation of the differences between the word sets. Such a manipulation would demonstrate whether specific words are driving the loss of effect for *word list 1*, or whether instead it is the confluence of all words included in that set (or alternatively, if the words in *word list 2* somehow overinflated the effect). The final experiment in this dissertation incorporates this manipulation while also investigating alternative forms of learning. Future work should more directly investigate the difference between competitor effects for different word sets in a situation akin to this experiment.

The stronger competition effects in the between-participants manipulation than in the within-participants study may suggest that altering the timing during training exerts global changes on the way that new words are learned. When referents are frequently unavailable until late in the lexical activation process, the learner may adopt a global strategy of waiting to perform any encoding. In Experiment 4, half of the trials used the *delay* timing, which may have triggered a decrease in the use of immediate encoding. The smaller competition effect for *word list 2* suggests that such a decrement is not complete for all words. In order to more fully investigate global vs. local effects of stimulus timing on word learning, more studies are needed. Adjusting the relative number of *synchronous* and *delay* trials could elucidate how the statistics of timing across trials affects global learning strategies.

CHAPTER 6

EXPERIMENT 5: ALTERNATIVE FORMS OF LEARNING

The experiments to this point have demonstrated that under certain circumstances, a word learning task that encourages use of unsupervised learning leads to temporally-continuous learning during periods of real-time lexical competition. When learners are presented with auditory and visual information simultaneously and not given a chance for lexical competition to suppress competing word-forms after trials, they form associations with referents and competing word-forms that were activated in parallel. This offers evidence that, at least in some circumstances, unsupervised learning is continuous in time, without any monitor or thresholding mechanism to determine when competition is resolved to initiate learning. These results additionally suggest that adult learners at least sometimes use unsupervised associative mechanisms to learn novel words, and that they maintain multiple potential word-referent mappings for the same referent simultaneously.

Still unexplored is how ubiquitous temporally-continuous learning effects are when other forms of learning are encouraged. The previous experiments all relied on a purely passive learning task, in which participants simply looked and listened as word-referent pairings were displayed. Although they were told to try to learn the correspondences between specific words and referents, they did not have to make responses on most trials (although they did on the occasional interim trials), which may have encouraged them to rely more heavily on unsupervised mechanisms.

Even absent feedback, the simple act of making a response could conceivably affect learning. Response generation may affect the form that learning takes in several ways. First, when making a response, the learner marks a particular time as the appropriate time to perform learning. The participant could wait to build or update word-referent mappings until she makes a response, which is likely after most competition has resolved. Second, the act of forming the response itself requires fairly comprehensive stimulus processing, such that the participant can identify the referent in order to determine which response is most appropriate. This forced competition resolution could lead learning to have a stronger basis in post-competition lexical activation, thereby limiting the formation of spurious associations (as in McMurray et al, 2012). Finally, response generation may enable a completely different form of learning in some paradigms. Responses can be judged with feedback to determine whether they were based on effective associations. This feedback could be an external teaching signal identifying whether the response was correct, or it could be more of the prediction error variety, where responses are maintained across trials and judged at later events when the available stimuli either confirm or disconfirm the linkages that were used to make these responses (Roembke & McMurray, submitted).

Even as the response may shape learning in multiple ways, as a participant is generating a response continuous unsupervised learning may occur. As a result, while processing the visual and acoustic stimuli in order to form a response for the trial, the learner may be forming associations on the basis of active representations. Response generation may thus occur independently of learning. It is therefore an open question whether response generation affects the way that temporal processing dynamics interact with learning. Experiment 5 provides a detailed analysis of whether other learning tasks that require responses lead to evidence of continuous encoding during competition resolution.

6.1 Introduction

The preceding experiments rely on a learning task that highly emphasizes the use of passive, unsupervised learning. During the learning trials, participants were presented with multiple potential referents, forcing them to learn words through associations formed across trials. Because no responses were required on the primary learning trials, this was a fairly passive process; the learners were able to form word-referent mappings by passively tracking the co-occurrence statistics of word-forms and referents, and never had to generate a response to signify which of the possible referents the word referred to. This design was chosen because it provided the strongest opportunity to measure continuous encoding during unsupervised learning without any events to favor specific points in time (or processing). Indeed, the design was fairly successful in this goal, exhibiting strong evidence that learning happens continuously during processing for between-participants designs and more modest evidence in within-participants designs. These experiments have demonstrated that under some circumstances, word learning proceeds continuously in time, with mapping occurring during periods of parallel activation.

However, alternative forms of learning may alleviate this effect. Specifically, when the learning task signals a specific time to form mappings, learners may withhold learning until this signal occurs. Such a signal is readily apparent in a learning task that incorporates feedback; when the signal identifying the accuracy of the learner's choice is received, learning can proceed. Indeed, in error driven learning rules, mappings cannot be updated without both the response and the feedback, so no learning can take place until this point in time. As discussed in Chapters 1 and 2, this learning is likely to operate on post-competition representations, and it is therefore somewhat removed from the dynamics of lexical competition. Yet other signals within the trial could also operate as cues to learn after competition is resolved. When the task calls for a response even without feedback, the learner could base her mapping on the active representations when this response is given. These responses likely occur late in the lexical activation process, when most competitors have been suppressed. However, there remains ambiguity in the mapping, as no feedback signals whether the learner made an accurate response. Thus the learning remains unsupervised in some ways despite having a potential signal for when to initiate learning.

Experiment 5 investigates how response generation and feedback affect the interaction between word learning and the temporal dynamics of lexical activation. Participants in this experiment made responses on every learning trial signaling which referent they believe matches the auditory word-form they heard. Half of the participants received feedback on their selection. As in Experiment 4, stimulus timing was manipulated within-participants, with half the words trained with the *synchronous* timing of Experiments 1 and 4, and half trained with the *delay* timing of these experiments. This experiment allows differentiation of the representations formed when a response is given, and offers insight into whether feedback is required to eliminate continuous learning.

If response generation itself cause learners to delay learning until a response is made, neither participants receiving feedback nor those without feedback should show evidence of forming spurious associations. However, if the formation of a response on its own doesn't affect learning, the participants without feedback are predicted to show evidence of spurious associations, while those who receive feedback are predicted to show no such associations. It is not predicted that those learning with supervision will show spurious associations, as this form of learning provides information to update learned mappings only quite late in the lexical competition process (as described in Chapters 1 and 2).

6.2 Methods

6.2.1 Participants

Thirty-five participants from the University of Iowa community completed this experiment. Participants were paid $15 for their participation. All participants self-reported normal hearing and normal or corrected-to-normal vision. Three participants were excluded for low accuracy at test (below 75% on all four *onset competitor* pairs).

6.2.2 Design

The design mirrored that of Experiment 4 in many respects. Participants performed a phoneme-monitoring task before the word-referent training to become familiarized with the auditory word-forms. They then learned the words in a cross-situational learning task with two visual referents presented on each trial. The relative timing of the auditory and visual stimuli was manipulated within-participants, using the same timing as in Experiment 4; half of the words were learning with *synchronous* timing and half with *delay* timing. After learning, participants' word-referent mappings were tested using the VWP to determine whether the degree of interference was larger under conditions of *synchronous* training.

The training trials in Experiment 5 added a response component that did not occur in previous experiments (although it is quite similar to the interim trials in Experiment 4). On every trial, participants indicated which referent they believed matched the auditory word played. This response was made after the visual stimuli were removed from the screen; two open boxes replaced the images, and the participants clicked in the box where they believed the correct referent had occurred. Half of these participants received feedback indicating whether their choice was correct, consisting of a "ding" if they chose the correct referent and a "buzz" if they chose the incorrect referent. The other participants received no feedback. Because responses were made on every training trial, no interim trials were necessary in this experiment.

6.2.3 Stimuli

The stimuli were identical to those used in Experiment 4. However, given the quite significant differences between the two lists in Experiment 4, I constructed two additional word lists by swapping words between the two original word lists to ensure greater variety in the particular sets of words seen by any given participant (Table 6-1). These additional lists included slight overlap between some phonemes across competitor

Table 6-1: Phonetic transcription of additional word lists used in Experiment 5.

WORD SET 3				WORD SET 4			
Onset competitors		Offset competitors		Onset competitors		Offset competitors	
rubsɪf	rubsʌp	zɑimpæf	jempæf	ɵupsɛd	ɵupsiv	roikzɪb	boʊkzɪb
vimbɑim	vimbæl			bivdup	bivdʌf		
kælboʊm	kælboit	dævlik	hɛvlik	zeɪftɑd	zeɪftug	mɑgrʌs	dʌgrʌs
vɪsmeɪv	vɪsmɑk			lɪdzoʊv	lɪdoil		

pairs; this was unavoidable given the initial words. The same set of auditory stimuli from Experiment 4 were used, and these stimuli were distributed between tasks in the same way (10 exemplars used in word-referent training, five alternative exemplars use in VWP testing).

6.2.4 Procedure

Many aspects of the procedure of Experiment 5 were identical to that of Experiment 4. The pre-exposure phase was identical, except that the error which prevented two of the *offset competitor* words from being presented during pre-exposure was fixed so that all 12 words were heard eight times each. This increased the number of trials from 80 to 96. The VWP was identical in all respects to Experiment 4. However, the word-referent training differed substantially for Experiment 5.

Participants initiated word-referent training trials by clicking on the blue dot in the center of the screen. After 100 ms, the auditory stimulus was played. For the *synchronous* items, the two visual referents were displayed simultaneously with the auditory stimulus; in the *delay* condition, 1000 ms elapsed between the onset of the auditory stimulus and the display of the referents. The alternative referent in this experiment was selected as in Experiment 4; it was never the referent of the phonological competitor, and it was never the referent of either member of the yoked competitor pair for the target. In both conditions, the visual images remained on the screen for 800 ms. After visual stimulus offset, the images were removed and open black boxes were displayed surrounding the locations were the referents had been displayed. The mouse cursor was made visible simultaneously to these boxes, and participants were told to click in the box that contained the picture they think the auditory stimulus identified. Only mouse clicks after the black boxes appeared were registered.

In both the *active unsupervised* (no feedback) and the *supervised* (feedback) condition, after the participant clicked within one of the boxes, 300 ms elapsed and then the screen went blank. In only the *supervised* condition, auditory feedback was provided immediately after the click during this intervening period. This feedback consisted of a positive “ding” sound if the participant made the correct choice and a negative “buzz” sound if they chose incorrectly. For both conditions, after this 300 ms period (whether or not feedback occurred), a further 250 ms inter-trial interval separated the trials; this equated the total trial duration to that of the passive learning in the previous experiments.

The nature of the learning task was predicted to greatly accelerate learning speed; in recent cross-situational learning studies using highly similar stimuli, Roembke and McMurray (submitted) found that learning was quite rapid when participants made a response on every trial, even without feedback. As such, this experiment only included 20 blocks of trials (12 trials per block) rather than the 32 blocks used in previous experiments. Thus there were 240 training trials presented during word-referent training.

6.3 Results

6.3.1 Exclusions

As in Experiment 4, only *onset competitors* were included in the bulk of analyses here (both word-types were included in analysis of the pre-exposure data). Within these *onset competitors*, exclusion criteria for Experiment 5 were identical to those used in Experiment 4. For each participant, the accuracy of choosing the correct target during the

VWP testing trials was examined for each word pair. Any pair whose accuracy was below 75% was excluded for that participant. For three participants, all four *onset competitor* pairs fell below this threshold, so these participants were completely excluded from analysis. Two of these participants were in the *supervised* condition. For the remaining 32 participants, there were 128 possible pairs of *onset competitors* (32 participants × four pairs). Of these, 101 had accuracy that exceeded the threshold for inclusion in analysis. The distribution of excluded pairs is displayed in Figure 6-1. Most of these pairs were in the *supervised* condition (19 of the 28); combined with the two participants dropped completely from analysis, this suggests that feedback may have impaired learning. Words in *wordlist 2* were the most likely to be learned poorly (12 pairs in this list did not reach criterion); *wordlist 1* was the most accurately identified (only one excluded pair). All wordpairs were excluded for at least one participant; however, the pair *lidzoil* and *lidzove* was excluded more than any other pairs (seven exclusions; no other word had more than four). Nearly every included participant had at least one training pair exceed the accuracy threshold in each training condition. Of the *supervised* participants, one had no pairs for *synchronous* and one had no pairs for *delay*. Of the *active unsupervised* participants, two had no pairs for *delay*. The remaining 28 participants had included data in both conditions.

6.3.2 Pre-exposure

Overall, participants performed quite well during the pre-exposure task. Participants correctly identified the words that had a “V” sound in 94% of trials, and showed a false alarm rate on trials without a “V” sound of around 11%. These error rates are well in line with those of previous experiments (e.g. Experiment 1: 93% correct recognition; 10% false alarm rate).

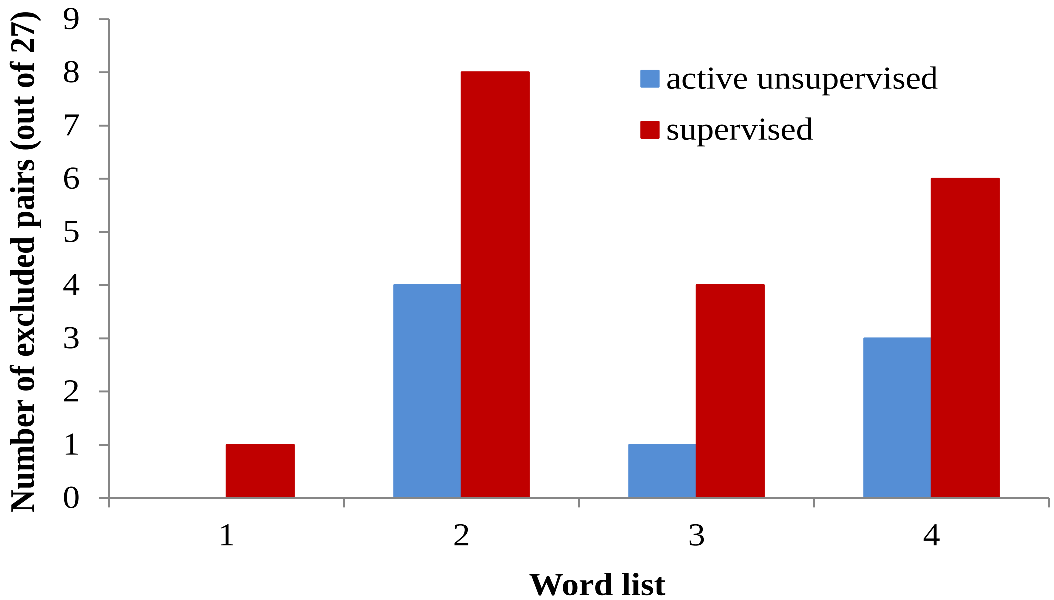

Figure 6-1: Experiment 5 – Number of word pairs excluded from analysis due to low accuracy at VWP test, by word list and condition. Across all conditions, 27 pairs were excluded.

6.3.3 Word-referent training

Experiment 5 allowed tracking of learning throughout word-referent training, as participants made a response on every trial. Thus the training data serve as a measure of the trajectory of learning, much as the interim trials offered this in previous experiments. In these analyses, as in those for the interim trials of Experiment 4, we consider separately those word pairs that were learned effectively and those that were not learned at the time of VWP testing (Figure 6-2).

The words that showed accurate performance at test were learned quite effectively. As in the previous experiment, those words that had poor accuracy during the 4AFC VWP trials also showed quite high performance during 2AFC training. However, these words were learned more slowly and did not reach ceiling (Figure 6-2). Because there were very few word pairs that did not meet the criteria for acceptance, no statistical analyses were conducted on these data.

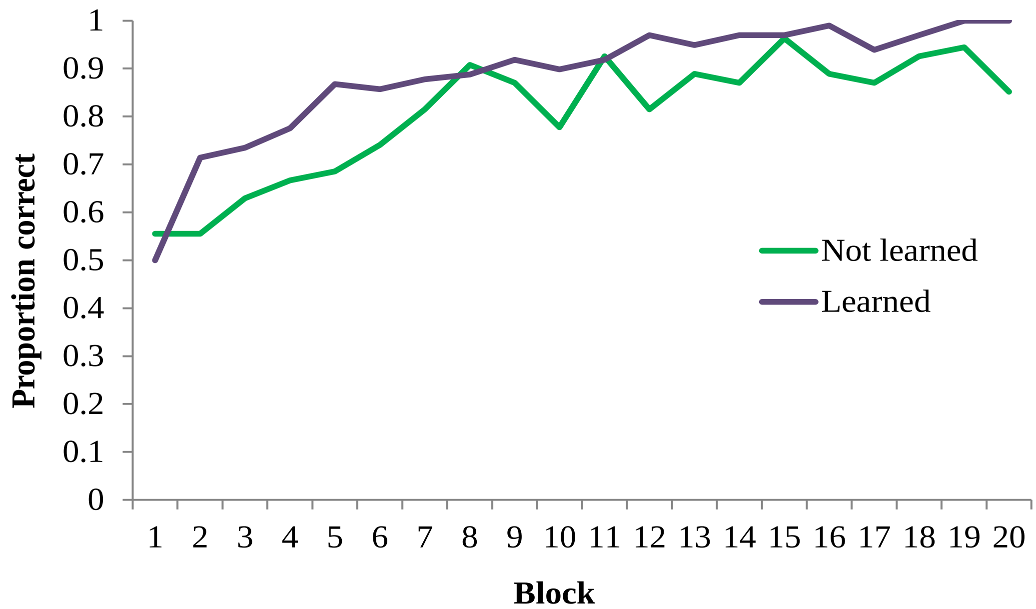

Figure 6-2: Experiment 5 – Accuracy of responses during training trials for Experiment 5 words that were accurately identified at test and those that were poorly identified at test.

Analyses on these data were conducted in two ways. First, all word pairs with below-criterion accuracy in the VWP trials were excluded to compare learning only for those words that were included in the analyses for the other tasks. Second, all word pairs were included to determine how learning differed as a function of timing- and feedback-condition, independently of later performance.

For both analyses, learning was considered as a function of both training-condition (*synchronous* or *delay*) and feedback-condition (*active unsupervised* or *supervised*; Figure 6-3A and B excluding low-performance pairs; Figure 6-3C and D including all word pairs). Participants in all cells reached high levels of performance during training. These data were analyzed using two separate mixed effects models for the two ways of analyzing the data. Each included training-condition (contrast coded: *synchronous*: -.5; *delay*: +.5) and feedback-condition (contrast coded: *active unsupervised*: -.5; *supervised*: +.5) as fixed factors, and participant, word-list and

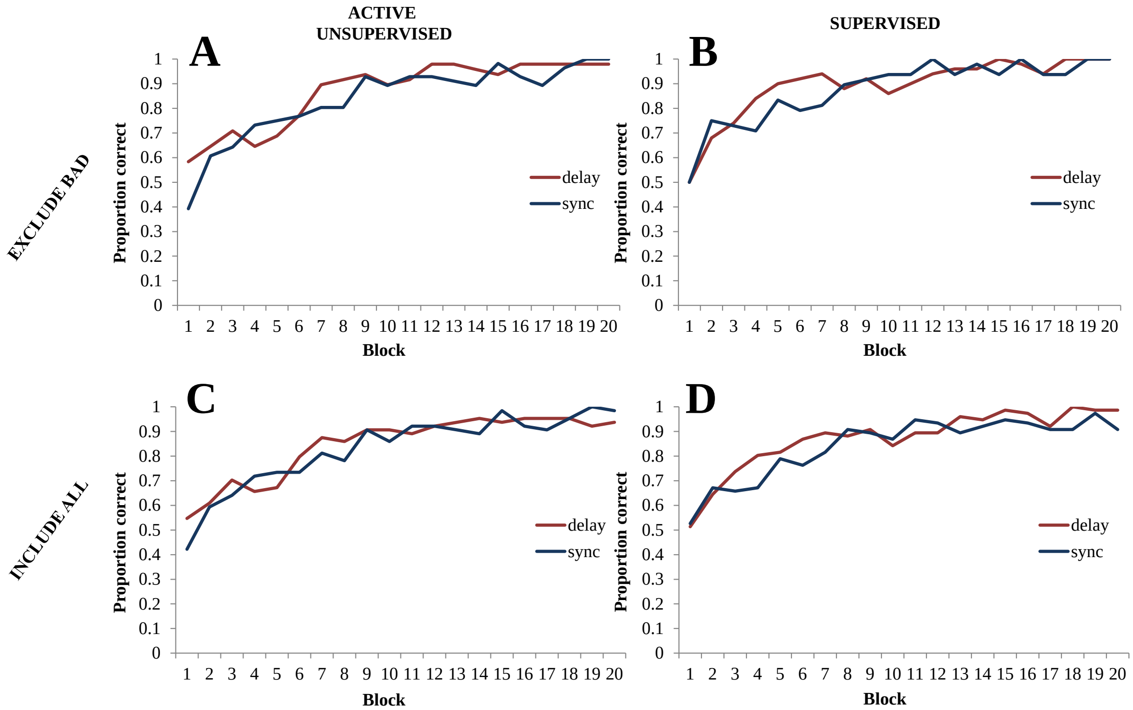

Figure 6-3: Experiment 5 – Accuracy of responses during word-referent training trials across blocks, by training-condition and feedback-condition. A) Active unsupervised, excluding low accuracy pairs. B) Supervised, excluding low accuracy pairs. C) Active unsupervised, including all pairs. D) Supervised, including all pairs.

auditory-word as random intercepts. No random slopes were used as adding random slopes of timing condition by participants did not improve fit for analysis of the eye-tracking data (by χ^2 test, p=.93)[1]. The fixed factors were not correlated (all R<.03). The results of these analyses are displayed in Table 6-2 for analysis excluding poorly-learned pairs, and Table 6-3 including all pairs.

In the analysis excluding the poorly-learned pairs, the main effect of block was significant, as participants improved on later blocks. No other main effects or interactions approached significance, signaling that all words were learned approximately as quickly as one another. However, when including all word pairs, block was significant, but there were also significant interactions of feedback-condition × training-condition and

1 We maintain this model structure throughout all analyses of this experiment for consistency.

Table 6-2: Results of statistical analysis of word-referent training trials in Experiment 5, excluding word pairs below accuracy criterion at VWP test.

Factor	*B*	*SE*	*z*	*p*
Timing condition	.091	.15	.63	.53
Feedback Condition	.39	.26	1.53	.13
Block	.22	.0091	24.0	<.0001
Feedback × Timing	-.26	.28	-.90	.37
Timing × block	-.0027	.018	-.15	.88
Feedback × block	-.013	.018	-.69	.49
Timing × feedback × block	.042	.036	1.18	.24

Table 6-3: Results of statistical analysis of word-referent training trials in Experiment 5, including all word pairs.

Factor	*B*	*SE*	*z*	*p*
Timing condition	.14	.12	1.10	.27
Feedback Condition	.42	.27	1.59	.11
Block	.19	.0072	26.7	<.0001
Feedback × Timing	-.48	.24	-2.01	.044
Timing × block	-.0002	.014	-.012	.99
Feedback × block	-.016	.014	-1.09	.27
Timing × feedback × block	.082	.028	2.90	.0038

feedback-condition × training-condition × block. Simple effects analyses investigated this latter interaction by comparing training-condition and block separately in each feedback-condition (using identical model structure, but without feedback-condition as a factor). The model for the *active unsupervised* participants showed a significant effect of training-condition (B=.40, SE=.18, Z=2.27, p=.023), with better performance for *delay* words, and a significant effect of block (B=.20, SE=.010, Z=19.37, p<.0001). The interaction was also significant (B=-.040, SE=.020, Z=-1.98, p=.048). Inspection of the data across blocks suggests that this interaction may have emerged from poorer performance for *synchronous* words in early blocks but equivalent performance in later blocks. The *supervised* condition, meanwhile, showed no main effect of training-condition (B=-.12, SE=.17, Z=-.67, p=.50). The effect of block was significant (B=.18, SE=.010, Z=18.38, p<.0001). The interaction was also significant (B=.041, SE=.020, Z=2.06, p=.039). However, the interaction was in the opposite direction as that of the *active unsupervised* condition. The participants in the *supervised* condition appeared to learn the *delay* words slightly faster.

6.3.4 VWP testing trials

As in the analysis of Experiment 4, VWP analyses focused on *onset competitor* trials, as there were insufficient *offset competitors* to conduct an appropriate analysis of performance. Analyses of the VWP data only used words that exceeded the accuracy criteria detailed in ***6.3.1***.

6.3.4.1 Accuracy

Overall accuracy on the VWP trials was high for all words (M=94.9% correct). This level of accuracy was consistent across the different timing and feedback conditions (Figure 6-4). These data were analyzed using a mixed effects model (binomial linking function) with training-condition and feedback-condition (contrast coded) as fixed factors, and participant, auditory word, and word list as random intercepts. The fixed

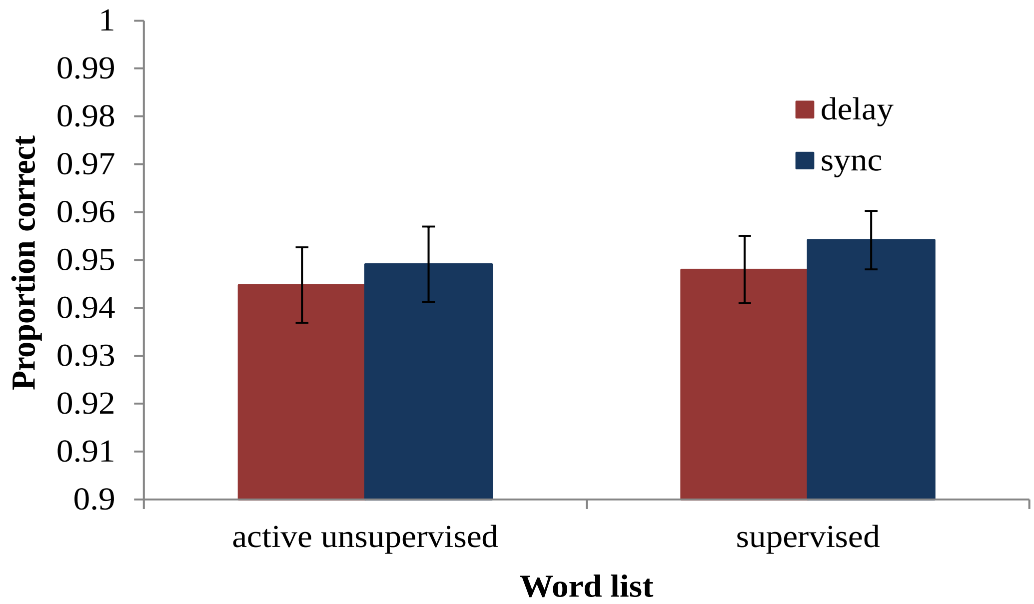

Figure 6-4: Experiment 5 – Accuracy for included word pairs in the VWP trials of Experiment 5, by training-condition and feedback-condition. Error bars represent standard error for that condition.

factors were not correlated (*R*=-.084). This analysis did not reveal any difference between training-conditions (*B*=-.041, *SE*=.029, *Z*=-1.41, *p*=.16) nor of feedback-conditions (*B*=.26, *SE*=.37, *Z*=.70, *p*=.49), and the interaction was not significant (*B*=.020, *SE*=.057, *Z*=.36, *p*=.72). Thus accuracy was consistent across the different comparisons.

6.3.4.2 Eye movements

As in previous experiments, eye movements were analyzed by considering looks to the competitor relative to looks to the average unrelated item. This was accomplished using log-odds-ratios (Figure 6-5 across time; Figure 6-6 averaged across the analysis window). These analyses used the same window as in previous experiments (500-1500 ms). The average log-odds-ratio (as seen in Figure 6-6) was used as the DV in the analysis. These data were entered into a mixed effects model (linear linking function) with training-condition and feedback-condition as fixed factors (both contrast coded) and

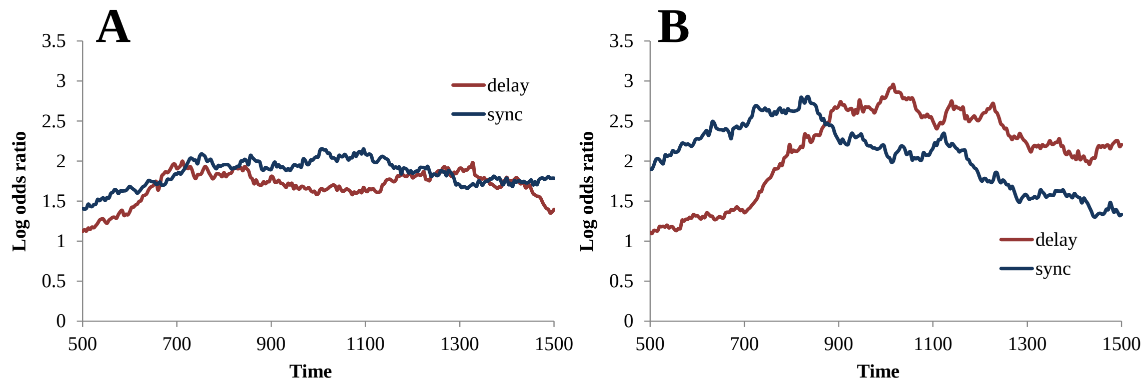

Figure 6-5: Experiment 5 – Log-odds-ratio of proportion of competitor looks to proportion of looks to average unrelated item across time, by timing condition. A) Active unsupervised participants. B) Supervised participants.

participant, auditory word and word list as random intercepts. The fixed factors were not correlated (R=-.076). MCMC simulations were used to estimate significance values. This analysis did not reveal a main effect of training-condition (B=-.034, SE=.045, p_{mcmc}=.53) nor of feedback-condition (B=.078, SE=.063, p_{mcmc}=.20). The interaction also was not significant (B=.067, SE=.089, p_{mcmc}=.50). Inspection of the data suggests a trend toward an effect for the *active unsupervised* group; to investigate this trend, simple effects were conducted using only the *active unsupervised* data and including only training condition as a factor. All other model features were kept as above. The effect of training-condition was again non-significant (B=-.063, SE=.065, p_{mcmc}=.38). Thus there appeared to be no difference in the degree of competitor effects between the timing conditions for either those in the *supervised* or in the *active unsupervised* conditions.

Target looks were analyzed as in previous experiments, using an empirical logit transform of the raw target looks in the different conditions as the DV. Target looks were considered as a function of training-condition and feedback-condition (Figure 6-7). These data were entered into a mixed effects model of the same structure as the analysis of competitor looks. The fixed effects were not correlated (R=-.043). This analysis revealed no difference in target looks as a function of training-condition (B=-.020, SE=.031,

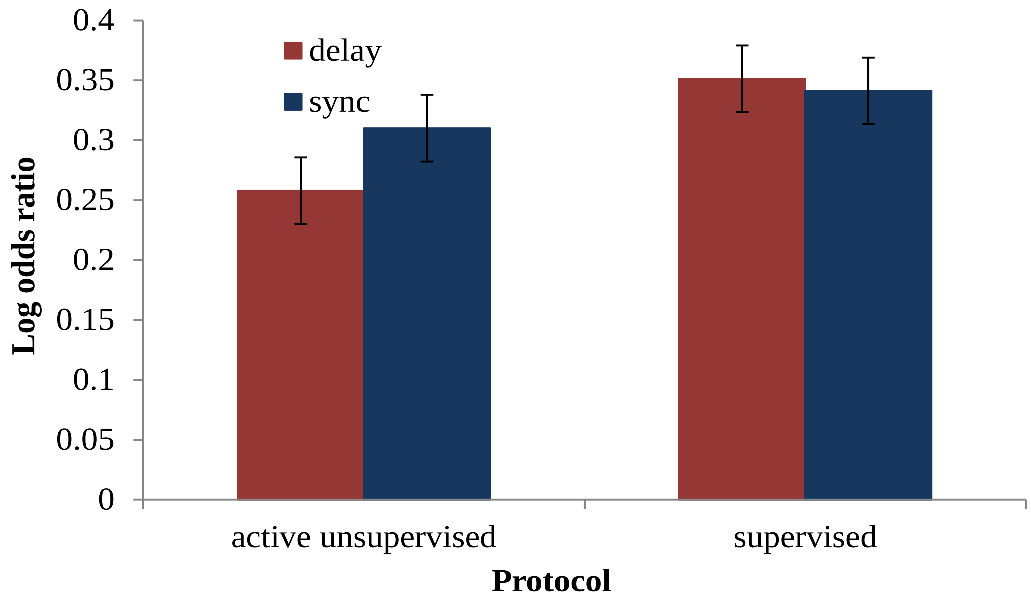

Figure 6-6: Experiment 5 – Log-odds-ratio within the 500-1500 ms analysis window by training-condition and feedback-condition. Error bars represent standard error for that condition.

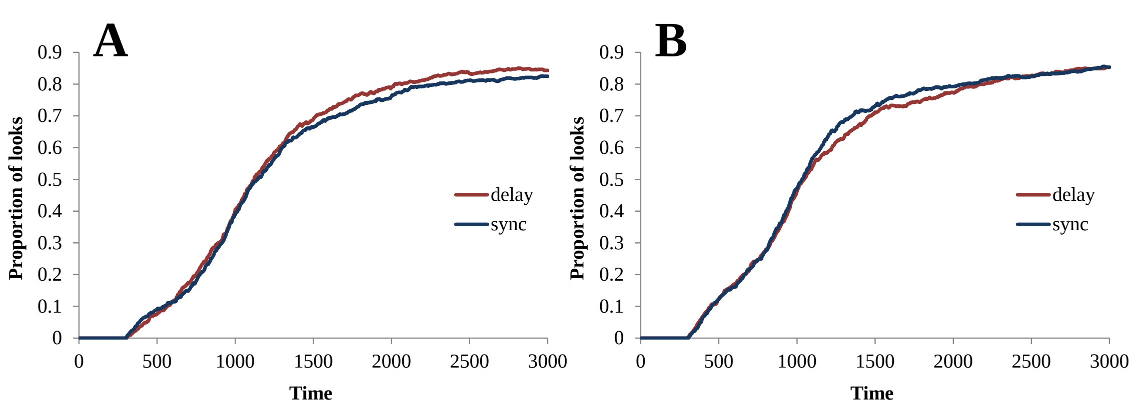

Figure 6-7: Experiment 5 – Proportion of looks to target item across time, by training-condition. A) Active unsupervised participants. B) Supervised participants.

p_{mcmc}=.68). There was a marginally-significant of feedback-condition (*B*=.10, *SE*=.078, p_{mcmc}=.064), as participants in the *supervised* group showed slightly more fixations to the target items. The interaction was not significant (*B*=-.067, *SE*=.061, p_{mcmc}=.32). Despite the slight increase in target looks for the *supervised* group, training condition never had an effect on degree of target looks.

6.4 Discussion

Experiment 5 set out to determine whether generating a response alters the timing of word learning. Specifically, response generation might delay the time at which participants update word-referent mappings, eliminating formation of such mappings during periods of lexical competition. The experiment simultaneously explored how feedback affects this process; although all participants in this experiment made responses on every word-referent training trial, only half of them received feedback. A feedback signal offers the opportunity for classic supervised learning to occur, which is predicted to eliminate learning during lexical competition; without feedback, learning could occur earlier. In both cases, no evidence was found for continuous learning during lexical processing; for both feedback conditions, the difference in competition (on the VWP) as a function of timing (during training). This suggests that some aspect of response generation causes participants to delay their learning until after competition resolves[2].

Response generation may have affected learning as it provided learners with a specific time to update their representations. The learner must resolve competition well enough to select one of the two available referents on the display. This forced

[2] To further analyze the effect that response generation had on learning, an analysis was conducted comparing the *active unsupervised* participants in this experiment to the participants in Experiment 4, who learned the same words with the same timing conditions, but did not make responses. This analysis revealed no significant effects, suggesting that responding did not lead to different learning. However, this finding may be a result of insufficient power, as Experiment 4 showed some evidence of spurious associations, whereas Experiment 5 showed no such evidence. Additional participants need to be run in these experiments and further studies need to be done to more fully assess the effect of adding a response.

competition resolution may result in learners adjusting their learning to be more akin to supervised situations, when learning waits for an error signal. In this case, although no error signal is given in the *active unsupervised* condition, they may still wait to learn until competition has settled well enough to make a decision. Associations would thus be formed on the basis of post-competition activations, rather than learning while competition was resolving.

However, the results of this study call for more thorough investigation in the future. A (non-significant) trend toward an effect of training condition emerged in the *active unsupervised* condition, suggesting that response generation in this condition may have less of an effect on learning; participants may still conduct some continuous learning, but then continue updating these learned mappings as a response is made. Future studies can help determine whether this trend is noise in the data, or whether the experiment was insufficiently powerful to detect the effect. Additionally, the words used in this study were the same used in Experiment 4, some of which showed only weak evidence of continuous learning even when no response was required. As discussed in Chapter 5, the longer duration of these stimuli and the more protracted overlap between competitors may cause coactivation to last well into the visual presentation even for *delay* presentation formats. Studies that more precisely control the timing relative to stimulus duration could help determine whether a slower delay condition shows smaller competition effects.

A further concern is that response generation may not be causing the decrease in spurious associations, but instead the additional processing time given participants to make responses causes it. In order to ensure that participants were not under undue time pressures in the *synchronous* trials, responses were always self-paced. During this delay, participants may continue updating representations, as in Experiment 2. As such, the response may have been irrelevant to the lack of spurious associations; instead, the additional processing may have caused it. Creating a task where responses are more rapidly generated without adding artificial time pressures for *synchronous* trials is critical to determining whether the response itself changed learning. Having participants respond with a button press rather than clicking on the display may help speed the responses, eliminating some of the additional processing time that can occur when generating a response.

Another possibility is that the decreased number of training trials in this experiment reduced the degree of spurious associations. Adding responses was predicted to improve learning overall, and participants in this experiment showed overall comparable learning performance to those in previous experiments. While the number of word-referent training trials was sufficient for participants to learn the words, it may not have been sufficient to strengthen associations with competitors to a point where the VWP could detect them. Although participants learned the words effectively, they may have continued to form and update associations with additional training.

Experiment 5 thus provides modest initial support that response generation alters the way that lexical processes interact with word learning. Participants making responses showed very little evidence of generating spurious associations; however, it is as yet unclear what aspect of response generation blocks these associations.

CHAPTER 7

GENERAL DISCUSSION

7.1 Summary of results

In a series of studies, this dissertation uncovered evidence that unsupervised learning occurs continuously in time, with learners forming associations whenever representations of stimuli are active. Using word learning as a model test case, these studies found that situations that emphasize learning during periods of lexical competition lead to spurious associations with the phonological competitors to the target word that are partially active. This argues that learners are forming associations while multiple word-forms are active rather than waiting for competition to resolve.

This effect was most pronounced when all words were trained with a similar stimulus timing that encourages the formation of parallel associations. In Experiment 1, participants who learned a set of words such that all words were presented synchronously with potential referents showed evidence of spurious associations; these participants at test showed greater consideration of the referents of competitor word-forms than would be expected from online competition, signaling an association between the word-form and the competitor referents. Meanwhile, participants who learned words such that the entire word was heard before any referents were presented showed no such evidence; these participants showed reduced competition relative to those in the *synchronous* condition. Experiment 4 showed modest support for such effects when a within-participants design was used. For one of the word lists used in this study, learners showed increased competition at test for words trained with *synchronous* timing; the other list showed no such effect. While it is unclear what conditions determine whether such effects emerge, the presence of increased competition for some words demonstrates that continuous encoding within-participants is possible in some circumstances. Although this effect is somewhat limited, it does demonstrate that unsupervised learning can occur in such learning tasks.

This experiment demonstrates that learning that is continuous in time leads to co-active word-forms becoming associated with a given word's referent. However, additional processing time after word offset can diminish these associations, at least for some forms of competitors. In Experiment 2, participants learned words with the same auditory-visual synchrony of the *synchronous* condition of Experiment 1, which showed evidence of spurious associations. However, Experiment 2 provided an additonal delay between trials during which participants could continue updating the activation states of the words/objects, as well as their learned associations. Although no further auditory or visual information was presented during this delay, participants in this experiment showed no evidence of spurious associations for *onset competitors*; associations formed with *offset competitors* were found. This blank delay between trials thus altered the associations that formed, such that no links with words that overlapped at onset were observed. As participants continue to process the stimuli, they appear to continue to update their learned associations; for *onset competitors*, this leads to an elimination of spurious associations as these types of competitors are suppressed. However, the later activation for *offset competitors* may require even further processing time to suppress spurious associations formed during parallel activation (or they may not suppress at all, as no further information arrives to signal that these activations are inaccurate). Experiment 2 thus showed that continuous learning extends even after the stimuli are removed, as learned associations are updated throughout processing.

Yet the dynamics of stimulus processing are not isolated to the activation of the auditory word-form. Visual information takes time to process (e.g., Treisman & Gelade, 1980; Wolfe, 1994), and visual processing effects can cascade to affect phonological processes (Chen & Mirman, 2012). Experiment 3 seems to show evidence of such interactions between visual and auditory activation. Participants were presented with visual referents before the onset of auditory information. This presentation format ensured that referents were on the display while lexical processing was ongoing, but it also provided a chance for learners to form predictions about which words could occur. When the two referents appeared before the auditory stimulus, learners could activate the names of each referent, and use these names to narrow the co-activation from competing names. This would reduce the formation of parallel associations, as participants could begin suppressing competitors even before auditory information is received.

Finally, these increased competitor effects appear to arise primarily in learning tasks that emphasize the use of unsupervised learning mechanisms. In Experiment 5, participants received similar trial structures to those in Experiment 4, but were forced to make a response on every trial. Whether or not participants received feedback indicating whether their responses were correct, they showed little evidence of forming spurious associations (though there was a trend in the direction of spurious associations in the *no feedback* case). This suggests that the process of generating a response alters when learning occurs; when participants are forced to fully process the stimuli in order to respond, they may delay their learning until the response is made. However, such effects may occur simply because response generation increases trial duration; it may be the added time to process while generating the response that causes this difference, rather than actual changes in the form of learning.

These experiments thus show a complex relationship between stimulus processing and learning. Certain learning paradigms allow associations to be built throughout processing; however, the representations formed during this learning are updated with ongoing processing. Other forms of learning appear to eliminate temporally-continuous learning, and instead focus learning on the end-product of competition processes. Thus learning, even just within the domain of word learning, is quite complex, with numerous sources of information (from the activations of both auditory and visual stimuli as well as from the nature of the learning task) simultaneously impacting the way that the learner forms associations.

7.1.1 From parallel activations to parallel associations

As listeners hear auditory stimuli, they simultaneously activate several words that are consistent with the acoustic signal (Allopenna et al., 1998; Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, 1987). These words compete for recognition; the competition processes extend across time, with listeners sometimes needing more than the entire word to identify the correct word-form (Bard et al., 1988; Grosjean, 1985; McMurray et al., 2009). While these words are simultaneously active, the listeners can associate the word-forms with available referents; although competition is ongoing, learners appear to start forming word-referent mappings immediately. Not only is word recognition incremental, but as word learning is continuously coupled to these dynamic processes, it comes to reflect this incrementality.

Parallel activation of word-forms thus leads to parallel associations with referents if the referents are present while competition is ongoing. However, this learning process is more complex than simply tracking stimulus co-occurrence. Instead, learning continues throughout *processing*, not just throughout stimulus presentation. Ongoing processing dynamics continue to affect the learned representations, with mappings updating beyond what was learned during stimulus presentation.

This ongoing updating forces a reconsideration of how learning proceeds. Whereas the intuitive conception of learning is a single update per individual learning event, some forms of learning are better conceptualized as repeated changes in the mapping across time within an event. Such continuous learning integrates information immediately as it comes in, but then continues to use additional information to refine the learned mappings. Although some forms of learning, such as those reliant on an error signal that occurs at a prescribed time, implement learning that occurs as a single update to weights or connections, unsupervised learning includes repeated updating. This is more akin to a recurrent network that learns in many cycles as it processes stimuli (McMurray et al., 2012; Thelen, Schoner, Scheier, & Smith, 2001). This changes the idea of a learning event; learning is a protracted process that extends across the event, and is dynamically influenced by the state of activation of the stimuli.

7.2 Limitations and future directions

The experiments in this dissertation offer intriguing evidence that learning is continuous in time during lexical processing, however they also leave many questions to be answered in future research. The conditions under which spurious associations were found were somewhat limited; they were more apparent when timing was manipulated between-participants and no additional processing time for either visual or auditory stimuli was available. The variability of the effect in the within-participants design is particularly vexing, as it suggests that some aspects of specific stimuli may affect the form that learning takes. Future research should investigate what characteristics of individual stimuli leads to variation in evidence of spurious associations. Perhaps the degree of overlap or the duration of certain stimuli leads to less differentiation in the associations formed in the two timing conditions. Similarly, it is important to investigate how different trial structures alter the learning process, such as different forms of responses or different types of feedback.

7.2.1 Learning based on internal representations

The alternative timing conditions examined in Experiments 2 and 3 lead to additional questions about the nature of the interactions between visual and auditory stimulus processing, and how these processes affect learning. These experiments demonstrated that learning is not confined to times when stimuli are physically present. Instead, learning occurs whenever some mental representations of the stimuli are available to update the mappings between them. Learning is thus not a single-pass event, but a continuous process of forming and updating the links between stimuli even after the perceptual events have passed. As competition continues to resolve, learners can continue to augment their mappings. However, it remains to be seen how protracted this process is. Experiment 2 showed that adding a delay at the end of auditory-visual presentation eliminated spurious associations with *onset competitors*, yet it did not affect those formed with *offset competitors*. It is possible that additional time might lead to elimination of these associations as well; alternatively, because these associations form at the end of words, with no successive information overruling them, they may endure regardless.

Related to questions of ongoing learning are issues of sleep-based consolidation. When learning words, sleeping seems to help these words become more embedded in the lexicon (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003), although sleep is not required for at least some forms of consolidation (Kapnoula, Gupta, Packard, & McMurray, in press; Lindsay & Gaskell, 2013); such effects signal ongoing learning about words well after stimulus presentation. It is unclear what effect consolidation would have on the weak spurious associations formed during parallel activation. These associations may become more embedded in lexical knowledge, showing lasting effects on processing. Alternatively, consolidation might prune such weak associations, helping the learner overcome the inappropriate learned associations.

Experiment 3 showed that when learners are given a preview period to see the pictures before learning, spurious associations are eliminated. These results could signal that longer trials overall block the formation of spurious associations, but as described in the discussion for that experiment, this explanation may be untenable, as few changes in learning “strategies” seem to predict the results of both Experiments 2 and 3. Instead, it was suggested that these results imply effects of prediction of the upcoming word-form blocking competitor words from becoming activated. While a viable explanation, significant additional work is necessary to determine how such predictions affect learning. There is evidence that even when context constrains the likely upcoming words, listeners still activate phonological competitors (Zwitserlood, 1989), suggesting that the pre-exposure to the referents may not alter coactivation of competing word-forms. However, other methods show that listeners readily use verb information to constrain which the range of words that are considered during lexical activation (Dahan & Tanenhaus, 2004). Similarly, listeners use knowledge of the world to bias eye-movements toward likely referents of a sentence (e.g. when hearing “the man will drink the…” participants are more likely to fixate *beer* than *wine*; Altmann & Kamide, 2007).

Similarly debates about the capacity for context to constrain stimulus processing arise regarding the VWP as a tool for language research. There is concern over whether providing a preview period when participants can pre-name the objects leads them to rely on a less natural mode of language processing. These debates are quite active, with many researchers expressing confidence that pre-naming does not narrow the scope of lexical competition. It thus remains an open question to what extent the learners can use the pre-exposure to narrow the parallel activation that occurs during word recognition. Further studies thus are essential to detail why learning changes when given a preview period.

7.2.2 Forms of learning

Experiment 5 offers some initial evidence that response generation affects the time at which associations are formed. When participants responded, regardless of whether or not feedback was present, no reliable difference was seen between *synchronous* and *delay* presentation. However, much additional work needs to be done to better understand these results. The failure to find effects in this study might arise from variability between words included in the study; as described in the discussion of Experiment 4, durational differences may have led to differences in words’ capacity to show spurious associations. Additionally, the inclusion of a response affects the stimulus timing in ways that may alter learning for reasons other than the response itself. In order to ensure that the response was not adding undue time pressures on participants in the *synchronous* condition, it was necessary to allow participants sufficient time after the trial to make their responses. However, this time also may provide additional processing time, as in Experiment 2. This could lead to continued learning during the interim period that eliminates spurious associations independently of the response. This suggests that investigating *offset competitors* rather than *onset competitors* in these studies may have shown more pronounced evidence of spurious associations in this experiment. Other methods that incorporate response generation without additional processing time could also more fully investigate whether response generation is having a direct effect on the form that learning takes.

Considering word learning more generally, the experiments in this dissertation demonstrate some use of unsupervised associative mechanisms by adults to learn new words. While this contrasts with theories that adults never use such mechanisms (Medina et al., 2011; Trueswell et al., 2013), it does not specify whether this is the typical form of word learning engaged in by adults. Work within the cross-situational word learning paradigm is designed to more closely simulate natural learning contexts (Yu & Smith, 2007), and recent work within this paradigm suggests associative (rather than hypothesis-testing) learning in this task (Roembke & McMurray, submitted). However, other word learning situations may elicit quite different forms of learning, which may show distinct forms of interaction with lexical activation dynamics. For example, words can be learned through ostensive definition, where the learner is explicitly told the word-referent mapping. Such learning may obviate the need to use unsupervised learning mechanisms, and thereby discourage continuous learning. Although Experiment 5 began exploring alternative forms of learning, more comprehensive investigations into learning in other paradigms can offer a more complete picture of when encoding occurs in different word learning situations.

7.2.3 Processing dynamics and learning outside language

This dissertation used word learning as test case for understanding how learning interacts with perceptual processing more generally. Word learning was selected because the temporal dynamics of processing are quite well understood, and because of the inherently temporal nature of spoken language. These characteristics allowed the generation of precise predictions and the development of methodologies to test such predictions. However, other domains have more opaque processing dynamics, complicating understanding of how processing might impact learning. The form that learning takes differs based on domain, making it essential to understand both the form of learning within the domain and the way that processing plays out. Thorough investigations of interactions between stimulus processing and learning in other domains can thus further knowledge of domain-specific learning and domain-general properties that carry through different domains.

7.3 Theories of word learning

Although nearly all researchers of word learning acknowledge some associative aspects to word learning, many suggest that the associative nature of learning dissipates as the learner becomes more skilled (Golinkoff & Hirsh-Pasek, 2006; Namy, 2012; Trueswell et al., 2013; Waxman & Gelman, 2009). Under such assumptions, the more advanced word learner relies on more explicit word learning tools in order to acquire word-referent mappings; associative mechanisms are no longer necessary as learners come to better understand the nature of word learning.

Such explicit word learning theories suggest that learning operates on single-word/single-referent mappings. The learner forms hypotheses regarding which words map to which referents, then maintains these unitary hypotheses until counterevidence is received (Medina et al., 2011; Trueswell et al., 2013). The results of the current study are incompatible with such theories. In the *synchronous* presentation cases of these experiments, participants formed a mapping between the correct word-form and the referent, but also built parallel associations between competing word-forms and the same referent. These parallel associations entail multiple links between word-forms and referents, counter to the prediction that learners only entertain a single hypothesis at a time.

These findings thus argue for an associative form of word learning for advanced learners. The participants in these studies were all adults with fully-formed lexica; if people change word learning strategies as they become more adept at learning, they should surely have fully adopted the new strategy by adulthood. Instead, they appear able to continue to rely on associative mechanisms in at least some word learning tasks. Although other forms of word learning may use more explicit mechanisms, the evidence of associative learning in the present dissertation confirms that adults have not lost the capacity to use implicit forms of word learning.

Perhaps more importantly, these results show that associative word learning is a much more sophisticated process than often suggested. In critiques of associative approaches to word learning, many researchers suggest that associative learning is a dumb process of linking perceptual representations (Gelman & Waxman, 2009; Golinkoff & Hirsh-Pasek, 2006; Waxman & Gelman, 2009). This dissertation presents several cases where associations are formed based on internal representations. In Experiment 1, participants in the *delay* condition form associations after auditory offset, relying on some memory representation of the stimulus. Experiment 2 showed continued updating of associations after both auditory and visual offset, showing that learning continues to occur as representations are stored in working memory storage (and as these representations continue to update to resolve ongoing competition even as they are simultaneously linked via associative mechanisms). Experiment 3 showed that information that precedes auditory-visual pairings nevertheless affects the way that associations are formed; although the physical co-occurrence between Experiment 1's *synchronous* condition and Experiment 3 was identical, stimulus processing dynamics external to the co-occurrence event affect the formation of associations.

This suggests a form of associative learning that is much deeper and more complex than the simple stimulus-stimulus mappings suggested by advocates of inferential approaches to word learning. Associations can form not just between what is currently being perceived, but also between internal representations of stimuli. Associative learning is richer than behaviorism (McMurray, Zhao, Kucker, & Samuelson, 2013; Rescorla, 1988; Sloutsky, 2009; L. B. Smith et al., 2003; Wasserman & Miller, 1997).

An interesting finding from these experiments is that word learning appears to function more effectively when words and their referents are not presented simultaneously. *Synchronous* stimulus presentation caused learners to form spurious associations with competitors, which led to increased competition at later recall; meanwhile, *delay* presentation reduced this competition, signaling that the correct associations between the word and its referent dominated what was learned. This finding is rather counterintuitive. Typically, synchrony is thought essential to learning words; when the learner hears “dog,” she is best able to map this to a referent if she is currently looking at a dog. However, it appears that a slight delay after auditory presentation can be quite valuable to avoid forming weak word-referent mappings with competitors. In many cases, facets of the natural learning situation may ameliorate these weak associations, as learners can continue to fixate the referent as they resolve competition (and they may have begun processing the referent before hearing the target, as in Experiment 3). However, the finding that *synchrony* is not necessarily the ideal way to teach word-referent mappings is quite novel.

7.3.1 Caveat on natural word learning

If learners use unsupervised methods to learn words, and if such unsupervised learning leads to spurious associations with competing word-forms, how are words learned as effectively as they are? Associative word learning is thought to be most apparent for young children (e.g., Golinkoff & Hirsh-Pasek, 2006; Namy, 2012). Children also appear to have slower lexical activation processes and increased competition (Fernald et al., 1998). Meanwhile, children are learning a huge number of words in a short period of time, and they are learning many words in parallel (Ganger & Brent, 2004; McMurray, 2007). This confluence suggests that children may be particularly at risk of forming false associations from competing word-forms as they learn. Indeed, the Language Restructuring Model argues that children need to develop more precise phonological representations to deal with such risks of conflated word-referent pairings as the lexicon becomes denser (Metsala & Walley, 1998; Walley et al., 2003).

However, such issues may also be quite minor in ecological learning situations. In order for these types of spurious associations to affect lexical processing, the learner must be learning two overlapping word-forms concurrently. During the learning event, both of these word-forms must become active frequently enough when referents are present to lead to strong associations for both word-forms. However, such parallel learning of phonologically-related word-forms is likely quite rare; Swingley and Aslin found that the lexical inventories of 14-15 month-old infants show few examples of overlapping words, suggesting that at least early in word learning, competitors are less common (Swingley & Aslin, 2002). Similarly, during learning, the timing of word-referent pairings likely varies quite widely. Although in some instances the learner is attending to the referent during periods of high ambiguity, there are likely many other instances in which the referent is not identified until after competition has resolved (for example in a highly cluttered visual scene), or the range of potential referents has been identified (or narrowed down) by things like social cues well before the word is heard (an analogue of Experiment 3). In most instances, the learner likely can continue to consider the word-referent pairing after word offset, leading to ongoing updating of the learned associations (an analogue of Experiment 2). This variability in timing would lessen the false associations formed with overlapping word-forms.

Investigating the role of temporal structure of language in word learning may not offer a major contribution to our understanding of the development of specific lexical items. However, by creating scenarios in which the interactions between temporal dynamics and learning are emphasized, these studies provide insight into the nature of word learning more broadly, as well as into more general learning mechanisms. Although formation of spurious associations as a result of parallel activation likely does not lead to massive competition from countless spurious associations throughout life, the evidence found showing that such associations do arise indicates that unsupervised learning is active during word learning, and that this learning occurs continuously in time, even during periods of competition.

7.4 Perceptual processing and learning

Perceptual processing is rife with temporal dynamics. Reaction time is a standard tool in cognitive research as it offers a glimpse into the dynamics at play in cognitive processing (Holden, Van Orden, & Turvey, 2009). In an array of domains, the measure of the temporal components of processing is central to understanding the stages that occur during those processes. For example, the Stroop task uses the speed with which participants can name the color that a word is printed in (relative to the speed with which they can read the word) as a way to understand the components that play into inhibition (MacLeod & Dunbar, 1988; MacLeod, 1991; Stroop, 1935). Similarly, in visual search tasks, the slope of RT with respect to number of displayed competitors is used as a way to understand how long it takes to process stimuli, as well as to gauge which aspects of stimuli are processed automatically (Treisman & Gelade, 1980; Wolfe, 1994).

Word recognition has long relied on such measures as a tool for understanding the various processes that impinge on lexical activation. For example, listeners typically are slower to recognize words with many phonological neighbors in the lexicon (Luce & Pisoni, 1998); this is taken as evidence of competition between similar word-forms as words are being activated. Similarly, semantic priming has a rich history as a measure of the speed of activating a words meaning (Andruski et al., 1994; Milberg, Blumstein, & Dworetzky, 1988; Zwitserlood, 1989); words that elicit speeded RTs for related primes are thought more effectively access their semantic network. Thus in word recognition, temporal dynamics are central to understanding of processing.

The temporal dynamics of lexical processing are particularly well-studied, as the temporal nature of spoken language lends itself nicely to investigation of processing dynamics. The use of methodologies that are sensitive to the temporal nature of spoken word recognition has led to a comprehensive understanding of the timecourse of recognition; tools such as gating (Bard et al., 1988; Grosjean, 1985; Marslen-Wilson, 1987) and eye-tracking (Allopenna et al., 1998; McMurray et al., 2008; Tanenhaus et al., 1995) provide detailed glimpses of how quickly listeners process acoustic information and when competitor effects occur. Quite comprehensive descriptions of word recognition are thus possible, and these descriptions include a complex pattern of parallel activation and competition throughout processing.

Yet despite decades of research on the timecourse of perceptual processing, the influence of the dynamics of processing on learning has not been thoroughly explored. In many instances, learning is studied in ways that remove effects of perceptual processing in order to more directly investigate other components of the process. When attempting to determine whether people learn categories on the basis of prototypes or exemplars, for example, how long it took the learner to process the features of the stimulus is irrelevant.

Yet studying the interaction between the dynamics of stimulus processing and learning offers insight into the way that learners form mappings in individual learning events.

7.5 Continuous unsupervised learning

The experiments in this dissertation demonstrate that learning is not functionally independent of processing dynamics. Instead, at least in some circumstances, learning proceeds while stimuli are being processed, and mappings begin forming while competition between perceptual forms is ongoing. Experiments 1-4 in this dissertation emphasized the use of unsupervised associative learning to acquire word-referent mappings for a new set of words. This form of learning is often thought to be implemented by a basic Hebbian learning process of strengthening connections between co-present stimuli and weakening connections between stimuli that do not appear together (Hebb, 1949). However, caution is always necessary when comparing theories of neural learning to theories of cognitive learning. Learning a word likely entails forming links between populations of neurons whose activation levels offer a representation of the word (though see, Bowers, 2009). Thus the parallel between Hebbian learning and associative word learning is best considered an analogy.

However, the results of the experiments in this dissertation strengthen this analogy. Rather than associative learning relying on smart mechanisms to gate learning to occur after competition is resolved, instead learning proceeds continuously, whenever representations of both a word-form and a referent are active. This is similar to the classic conception of Hebbian learning, in which the brain operates by strengthening the connections between any neurons that are co-active and weakening connections between those whose activations are not correlated (Hebb, 1949). As lexical processing is ongoing, the varied degrees of activation of competing word-forms lead to learning that is continuous across time.

Unsupervised learning mechanisms thus operate without the need for additional machinery to determine when learning can proceed. Although it is possible that simple competition monitors or activation thresholds could gate learning, there is no specific evidence for such mechanisms here, and as a result it appears that unsupervised learning works without these monitors, simply linking any stimuli that are co-active. This points to a crucial need to consider the processing dynamics in any learning domain; for example, if we are trying to teach a child to read, we need to consider not just the regularities linking orthographic and phonological forms, but also how the processing of stimuli in each domain might impact the learning (Harm, McCandliss, & Seidenberg, 2003; Harm & Seidenberg, 2004). Learning does not occur in a vacuum, but instead is influenced by the way that stimuli are processed.

However, it remains unclear whether unsupervised learning occurs in this way when a response is also required. During tasks with responses, two distinct modes of learning are possible. In one mode, learning occurs continuously in time, with processing in order to make a response a parallel process. This processing would influence the continued learning, particularly as competition resolves sufficiently to make a response. In the alternative case, the necessity to form a response may change the way learning proceeds altogether. Learners can use the response as a signal to update representations, without forming any associations until this response is made. This would entail a single instance of learning in each trial, and it would occur after competition is resolved. Supervised and unsupervised learning need not be functionally independent forms of learning; instead, they likely operate in tandem to drive learning (Munakata & O'Reilly, 2003; O'Reilly & Norman, 2002; O'Reilly & Soto, 2002; Zhu et al., 2007). During the lexical processing that precedes a response, listeners may thus be learning using unsupervised means, but then update this learning based on information available when responding.

7.6 The value of domain-specific investigations

Typically in studies of learning, researchers only examine domain-specific interactions with the learning task (e.g., the specifics of word learning) when questions about that domain are being asked. In contrast, if the goal is to explore more general aspects of learning system, researchers typically employ abstract stimuli that are quite removed from the kinds of stimuli learned in natural situations (Posner & Keele, 1968; Waldron & Ashby, 2001; Wifall, McMurray, & Hazeltine, 2012). Domain-specific investigations are quite useful for elaborating how learning occurs within that domain, but they often do not extend into general principles of learning (though some theoretical accounts offer notable exceptions: McMurray et al., 2012; Thiessen, 2011).

This dissertation took a quite different approach. By situating learning firmly within a domain in which the temporal dynamics of processing are well understood, these experiments offer insight into how the specific temporal properties within a domain impact domain general learning mechanisms. These experiments embrace the complex nature of stimulus processing within word learning as a tool to determine when learning proceeds more broadly. The results of these studies suggest that unsupervised learning mechanisms operate continuously in time, whenever stimulus representations are active. While these findings are informative for word learning, they offer a deeper value in expanding understanding of the mechanisms of learning.

This domain-specific approach is thus a valuable way to approach studies of learning. Although seeking domain-general principles is quite necessary, understanding how domain-specific information impacts learning is a useful tool to this end. By understanding how learning proceeds in the face of processing within a domain, we can better understand the basic characteristics of learning.

7.7 Conclusions

Studies of learning have thrived by asking questions that obviate concerns of processing dynamics. Yet the way that stimuli are processed can have major implications for how these stimuli are learned. By embracing the interactions between stimulus processing and learning, a more comprehensive description of learning is possible. This dissertation offers an initial foray into the interface between learning and processing. In considering the nature of word recognition processes as they relate to word-referent learning, these studies show that learning is a continuous process. Rather than noise that complicates understanding of learning, perceptual processes are valuable sources of information that impact how learning proceeds.

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*(4), 419–439. doi:10.1006/jmla.1997.2558

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502–518. doi:10.1016/j.jml.2006.12.004

Andruski, J. E., Blumstein, S. E., & Burton, M. W. (1994). The effect of subphonetic differences on lexical access. *Cognition*, *52*(3), 163–187.

Apfelbaum, K. S., Blumstein, S. E., & McMurray, B. (2011). Semantic priming is affected by real-time phonological competition: evidence for continuous cascading systems. *Psychonomic Bulletin & Review*, *18*(1), 141–9. doi:10.3758/s13423-010-0039-8

Apfelbaum, K. S., Hazeltine, E., & McMurray, B. (2012). Statistical Learning in Reading: Variability in Irrelevant Letters Helps Children Learn Phonics Skills. *Developmental Psychology*. doi:10.1037/a0029839

Apfelbaum, K. S., & McMurray, B. (2011). Using variability to guide dimensional weighting: associative mechanisms in early word learning. *Cognitive Science*, *35*(6), 1105–38. doi:10.1111/j.1551-6709.2011.01181.x

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

Bard, E. G., Shillcock, R. C., & Altmann, G. T. M. (1988). The recognition of words after their acoustic offsets in spontaneous speech : *Perception & Psychophysics*, *44*(5), 395–408.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*(4), 629–654.

Bates, D., & Sarkar, D. (2011). lme 4: Linear mixed-effects models using S4 classes. R package version 099875-9.

Berwick, R. C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science*, *35*(7), 1207–42. doi:10.1111/j.1551-6709.2011.01189.x

Boersma, Paul & Weenink, David (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.52

Booth, A. E., & Waxman, S. R. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology*, *38*(6), 948–957.

Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: some implications for categorical perception and levels of information processing. *Psychological Research*, *46*(3), 207–222.

Bortfeld, H., & Morgan, J. L. (2010). Is early word-form processing stress-full? How natural variability supports recognition. *Cognitive Psychology*. doi:10.1016/j.cogpsych.2010.01.002

Bowers, J. S. (2009). On the biological plausibility of grandmother cells: Implications for neural network theories in psychology and neuroscience. *Psychological Review*, *116*(1), 220–251.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, *61*(2), 206–219.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–44.

Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, *128*(3), 407–416. doi:10.1016/j.cognition.2013.05.010

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference*, *15*, 17–29.

Catalano, J., & Kleiner, B. (1984). Distant transfer in coincident timing as a function of variability of practice. *Perceptual and Motor Skills*, *58*, 851–856.

Cepeda, N. J., & Munakata, Y. (2007). Why do children perseverate when they seem to know better: graded working memory, or directed inhibition? *Psychonomic Bulletin & Review*, *14*(6), 1058–65.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science*, *26*, 609–651.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming Syntactic. *Psychological Review*, *113*(2), 234–272. doi:10.1037/0033-295X.113.2.234

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, *17*(1), 205–215.

Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language*, *22*(3), 727–735.

Chater, N., & Christiansen, M. H. (2010). Language Acquisition Meets Language Evolution. *Cognitive Science*, *34*, 1131–1157. doi:10.1111/j.1551-6709.2009.01049.x

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*(2), 417–30. doi:10.1037/a0027175

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, *30*(3), 637–669. doi:10.1017/S0305000903005701

Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to Segment Speech Using Multiple Cues: A Connectionist Model. *Language and Cognitive Processes*, *13*(2), 221–268.

Clark, A. (in press). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 1–86.

Cluff, M. S., & Luce, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: retroactive effects on multiple activation. *Journal of Experimental Psychology. Human Perception and Performance*, *16*(3), 551–63.

Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constrainst. *Journal of Memory and Language*, *30*(2), 234–250. doi:10.1016/0749-596X(91)90005-5

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*(2), 193-210.

Creel, S. C. (2012). Phonological similarity and mutual exclusivity: on-line recognition of atypical pronunciations in 3-5-year-olds. *Developmental Science, 15*(5), 697-713. doi:10.1111/j.1467-7687.2012.01173.x

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, *106*(2), 633–664. doi:10.1016/j.cognition.2007.03.013

Creel, S. C., & Dahan, D. (2010). The effect of the temporal structure of spoken words on paired-associate learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 110–122. doi:10.1037/a0017527

Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*(4), 492–504. doi:10.1016/j.wocn.2010.12.002

Dahan, D., & Gaskell, M. G. (2007). The temporal dynamics of ambiguity resolution: Evidence from spoken-word recognition. *Journal of Memory and Language*, *57*(4), 483–501. doi:10.1016/j.jml.2007.01.001

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: evidence from eye movements. *Cognit Psychol*, *42*(4), 317–367. doi:10.1006/cogp.2001.0750

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

Dahan, D., & Tanenhaus, M. K. (2004). 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*, *30*(2), 498–513. doi:10.1037/0278-7393.30.2.498

De Sa, V., & Ballard, D. H. (1998). Category Learning Through Multi-Modality Sensing. *Neural Computation*, *10*(5), 1097–1117

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35–9. doi:10.1111/j.1467-9280.2007.01845.x

Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, *14*(2), 179–211. doi:10.1207/s15516709cog1402_1

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*(4), 547–582. doi:10.1111/j.1551-6709.2009.01023.x

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376–83. doi:10.1111/j.1467-8624.2010.01479.x

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid Gains in Speed of Verbal Processing by Infants in the 2nd Year. *Psychological Science*, *9*(3), 228–231. doi:10.1111/1467-9280.00044

Fisher, S. E., & Marcus, G. F. (2006). The eloquent ape: genes, brains and the evolution of language. *Nature Reviews. Genetics*, *7*(1), 9–20. doi:10.1038/nrg1747

Fodor, J. A., Gaskell, M. G., & Brill, S. L. (1975). Pi ka pu: The perception of speech sounds by prelinguistic infants. *Perception and Psychophysics*, *18*(2), 74–78.

Fodor, J. A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using Speakers' Referential Intentions to Model Early Cross-Situational Word Learning. *Psychological Science*, *20*(5), 578–585. doi:10.1111/j.1467-9280.2009.02335.x

Freeman, J. H., & Steinmetz, A. B. (2011). Neural circuitry and plasticity mechanisms underlying delay eyeblink conditioning. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(10), 666–77. doi:10.1101/lm.2023011

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental psychology*, *40*(4), 621–632. doi:10.1037/0012-1649.40.4.621

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology. Human Perception and Performance*, *6*(1), 110–25.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*(2), 105–132. doi:10.1016/S0010-0277(03)00070-2

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating Form and Meaning: A Distributed Model of Speech Perception. *Language and Cognitive Processes*, *12*(5/6), 613–656. doi:10.1080/016909697386646

Gaskell, M. G., & Marslen-Wilson, W. D. (1999). Ambiguity, competition, and blending in spoken word recognition. *Cognitive Science*, *23*(4), 439–462. doi:10.1207/s15516709cog2304_3

Gelman, S. A., & Waxman, S. R. (2009). Response to Sloutsky: Taking development seriously: theories cannot emerge from associations alone. *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2009.05.004

Gliozzi, V., Mayor, J., Hu, J.-F., & Plunkett, K. (2009). Labels as Features (Not Names) for Infant Categorization: A Neurocomputational Approach. *Cognitive Science*, *33*(4), 709–738. doi:10.1111/j.1551-6709.2009.01026.x

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*, 501–518.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 8030–8035. doi:10.1073/pnas.1332441100

Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Baby Wordsmith: From Associationist to Social Sophisticate. *Current Directions in Psychological Science*, *15*(1), 30–33. doi:10.1111/j.0963-7214.2006.00401.x

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (2008). Early object labels: the case for a developmental lexical principles framework. *Journal of Child Language*, *21*(01). doi:10.1017/S0305000900008692

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211–244. doi:10.1037/0033-295X.114.2.211

Grosjean, F. (1985). The recognition of words after their acoustic offset: evidence and implications. *Perception & Psychophysics*, *38*(4), 299–310.

Gruendel, J. (1977). Referential extension in early language development. *Child Development*, *48*(4), 1567–1576.

Hall, D. G., Williams, S., & Belanger, J. (2010). Learning Count Nouns and Adjectives: Understanding the Contributions of Lexical Form Class and Social-Pragmatic Cues. *Journal of Cognition and Development*, *11*(1), 86–120. doi:10.1080/15248370903453592

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*(2), 155–182.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, *106*(3), 491–528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychol Rev*, *111*(3), 662–720. doi:10.1037/0033-295X.111.3.662

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, *31*(3-4), 373–405. doi:10.1016/j.wocn.2003.09.006

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley and Sons.

Hirsh-Pasek, K., Nelson, D. G. K., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.

Holden, J. G., Van Orden, G. C., & Turvey, M. T. (2009). Dispersion of response times reveals cognitive dynamics. *Psychological Review*, *116*(2), 318–342. doi:10.1037/a0014849

Hoover, J. R., Storkel, H. L., & Hogan, T. P. (2010). A cross-sectional comparison of the effects of phonotactic probability and neighborhood density on word learning by preschool children. *Journal of Memory and Language*, *63*(1), 100–116. doi:10.1016/j.jml.2010.02.003

Horst, J., & Samuelson, L. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157. doi:10.1080/1525O(xH)701795598

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language Experience Shapes the Development of the Mutual Exclusivity Bias. *Infancy*, *15*(2), 125–150. doi:10.1111/j.1532-7078.2009.00009.x

Hsu, A. S., & Chater, N. (2010). The logical problem of language acquisition: a probabilistic perspective. *Cognitive Science*, *34*(6), 972–1016. doi:10.1111/j.1551-6709.2010.01117.x

Huet, M., Jacobs, D. M., Camachon, C., Missenard, O., Gray, R., & Montagne, G. (2011). The education of attention as explanation of variability of practice effects: Learning the final approach phase in a flight simulator. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(6), 1841–54. doi:10.1037/a0024386

Huette, S., & McMurray, B. (2010). Continuous dynamics of color categorization. *Psychonomic Bulletin & Review*, *17*(3), 348–54. doi:10.3758/PBR.17.3.348

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*(4), 460–482. doi:10.1016/j.jml.2007.02.001

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*(4), 434–446. doi:10.1016/j.jml.2007.11.007

Johnson, S. C., Ok, S.-J., & Luo, Y. (2007). The attribution of attention: 9-month-olds' interpretation of gaze as goal-directed action. *Developmental Science*, *10*(5), 530–7. doi:10.1111/j.1467-7687.2007.00606.x

Ju, M., & Luce, P. A. (2006). Representational specificity of within-category phonetic variation in the long-term mental lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 120–138. doi:10.1037/0096-1523.32.1.120

Kachergis, G., Yu, C., & Shiffrin, R. (2012a). An associative model of adaptive inference for learning word–referent mappings. *Psychonomic Bulletin & Rreview*, *19*(2), 317–324.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2012b). Cross-situational word learning is better modeled by associations than hypotheses. *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, 1–6. doi:10.1109/DevLrn.2012.6400861

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11163–11170. doi:10.1073/pnas.1005062107

Kapnoula, E., Gupta, P., Packard, S., & McMurray, B. (submitted). Immediate lexical integration of novel word forms.

Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44.

Lindsay, S., & Gaskell, M. G. (2013). Lexical integration of novel words without sleep. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(2), 608–22. doi:10.1037/a0029243

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36.

MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin*, *109*, 163–203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: evidence for a continuum of automaticity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *14*(1), 126–135.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*(1), 133–156.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The Time Course of Spoken Word Learning and Recognition: Studies With Artificial Lexicons. *Journal of Experimental Psychology: General*, *132*(2), 202–227.

Mani, N., & Plunkett, K. (2010). Twelve-Month-Olds Know Their Cups From Their Keps and Tups. *Infancy*, *15*(5), 445–470. doi:10.1111/j.1532-7078.2009.00027.x

Marcus, G. (1993). Negative evidence in language acquisition. *Cognition*, *46*, 53–85.

Marcus, G., Pinker, S., & Ullman, M. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4).

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77. doi:10.1016/0364-0213(90)90026-S

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*, 1–27.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157. doi:10.1016/0010-0285(88)90017-5

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102.

Marslen-Wilson, W. D., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1376-1392.

Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576–585.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: information-processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372–80.

Mayor, J., & Plunkett, K. (2007). Constraints on generalization in a self-organising model of early word learning. In R. M. French & E. Thomas (Eds.), *From Associations to Rules: Connectionist Models of Behavior and Cognition* (pp. 66–77).

McClelland, J. L. (2010). Emergence in cognitive science. *Topics in Cognitive Science*, 1–20. doi:10.1111/j.1756-8765.2010.01116.x

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

McClelland, J. L., St John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4*(3/4), 287–335.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*, 631.

McMurray, B., Clayards, M. a, Tanenhaus, M. K., & Aslin, R. N. (2008). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin & Review*, *15*(6), 1064–71. doi:10.3758/PBR.15.6.1064

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831–77. doi:10.1037/a0029872

McMurray, B., & Jongman, A. (2011). What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review*, *118*(2), 219–246. doi:10.1037/a0022325

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), B33–B42. doi:10.1016/S0010-0277(02)00157-9

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from lexical garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language*, *60*(1), 65–91. doi:10.1016/j.jml.2008.07.002

McMurray, B., Zhao, L., Kucker, S., & Samuelson, L. K. (2013). Pushing the envelope of associative learning: Internal representations and dynamic competition transform association into development. In G. Hollich & L. Gogate (Eds.), *Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence* (pp. 49–80). Hershey, PA: IGI Global.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(22), 9014–9. doi:10.1073/pnas.1105040108

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name--nameless category (N3C) principle. *Child Development*, *65*(6), 1646–62

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J L Metsala & L. C. Ehri (Eds.), *Word Recognition in Beginning Literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum Associates.

Milberg, W., Blumstein, S. E., & Dworetzky, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, *26*(4), 305–308.

Minda, J. P., & Smith, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(2), 275–292.

Morrison, C. M., & Ellis, A. W. (1995). Roles of Word Frequency and Age of Acquisition in Word Naming and Lexical Decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 116–133.

Munakata, Y., & O'Reilly, R. C. (2003). Developmental and computational neuroscience approaches to cognition: The case of generalization. *Cognitive Studies*, *10*, 76–92.

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing Cognitive Control: Three Key Transitions. *Current Directions in Psychological Science*, *21*(2), 71–77. doi:10.1177/0963721412436807

Namy, L. L. (2012). Getting Specific: Early General Mechanisms Give Rise to Domain-Specific Expertise in Word Learning. *Language Learning and Development*, *8*(1), 47–60. doi:10.1080/15475441.2011.617235

Ng, K. H., & Freeman, J. H. (2012). Developmental changes in medial auditory thalamic contributions to associative motor learning. *The Journal of Neuroscience*, *32*(20), 6841–50. doi:10.1523/JNEUROSCI.0284-12.2012

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234. doi:10.1016/0010-0277(94)90043-4

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology. General*, *115*(1), 39–61.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 54–65.

Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*(2), 211–233.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: the generalized recirculation algorithm. *Neural Computation*, *8*(5), 895–938.

O'Reilly, R. C. (2001). Generalization in interactive networks: the benefits of inhibitory competition and Hebbian learning. *Neural Computation*, *13*(6), 1199–1241.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, *6*(12), 505–510.

O'Reilly, R. C., & Soto, R. (2002). A model of the phonological loop: Generalization and binding. In Z. Dietterich, T.G., Becker, S., Ghahramani (Ed.), *Advances in Neural Information Processing Systems* (Vol. 14). Cambridge, MA: MIT Press.

Pater, J., Stager, C. L., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*(3), 384–402.

Pauli, W. M., & O'Reilly, R. C. (2008). Attentional control of associative learning--a possible role of the central cholinergic system. *Brain Research*, *1202*, 43–53. doi:10.1016/j.brainres.2007.06.097

Pierrehumbert, J. B. (2003). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. *Language and Speech*, *46*(2-3), 115–154. doi:10.1177/00238309030460020501

Pinker, S., & Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*(1-2), 73–193.

Plaut, D. C. (1997). Structure and Function in the Lexical System: Insights from Distributed Models of Word Reading and Lexical Decision. *Language and Cognitive Processes*, *12*(5 & 6), 765–806. doi:10.1080/016909697386682

Plunkett, K., & Marchman, V. A. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, *48*(1), 21–69.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3), 353–363.

Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchison, E., & Britton, B. (2006). An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia*, *44*(12), 2209–2221. doi:10.1016/j.neuropsychologia.2006.05.025

Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: MIT Press.

Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science*, *29*, 819–865.

Rescorla, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *The American Psychologist*, *43*(3), 151–60.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II* (pp. 64–99). Appleton-Century-Crofts.

Roembke, T. C., & McMurray, B. (submitted). Learners consider multiple hypotheses in parallel during cross-situational learning.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339–349. doi:10.1111/j.1467-7687.2008.00786.x

Rost, G. C., & McMurray, B. (2010). Finding the Signal by Adding Noise: The Role of Noncontrastive Phonetic Variability in Early Word Learning. *Infancy*, 1–28. doi:10.1111/j.1532-7078.2010.00033.x

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90*(1), 51–89. doi:10.1016/S0010-0277(03)00139-2

Salverda, A. P., Dahan, D., Tanenhaus, M. K., Crosswhite, K., Masharov, M., & McDonough, J. (2006). Effects of prosodically modulated sub-phonetic variation on lexical competition. *Cognition*. doi:10.1016/j.cognition.2006.10.008

Samuelson, L. K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15-20-month-olds. *Developmental Psychology*, *38*(6), 1016–37.

Samuelson, L. K., Schutte, A. R., & Horst, J. S. (2009). The dynamic nature of knowledge: Insights from a dynamic field model of children's novel noun generalization. *Cognition*, *110*(3), 322–345. doi:10.1016/j.cognition.2008.10.017

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: an alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*(1), 94–104.

Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding word learning in space. *PloS one*, *6*(12), e28095. doi:10.1371/journal.pone.0028095

Schafer, G., & Mareschal, D. (2001). Modeling infant speech sound discrimination using simple associative networks. *Infancy*, *2*(1), 7–28.

Schutte, A. R., Spencer, J. P., & Schöner, G. (2003). Testing the dynamic field theory: working memory for locations becomes more spatially precise over development. *Child Development*, *74*(5), 1393–417.

Seidenberg, M. S., & McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming,. *Psychological Review*, *96*(4), 523–568.

Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, *106*, 833–870.

Singh, L., White, K. S., & Morgan, J. L. (2008). Building a Word-Form Lexicon in the Face of Variable Input: Influences of Pitch and Amplitude on Early Spoken Word Recognition. *Language Learning and Development*, *4*(2), 157–178. doi:10.1080/15475440801922131

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*, 39–91.

Sloutsky, V. M. (2009). Theories about "theories": where is the explanation? Comment on Waxman and Gelman. *Trends in Cognitive Sciences*, *13*(8), 331–332. doi:10.1016/j.tics.2009.05.003

Sloutsky, V. M., & Robinson, C. W. (2013). Redundancy matters: flexible learning of multiple contingencies in infants. *Cognition*, *126*(2), 156–64. doi:10.1016/j.cognition.2012.09.016

Smith, J. D. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science*, *13*(5), 437–442.

Smith, K., Smith, A. D. M., & Blythe, R. A. (2011). Cross-Situational Learning: An Experimental Study of Word-Learning Mechanisms. *Cognitive Science*, *35*(3), 480-498. doi:10.1111/j.1551-6709.2010.01158.x

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19. s

Smith, L. B., Jones, S. S., Yoshida, H., & Colunga, E. (2003). Whose DAM account? Attentional learning explains Booth and Waxman. *Cognition*, *87*(3), 209–213. doi:10.1016/s0010-0277(02)00236-6

Spencer, J. P., Blumberg, M. S., McMurray, B., Robinson, S. R., Samuelson, L. K., & Tomblin, J. B. (2009). Short arms and talking eggs: Why we should no longer abide the nativist-empiricist debate. *Child Development Perspectives*, *3*(2), 79–87. doi:10.1111/j.1750-8606.2009.00081.x

Spencer, J. P., Samuelson, L. K., Blumberg, M. S., McMurray, B., Robinson, S. R., & Tomblin, J. B. (2009). Seeing the world through a third eye: Developmental systems theory looks beyond the nativist-empiricist debate. *Child Development Perspectives*, *3*(2), 103–105. doi:10.1111/j.1750-8606.2009.00087.x

Spiegel, C., & Halberda, J. (2010). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology*, *109*(1), 132–140. doi:10.1016/j.jecp.2010.10.013

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381–382. doi:10.1038/41102

Storkel, H. L. (2001). Learning new words: phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, *44*(6), 1321–1337.

Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language*, *29*(2), 251–274.

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1175–1192. doi:10.1044/1092-4388(2006/085)

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. doi:10.1037/h0054651

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*. doi:10.1016/j.jml.2008.11.003

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *75*, 1–20.

Swingley, D., & Aslin, R. N. (2002). Lexical Neighborhoods and the Word-Form Representations of 14-Month-olds. *Psychological Science*, *13*(5), 480–484. doi:10.1111/1467-9280.00485

Swingley, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*(2), 99–132. doi:10.1016/j.cogpsych.2006.05.001

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. doi:10.1126/science.7777863

Thelen, E., Schoner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment. A filed theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*, 1-86.

Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language*, *56*(1), 16–34. doi:10.1016/j.jml.2006.07.002

Thiessen, E. D. (2011). Domain General Constraints on Statistical Learning. *Child Development*, *82*(2), 462-470. doi:10.1111/j.1467-8624.2010.01522.x

Toscano, J. C., Anderson, N. D., & McMurray, B. (2013). Reconsidering the role of temporal order in spoken word recognition. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-013-0417-0

Toscano, J. C., & McMurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science*, *34*(3), 434–464. doi:10.1111/j.1551-6709.2009.01077.x

Toscano, J. C., & McMurray, B. (2012). Cue-integration and context effects in speech: Evidence against speaking-rate normalization. *Attention, Perception & Psychophysics*, *74*(6), 1284–301. doi:10.3758/s13414-012-0306-z

Treisman, A. M. (1986). Features and objects in visual processing. *Scientific American*, *255*(5), 114–125.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–156. doi:10.1016/j.cogpsych.2012.10.001

Utman, J. A., Blumstein, S. E., & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics*, *62*(6), 1297–1311.

Vitevitch, M. S., & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological Science*, *9*(4), 325–329. doi:10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of Memory and Language*, *40*(3), 374–408. doi:10.1006/jmla.1998.2618

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, Neighborhood Activation, and Lexical Access for Spoken Words. *Brain and Language*, *68*(1-2), 306–311. doi:10.1006/brln.1999.2116

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, *8*(1), 168–76.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing*, *16*(1), 5–20. doi:10.1023/A:1021789804977

Wasserman, E. A., & Miller, R. R. (1997). What's elementary about associative learning? *Annual Review of Psychology*, *48*, 573–607. doi:10.1146/annurev.psych.48.1.573

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Science, 13*(6), 258-263.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' Ability to Learn Phonetically Similar Words: Effects of Age and Vocabulary Size. *Infancy*, *3*(1), 1–30. doi:10.1207/S15327078IN0301_1

Westermann, G., & Mareschal, D. (2004). From parts to wholes: Mechanisms of development in infant visual object processing. *Infancy*, *5*(2), 131–151.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*(1), 114–132. doi:10.1016/j.jml.2008.03.001

White, K. S., Yee, E., Blumstein, S. E., & Morgan, J. L. (2013). Adults show less sensitivity to phonetic detail in unfamiliar words, too. *Journal of Memory and Language*, *68*(4), 362–378. doi:10.1016/j.jml.2013.01.003

Wifall, T., McMurray, B., & Hazeltine, E. (2012). Perceptual Similarity Affects the Learning Curve (but Not Necessarily Learning). *Journal of Experimental Psychology. General*. doi:10.1037/a0030865

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202–238.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(1), 1–14. doi:10.1037/0278-7393.32.1.1

Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, *70*(13-15), 2149–2165. doi:10.1016/j.neucom.2006.01.034

Yu, C., & Smith, L. B. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science, 18*(5), 414–420. doi:10.1111/j.1467-9280.2007.01915.x

Yu, C., & Smith, L. B. (2010). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, *9999*(9999). doi:10.1111/j.1467-7687.2010.00958.x

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: prior questions. *Psychological Review*, *119*(1), 21–39. doi:10.1037/a0026182

Yurovsky, D., Yu, C., & Smith, L. B. (in press). Competitive processes in cross-situational word learning. *Cognitive Science*.

Zaki, S. R., & Nosofsky, R. M. (2007). A high-distortion enhancement effect in the prototype-learning paradigm: dramatic effects of category learning during test. *Memory & Cognition*, *35*(8), 2088–2096.

Zénon, A., Hamed, S. B., Duhamel, J.-R., & Olivier, E. (2009). Attentional guidance relies on a winner-take-all mechanism. *Vision Research*, *49*(12), 1522–1531. doi:10.1016/j.visres.2009.03.010

Zhao, L. (2013). *The effect of referent similarity and phonological similarity on concurrent word learning*. (Unpublished doctoral dissertation). University of Iowa, Iowa City, IA.

Zhu, X., Rogers, T., Qian, R., & Kalish, C. (2007). Humans perform semi-supervised classification too. *AAAI Conference on Artificial Intelligence, 22*.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*(1), 25–64.