

**There's a lot on your plate: How phonological load affects language processing**

A dissertation submitted by

Ann Kochupurackal

in partial fulfillment of the requirements for the joint degree of

Doctor of Philosophy

in

Psychology and Cognitive Science

Tufts University

February 2026

© 2026, Ann Kochupurackal

Advisor: Ariel Goldberg, Ph.D.

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

## Abstract

This project investigates the impact of cognitive load on lexical access during language processing, specifically examining whether cognitive load affects the activation of target words and their competitors during comprehension and production. By exploring the nuances of lexical selection and competitor activation, this research reveals that lexical processing is not entirely automatic but is influenced by the resources available for managing competing demands. Drawing on previous research (Zhang and Samuel, 2018), this project comprises three studies: an eye-tracking experiment examining the effects of cognitive load on spoken word recognition, simulations exploring how inhibition and competition operate within the lexical system, and a production task analyzing the effects of cognitive load on voice onset times (VOTs).

Study 1 used a visual world paradigm to measure fixations to targets, cohort competitors, and rhyme competitors under phonological versus sham load. Results revealed no significant differences in target or competitor fixation dynamics between conditions. Study 2 TRACE simulations tested whether load affects recognition through domain general (resource depletion) or lexicon-internal (increased competitor activation) mechanisms; both accounts could accommodate the null behavioral effect at certain parameter levels. Study 3 used a spoken word paradigm to manipulate VOTs under the two load conditions. Results showed that phonological load reduced VOTs and selectively disrupted lexical context effects while preserving vowel context effects.

The findings have theoretical implications for understanding of lexical access and practical implications for everyday language use in multitasking scenarios – potentially informing strategies for improving communication under stress, especially for individuals facing lexical access challenges.

## Acknowledgements

In the midst of experiments that refused to run and analyses that refused to converge, I discovered that the best antidote to a failed analysis was a batch of baked goods. Baking became my parallel education: a lesson in how small variables matter, how patience yields results, and how even the flattest cookies teach you something for next time. Science, as it turns out, works kind of the same way – just with longer bake times and fewer immediate rewards. I am grateful to the many people who made this particular recipe possible.

First, to my advisor, Dr. Ariel Goldberg, who took a chance on me years ago. Your guidance has not only shaped this dissertation but the researcher I've become. Thank you for your patience, insight, and support throughout the years.

To my committee members – Drs. Nathan Ward, Gina Kuperberg, and Carrie Niziolek – thank you for helping me refine and strengthen this work from its earliest conceptualization to its final form. Special shoutout to Dr. Stephanie Badde and Vanalata Bulusu for letting me borrow the Multisensory Perception Laboratory's eye tracker, for helping me program Study 1, and for initial data preprocessing support.

To my cohort: Drs. Lauren Mason, Tom Wooten, and especially Dr. Young, my lab-mate. We were the COVID cohort – put into an unprecedented situation that could have isolated us but instead brought us closer together. I could not have done this without you.

To my old friends from NY and the new friends I've made throughout grad school – especially Dr. Sanjana Kadirvel and soon to be Drs. Gauri Harindranath and Keturah Ragland. Thank you for the laughter, distractions, and inside jokes ☺

To my sisters, Tina and Jenny, and my brother-in-law, Ajith: you are my role models, both academically and as human beings. You make me want to be better because you are, honestly, the

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

best people I know. Thank you for showing me what's possible and for believing I could get here too.

And lastly, to my mom. Everything I've accomplished, I owe to you. Thank you for every sacrifice, every word of encouragement, and every moment of care that has carried me to this point.

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

## Table of Contents

List of Tables	vii
List of Figures	viii
Acknowledgements	iii
1 Setting the table: A general introduction	2
An overview of lexical access	4
Observing lexical access: The visual world paradigm	6
2 Too many cooks in the kitchen: How load can impact a system	9
Cognitive load: Multi- and dual- tasking in everyday activities	13
Lexical access and cognitive load	16
Localizing the source of load effects	23
Purpose of Dissertation	26
3 Eye can't believe it's not butter: Phonological load and visual word recognition	28
Method	29
Results	35
Discussion	54
4 The proof is in the processing: Simulating load with jTRACE	63
Method	66
Procedures & Results	67
Discussion	83

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

5 Eating your words: Load effects on speech production _____	86
Method _____	88
Results _____	91
Discussion _____	100
6 Plating up: Bringing it all together _____	104
Distinct mechanisms of load effects _____	105
The resilience of target recognition _____	106
Bridging perception and production _____	107
Prioritizing goals _____	110
Conclusion _____	111
Appendix A. Study 1 Supplemental Information _____	113
Appendix B. Study 2 Supplemental Information _____	132
Appendix C. Study 3 Supplemental Information _____	135
References _____	140

**List of Tables**

Table 1	33
Table 2	37
Table 3	45
Table 4	49
Table 6	89
Table 7	92
Table 8	96

**List of Figures**

Figure 1	7
Figure 2	21
Figure 3	32
Figure 4	35
Figure 5	40
Figure 6A & B	41
Figure 7	44
Figure 8	46
Figure 9	61
Figure 10	65
Figure 11	68
Figure 12	72
Figure 13	76
Figure 14	80
Figure 15	81
Figure 16	90
Figure 17	94
Figure 18	95
Figure 19	97
Figure 20	99

## Setting the table: A general introduction

Imagine you're on a phone call when a text comes in: a colleague asking the name of a café you know well. This should be a simple question - but with your attention divided between the conversation and the text, you might find yourself struggling to retrieve a name that would normally come to you immediately. This common experience illustrates how cognitive load can disrupt lexical access – the process of retrieving individual words from memory – when mental resources are divided between competing tasks.

The purpose of my dissertation was to develop a model of how cognitive load impacts lexical access in spoken word recognition and production. While previous research has examined how competition is resolved *within* the language system (e.g., through lateral inhibition among lexical candidates; McClelland & Elman, 1986), less is known about how the language system interacts with broader cognitive demands. Understanding this interaction is essential for situating language processing within the larger cognitive ecosystem – recognizing that lexical access does not occur in isolation but alongside other processes competing for limited resources. Specifically, I examined how phonological load affects the sustained activation of lexical candidates under load. While previous research (e.g., Zhang & Samuel, 2015; 2018) has shown that cognitive load can impede competition, it remains unclear *how* and *where* load exerts its effects on speech processing – whether it operates within the lexical system itself, directly shaping activation dynamics and competition, or through domain-general control limitations that affect task performance more broadly. My dissertation addresses these questions by investigating the locus and mechanisms of phonological load effects across perception and production.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Understanding lexical access necessitates situating it within and connecting its functioning to the broader cognitive ecosystem. In the remainder of this chapter, I will provide an overview of lexical access and the visual world paradigm, a task used to research lexical activation and competition. This will set the stage for understanding how load can play a role in lexical processing. Chapter 2 presents an overview of cognitive load and the research on its impact on lexical competition. Despite the theoretical importance of understanding how these processes interact, relatively few studies have directly investigated this question. I examine the existing evidence and highlight the gaps that motivate the present research.

To develop a comprehensive account of how load mechanically impacts lexical competition, we investigated this phenomenon with three complementary approaches. Chapter 3 describes Study 1, which uses eye-tracking during spoken word recognition to examine whether phonological load modulates the time course of lexical competition. While fine-grained temporal dynamics and eye movements were not reliably affected, behavioral measures revealed clear load effects, suggesting that load impacts perception, but potentially at processing stages not captured by moment-to-moment competition measures. This pattern raises critical questions about the nature of load's influence: does it affect continuous activation dynamics within the lexical system, or does it reflect resource limitations that impact task performance more broadly (exogeneous)?

Following up on this analysis, Chapter 4 (Study 2) uses computational modeling with TRACE to explore these mechanistic possibilities. By simulating how lexicon-internal versus domain-general implementations of load would affect lexical processing, we generated testable predictions about what patterns should emerge under each account. This exploratory modeling

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

approach allowed us to characterize the signatures of different load mechanisms and evaluate which align with the fixation patterns observed in Study 1.

Chapter 5 (Study 3) extends our investigation to speech production, examining whether phonological load effects generalize across modalities and whether the pattern of results supports lexicon-internal or domain-general mechanisms. Production provides a critical test because while prior research has established load effects during spoken word recognition (Zhang & Samuel, 2015, 2018), examining production allows us to test whether these effects reflect perception-specific resource limitations or more general constraints on lexical access that operate across modalities.

Ultimately, this work helps refine models of spoken word recognition and production to account for how cognitive load interfaces with lexical access and selection, clarifying whether load acts as an architectural constraint within the language system or as a domain general resource limitation.

### **An overview of lexical access**

Lexical access is the mechanism in which language users retrieve words from their mental lexicon. In production, this involves selecting the appropriate word to convey an intended meaning (Dell, Chang & Griffin, 1999; Garrett, 1975; Levelt, 1989). In comprehension, lexical access refers to the process by which listeners map the incoming acoustic signal onto stored lexical representations (Marslen-Wilson, 1987). In both domains, lexical access involves two key stages: *lexical activation* and *lexical selection*.

*Lexical activation* refers to the initial stage where multiple word candidates are activated in response to an external stimulus (in comprehension) or a lexicon-internal meaning to be

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

expressed (in production). In spoken word comprehension, phonological information from an incoming stimulus is incrementally matched to words stored in long term memory, resulting in the activation of a pool of candidate words that are potential matches to the cue.<sup>1</sup> For example, upon hearing /pi.../, words containing, *peeve*, *peel*, etc.; Marslen-Wilson, 1987).<sup>2</sup> In the context of word production, the semantic features of a concept to be expressed are mapped to words that express one or more of those features, causing a set of candidate words to similarly be activated (e.g., [+small], [+domestic], [+feline] lead to the activation of words like *cat*, *dog*, *rabbit*, etc.).

*Lexical selection* occurs subsequently to lexical access and is where a single word is chosen from the activated candidates. Theories differ as to the exact mechanism of selection. According to some theories, lexical selection involves simply selecting the candidate with the highest activation—this is often referred to as the activation-based model (Levelt et al., 1999, Mahon et al., 2007). Other theories propose that selection involves competition among multiple activated candidates, which is resolved by selecting the word with the highest net activation after accounting for inhibitory effects of competitors (e.g., Dell, 1986; Roelofs, 1992). In a competitive selection system, the activation of non-target words decreases the speed and accuracy with which a target word is selected (Howard et al., 2006; Wheeldon & Monsell, 1994). Finally, lexical selection is also influenced by top-down factors such as context, syntax, and listener expectations about the speaker’s intent (e.g., Pickering & Garrod, 2013). These influences can pre-activate lexical items that are likely to occur, shaping competition before the

---

<sup>1</sup> It should be noted that a similar process is thought to occur when processing signs (Caselli & Cohen-Goldberg, 2014; Caselli, Emmorey, and Cohen-Goldberg 2021), however many details of how lexical access occurs in sign comprehension are yet unknown.

<sup>2</sup> Many if not most theories of lexical access propose that subsequent to the initial stimulus- or meaning-driven activation, related words – whether phonologically or semantically similar – also receive activation; a process termed ‘spreading activation’ (Dell, 1986; Luce & Pisoni, 1998; McClelland & Elman, 1986). For example, when /pi/ activates *peek*, activation may feed back from it to the phoneme /k/, which in turn would lead to activation of *beak* because of its partial match with the now activated /ik/. Activation could conceivably spread at semantic levels, leading *peek* to partially activate words like *look* and *glimpse*.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

acoustic signal fully unfolds. Some models formalize this through predictive coding mechanisms (e.g., Eddine et al., 2024), though top-down effects on lexical activation are broadly compatible with spreading activation frameworks as well.

### **Observing lexical access: The visual world paradigm**

One extremely useful tool for studying lexical access in comprehension is the visual world paradigm ('VWP'; e.g., Allopenna et al., 1998; McMurray, Farris-Trimble, & Rigler, 2017; Tanenhaus et al., 1995). In a standard VWP task, participants are shown an array of images and then presented with a spoken word and are tasked with clicking on the picture in the array corresponding to the heard word. During this process, the participant's eye movements are tracked, and research has shown that these fixations systematically reflect the internal processing of the auditory stimulus (Magnuson, 2019; McMurray, 2023). Figure 1 shows a representative pattern of fixations from a VWP study. The 'Target' line represents the proportion of fixations over time that participants direct toward the picture corresponding to the auditory stimulus. The 'Cohort' (competitors sharing onsets with the Target) and 'Rhyme' (competitors sharing endings with the Target) lines represent the proportion of fixations over time directed to phonological competitors of the Target, and the 'Unrelated' line represents fixations directed toward a phonologically unrelated distractor image. For example, if the Target on a given trial is *'peak'*, Cohort, Rhyme and Unrelated competitors could be *'peel'*, *'beak'*, and *'soap'* respectively. The graph represents fixations averaged over all participants and trials; time is given in milliseconds following the acoustic onset of the auditory stimulus.<sup>3</sup>

---

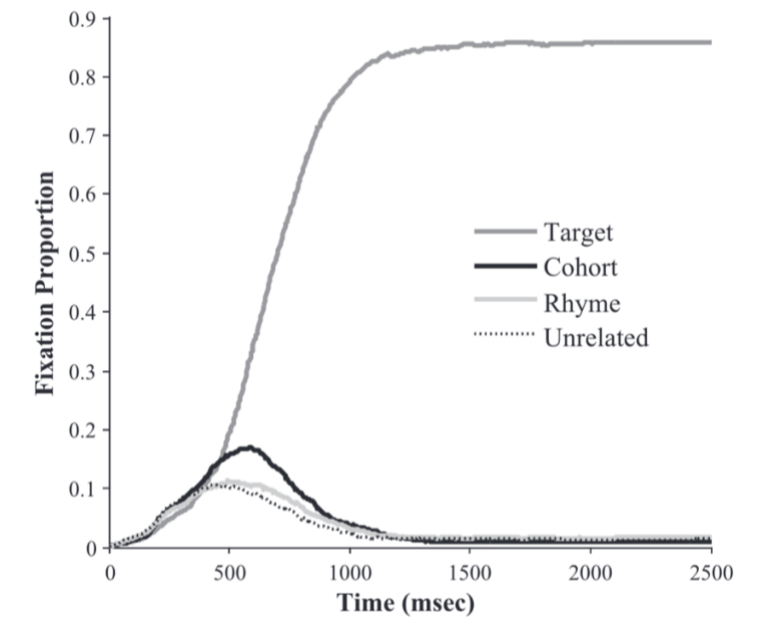
<sup>3</sup> Each auditory stimulus was 1000 ms long. There was 100 ms of silence at the onset and offset of each cue.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Initial fixations to Targets and Competitors can be viewed in Figure 1 in the time window of roughly 400-500 ms. At this point in time, fixations to the Target (*peak*) and Cohort competitor (*peel*) begin to diverge from the others, reflecting their match with the initial portion of the stimulus (e.g., /pi.../). This early parallel activation of onset-matching items is consistent with models of spoken word recognition in which lexical candidates are activated in proportion to their bottom-up fit with the unfolding speech signal (Allopenna et al., 1998; McClelland & Elman, 1986).

**Figure 1**

*Average timecourse of participant fixation in a standard Visual World Paradigm experiment*



*Note.* Data show the average proportion of fixations at each time point to each of the four types of onscreen stimuli. Graph adapted from McMurray, et al. (2017).

Subsequent feedback activation can be observed in the greater fixations to Rhyme competitors (*beak*) relative to Unrelated competitors (*soap*). Rhyme competitors don't match the

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

initial input and therefore receive no direct bottom-up support early in processing. Their later activation is thought to reflect top-down feedback from the lexical level to the phoneme level: as the target word becomes increasingly activated, it sends activation back to its constituent phonemes, which in turn activates other words sharing those phonemes, including rhymes (McClelland & Elman, 1986; McMurray et al., 2010).

We thus observe the general progression of lexical access described above: multiple candidates are activated initially followed by the eventual narrowing down and eventual asymptote of fixations to a single word, the Target.

**Too many cooks in the kitchen: How load can impact a system**

Psycholinguistic research has for the most part described lexical access as being *automatic* – operating without conscious effort or executive control and driven purely by lexicon-internal properties (e.g., activation) of lexical candidates (Dell, 1986; Forster, 1981; Posner & Snyder, 1975; Swinney, 1979). In most models, this competition is assumed to be managed entirely within the language system itself. For instance, interactive activation models posit lateral inhibitory connections between lexical candidates, such that as one word gains activation, it automatically suppresses its competitors (McClelland & Elman, 1986; McClelland & Rumelhart, 1981). On this view, selection emerges naturally from the dynamics of the network without requiring any external intervention. Similarly, in models that assume selection without competition, such as modular feedforward accounts (Forster, 1979; Levelt et al., 1999), the most activated candidate is simply selected based on fit with the input — again requiring no external resource to manage conflict. In either case, lexical access is presumed to be self-contained, operating independently of domain-general cognitive control. However, recent research suggests that it does in fact involve executive functions such as error monitoring, conflict resolution, and adjusting to different contexts, especially in more naturalistic language use environments.<sup>4</sup>

*Cognitive control* (also known as *executive function* in the multitasking literature) regulates the processes behind adaptability. It is a cluster of mental processes that enables individuals to adaptively manage their thoughts and behaviors, prioritizing relevant information and responses while disregarding distractions in dynamic environments (Miller & Cohen, 2001). Importantly, cognitive control facilitates goal-directed behavior by allowing individuals to set,

---

<sup>4</sup> While this project focuses on spoken language processing, evidence of this phenomenon has been amounting in reading as well (Besner & Roberts, 2003; Coltheart et al., 2001).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

maintain, and adjust their objectives in response to changing circumstances, ensuring that their actions remain aligned with their long-term aims. Converging evidence from the information monitoring, conflict resolution, and cognitive load literatures suggests that lexical access may not be as automatic as once believed and may incorporate cognitive control mechanisms to optimize processing.

*Conflict resolution* encompasses processes that range from resolving competition within a system to higher-level strategies that facilitate effective decision-making and comprehension in the face of ambiguity. At the low level, when competition arises from multiple competing representations, the system ramps up inhibitory mechanisms to manage selection pressure and reduce interference (Botvinick et al., 2001). This foundational process ensures that the most contextually relevant words are selected, particularly during language processing, where effective conflict resolution skills are critical for avoiding comprehension errors amidst ambiguity (Novick et al., 2005; Khanna & Boland, 2010). Steinberg Lowe and Buchwald (2023) had participants perform standard psycholinguistic tasks involving lexical access. They found the congruency sequence effect (CSE) in their data, a behavioral pattern that's believed to reflect conflict resolution processes; this suggests that conflict resolution processes play a role in lexical access.<sup>5</sup>

At the higher level, conflict resolution skills enable individuals to select the appropriate word among competitors and consider the perspective of their conversation partner to assess shared information during unscripted conversations (Khanna & Boland, 2010; Novick et al.,

---

<sup>5</sup> Gratton et al. (1992) were the first to demonstrate that the congruency sequence effect (CSE) is modulated by the amount of conflict in the preceding trial. Performance on a trial with conflict is faster and more accurate if it is preceded by another conflict trial, as compared to a non-conflict trial. Thus, the CSE is smaller following high-conflict trials than on trials following no (or low) conflict trials. This finding supports the notion that during word production, speakers engage cognitive control processes to monitor their productions and regulate their behavior online in an adaptive manner (Freund & Nozari, 2018; Lowe and Buchwald (2023)).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

2005, 2010; Hussey, 2013; Hussey & Novick, 2012; Thompson-Schill et al., 1997). Specifically, when faced with numerous lexical competitors, the activation of inhibitory mechanisms helps to enhance the selection of the most relevant word while overriding automatic responses in favor of less typical, contextually appropriate options. This multi-tiered approach ensures accuracy in language use by continuously evaluating the appropriateness of lexical choices and making necessary adjustments when errors are detected.

*Monitoring* refers to one's ability to track and evaluate a cognitive system's outputs and is typically seen as an error-prevention/correction mechanism. For example, when reaching for a coffee cup, monitoring allows you to notice that your hand is veering slightly off course and adjust mid-reach to grasp the handle accurately. In language processing, there is some suggestion that speakers may internally monitor their speech. Evidence for this can primarily be seen in studies eliciting galvanic skin responses when producing taboo words (Motley, Camden, & Baars, 1982) and the fact that speakers are occasionally able to catch (and interrupt) a speech error mid-articulation, a feat that is understood to be too quick to occur through external comprehension systems (Levelt, 1989). Other behavioral patterns such as a lexical bias in speech errors (i.e., phonological errors result in real words rather than non-words more often than expected by chance) and a phonetic similarity bias in speech errors (i.e., exchanging phonemes tend to share phonological features; Baars, Motley, & MacKay, 1975; Dell, 1986; Garrett, 1975; Oppenheim & Dell, 2010; Pinet & Nozari, 2018) are consistent though not decisive evidence of an internal monitor.<sup>6</sup> In spoken word comprehension, there is some suggestion that hearers are capable of shifting the way that incoming phonemic information is mapped onto lexical

---

<sup>6</sup> The idea that speakers internally monitor their speech is a topic of considerable debate among researchers (Nozari & Novick, 2017; Pickering & Garrod, 2013). While some argue that this internal monitoring is crucial for error correction and lexical selection (Levelt, 1989; Oppenheim & Dell, 2010), others contend that such processes may be less conscious or systematic (Schwartz, et al., 2006).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

candidates during lexical access under conditions of degraded input. For example, McMurray et al. (2009) found that under conditions of degraded input, participants appear to wait until all of the phonemes of a word have been perceived before activated lexical candidates, rather than activating them incrementally, as described above. This presumably could be the result of a system that monitors the success/efficacy of the spoken word comprehension system.

This growing body of evidence indicates that lexical access is a dynamic process influenced by various executive functions, including working memory (described in the next section), cognitive flexibility, and inhibition control, all of which play critical roles in language processing.

This raises the possibility that lexical selection may involve two distinct mechanisms. The first consists of inhibitory processes intrinsic to the language system – lateral inhibition between competing lexical representations that operates continuously during word recognition. The second is a domain-general conflict resolution system that may be recruited when competition becomes sufficiently intense or prolonged. If lexical processing draws on this shared resource pool, then taxing executive resources should lead to an overall reduction in the activation of lexical representations – both targets and competitors – as fewer resources are available to support the activation and maintenance of multiple candidates.

Understanding how these cognitive control mechanisms operate under different conditions, particularly in the presence of cognitive load, is crucial for gaining a more comprehensive picture of language processing in real-world settings. The next section will explore how cognitive load impacts lexical access and the implications for linguistic performance and adaptability.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

### **Cognitive load: Multi- and dual- tasking in everyday activities**

*Cognitive load* refers to any demands on working memory storage and the processing of information (Baddeley, 1992; Sweller, 1988; Sweller et al., 1990). It is essentially the total amount of mental effort required to process information, perform tasks, and solve problems. Cognitive load has been studied across many domains, including education (e.g., instructional design and learning strategies), clinical psychology (e.g., understanding the cognitive demands during therapy), human-computer interaction (e.g., the usability of software and user interface design), sports psychology (e.g., athletes' performance and quick decision making during high pressure situations), and workplace performance (e.g., productivity and decision-making in organizational settings).

Cognitive load refers to the mental effort required to process information and perform tasks, encompassing both working memory and attentional resources. Working memory can hold only a finite amount of information at any given time, and processing operations compete for access to this limited resource pool. Cognitive load, then, reflects how much of this capacity is currently occupied.

When load is low, ample resources remain available for ongoing processing. When load is high, the system approaches its capacity limit. Representations are maintained less precisely, processing slows as operations queue for limited resources, and interference between concurrent tasks increases. In the extreme, the system becomes overwhelmed, leading to decreased performance, information loss, and increased likelihood of errors.

This framework has important implications for tasks traditionally considered automatic. Even processes that typically proceed without deliberate effort may suffer under high load if they draw on the same limited resource pool. For example, trying to answer a simple text when you're

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

on a phone call becomes difficult not because answering the text is difficult, but because the resources normally devoted to this are being consumed by keeping track of the phone conversation.

Cognitive load plays a crucial role in determining how well individuals can manage multiple tasks at once. Multitasking is a normal part of everyday life, where thoughts and actions are seldom performed in isolation but rather in conjunction with multiple other thoughts and goals simultaneously. Multitasking involves rapidly switching between tasks (task switching) or performing multiple tasks concurrently (e.g., dual tasks), both of which increase cognitive load. When load is high, it impairs executive functions such as working memory, attention, and error monitoring – leading to diminished performance across tasks. It is also important to consider how it might affect processes seen as automatic, such as lexical access. Although lexical access is often viewed as an automatic function, it may still be influenced by the demands of cognitive load in complex or ambiguous situations. Understanding the mechanisms of cognitive load and its impact on multitasking is essential for developing strategies to improve efficiency and reduce errors in environments that require the simultaneous management of multiple tasks.

Dual tasking in particular offers an interesting example of what can happen when we're juggling multiple goals at once. This can lead to dual-task interference, in which performance on one or both tasks are affected (Bergen, Medeiros-Ward, Wheeler, Drews, & Strayer, 2012; Strayer & Drews, 2007). For example, attempting to engage in a conversation while composing an email can overload cognitive resources, resulting in slower response times and increased likelihood of errors in both activities.

One common explanation for this interference is a person's limited executive functioning resources that can be spread across multiple tasks; when two tasks require more of these

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

resources than are available, performance on one (or both) of these tasks are impaired (Kahneman, 1973). Although dual tasking and the effect of load has been researched primarily as a domain-general phenomenon, an alternative explanation that has been suggested is more domain-specific and relies on the *similarity* between the two tasks: the more similar two tasks are in terms of the processing required to perform them, the more interference we should observe (James, 1890; Kahneman, 1973; Pashler, 1994).

Imagine trying to listen to a podcast while reading a novel, versus listening to a podcast while reading a book. In the latter case, the tasks draw on largely distinct resources – auditory processing for the podcast, and motor/visual processing for the dishes – resulting in minimal interference. In the former case, both tasks require language processing and comprehension, creating competition for the same limited resources. The brain may struggle to decode both visual and auditory linguistic input simultaneously, leading to difficulties in retaining information from both sources. This results in decreased comprehension and retention for both reading the book and listening to the podcast, as the overlapping processes compete for shared capacity.

The interplay between cognitive load, lexical access, and multitasking further underscores the importance of cognitive control mechanisms in managing linguistic tasks. When lexical access is challenged by high cognitive load, such as in multitasking scenarios, the demands on cognitive control should, in theory, increase. This suggests that effective lexical access is not only a matter of automatic retrieval but also involves active management of cognitive resources to mitigate interference and maintain performance. This prediction arises from observations that high cognitive load and dual-task interference can impair performance on

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

linguistic tasks, indicating that cognitive control plays a critical role in facilitating successful lexical retrieval under demanding conditions.

Understanding how cognitive load influences lexical access can provide deeper insights into the broader dynamics of language processing and executive functioning. The following section explores how lexical access can be impacted by cognitive load, highlighting the crucial role of cognitive control in navigating these challenges, and how further research is needed to develop strategies for optimizing language performance in real-world, cognitively demanding situations.

### **Lexical access and cognitive load**

While psycholinguistic theories of language processing provide valuable insights into how lexical access operates, they have primarily been developed on the basis of internal lexical competition mechanisms. Such theories, while informative, may not fully capture the complexity of lexical access as it occurs in more dynamic, real-world contexts, where language processing is influenced by a wide range of cognitive and situational factors.

In naturalistic language use, lexical access is not an isolated process; it interacts with broader cognitive mechanisms, including executive functions, attentional control, and memory, which can influence both the activation and selection of words. Furthermore, variables such as cognitive load—the mental effort required for concurrent tasks—can significantly impact the efficiency and accuracy of lexical selection.

Important distinctions must be made regarding the type of load imposed. Lavie (2005, 2010) found evidence to suggest that perceptual load (physical complexity of stimuli) and cognitive load (working memory demands) affect performance differently: high perceptual load

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

increases focus on the main task, while high cognitive load reduces selective control and increases distraction.

Despite the extensive research on cognitive load theory and lexical access separately, there has been relatively little exploration regarding the interaction of the two, or how cognitive load impacts lexical processing. Understanding this interaction will shed light on how a cognitive control mechanism can support language processing under different cognitive demands, contributing to a more comprehensive theory of language and cognition.

There is some evidence that speech is processed in roughly the same way under cognitive load as under optimal conditions: for example, phonemic restoration in speech perception seems to function similarly under optimal conditions (Warren, 1970) and under load conditions (Mattys, Barden, & Samuel, 2014). However, evidence also exists suggesting that lexical access operates strikingly differently depending on the presence of cognitive load. Some studies have shown that while listeners rely more strongly on lexical-semantic knowledge than sub-lexical cues in optimal listening conditions, their relative reliance on some sub-lexical cues (rhythm and stress in particular) increases when sensory input is degraded by noise or atypical pronunciations (e.g., word recognition in noise task: Bradlow & Alexander, 2007; cross-modal priming with lexical decision task: Mattys et al., 2005).

Hsu and Novick (2016) investigated how cognitive control engagement influences real-time recovery from temporal sentence ambiguity. Participants completed Stroop trials immediately before listening to temporarily ambiguous spoken instructions (e.g., “Put the dumpling on the plate into the wok”) while viewing objects in a scene. Such sentences temporarily mislead listeners about the goal location – ‘on the plate’ initially seems like the destination, but ‘into the wok’ later reveals the true goal, requiring reinterpretation. When

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

instructions follow incongruent Stroop items – which require participants to suppress a prepotent response and thus engage in control mechanisms – listeners showed less consideration of the false interpretation, and dramatically earlier recovery of the correct one. Comprehension errors also decreased. The authors interpreted this as evidence that engaging cognitive control on the Stroop task left the control system in a heightened state of activation, which then facilitated conflict resolution during sentence reinterpretation. In other words, domain-general cognitive control resources recruited for the non-linguistic Stroop task remained engaged and supported the subsequent resolution of linguistic ambiguity. This demonstrates that engaging domain general cognitive control through a non-linguistic task can dynamically accelerate sentence interpretation processes, even as linguistic input is still unfolding.

Zhang and Samuel (2015) provided complementary evidence by examining load effects on different stages of lexical processing. They had participants perform a primed auditory lexical decision task. The auditory stimuli used as primes were specially selected so that they were ‘carrier’ words that contain embedded words that were semantically associated with the targets; the embedded word thus is the actual prime for the target. For example, one trial might consist of CARRIER: *napkin*, TARGET: *sleep*. In this trial, the carrier word *napkin* contains the embedded word prime *nap* which is a semantic associate of *sleep*. Another example is CARRIER: *patio*, TARGET: *burger*. In this case, the carrier word *patio* contains the embedded prime *patty* which is a semantic associate of *patty*. They found significant priming effects under optimal conditions with lexical decision reaction times to the target being faster when the carrier word contained a semantically embedded word (e.g., *nap*) versus when it contained an embedded word that was unrelated to the target. Importantly, this priming was observed when stimuli were noise-free and when participants were not under cognitive load. However, when noise was added to the primes,

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

when they were auditorily distorted, or when a cognitive load task was introduced, priming was not observed: *napkin* no longer primed *sleep*.<sup>7</sup> Interestingly, when embedded words were presented directly (e.g., *nap* presented alone), they still primed their associates under non-optimal conditions (e.g., *nap* primed *sleep*), suggesting that cognitive load does not interfere with the activation associated with the initial lexical activation phase but rather the subsequent maintenance of competitor activation over time. In the visual world paradigm (VWP), this maintenance process is reflected in the participants' eye movements: after initial activation of both targets and competitors, listeners continue to fixate competitor objects while competition is resolved (typically during the 500ms-1000ms time period following word onset; see in Figure 1). If cognitive load disrupts the maintenance of multiple active candidates, we expect to see reduced or shortened fixations to competitors under load – not because competitors fail to activate initially, but because their activation cannot be sustained with increased overall competition.

Considering these results together, we can see that lexical access is *not* automatic, and is subject to demands on cognitive resources. More specifically, the initial activation of lexical items may be automatic, but the subsequent processes involved in lexical selection (e.g., maintenance of competitor activation, competition, etc.) is not.

In their follow-up study, Zhang and Samuel (2018) replicated the effect and determined that the elimination of the priming effect is not due to cognitive load in general but specifically due to load that is phonological in nature. This study used the same primed auditory lexical decision task as before, except that the relationship between the prime and target was expanded. In some cases, the prime was embedded in the target word, as before (e.g., PRIME: *nap*,

---

<sup>7</sup> For the time being, we focused solely on the finding that cognitive load affects priming in this task. It is interesting and potentially very significant that auditory distortion had equivalent effects.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

TARGET: *napkin*). In other cases, the target was embedded in the prime (e.g., PRIME: *fancy*, TARGET: *fan*) and in others the prime was a Cohort competitor of the target (e.g., PRIME: *access*, TARGET: *accent*). Priming in these conditions was compared to an Unrelated condition where the target was semantically associated to a word that was not at all related to the carrier (e.g., CARRIER: *collar*, PRIME: *essay*, TARGET: *write*). Zhang and Samuel (2018) found that priming from non-word primes (which index initial lexical access without direct lexical competition) remained intact under cognitive load, whereas priming from word primes (which involve both lexical access and competition) was eliminated.

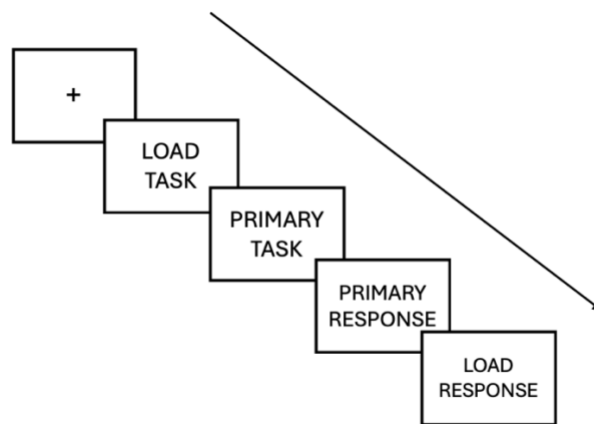
To investigate whether the elimination of the priming effect occurs under other types of load or whether it was specifically related to phonological load, Zhang and Samuel (2018) assigned participants to one of two conditions. In the phonological load condition, participants were shown a string of letters before the trial began (e.g., BCPY) and were instructed to remember them. They were then given the auditory lexical decision task. After making their lexical decision, participants were then presented with a letter (e.g., X) and asked to indicate whether it was one of the letters presented at the beginning of the trial. In the non-phonological load condition, the procedure was the same except a Chinese character (e.g., 𠄎) had to be retained in memory for the duration the trial. At the end of the trial, participants were presented with a single character and had to indicate whether it was the same as the character presented at the beginning of the trial. While the letter task could in principle be carried out visually (e.g., maintaining the visual form of the letters), the authors assumed (and this assumption appears to be born out) that participants would recode the letters in terms of their phonological form (e.g., /bi/). Meanwhile, as none of the participants were experienced with the Chinese writing system, the characters had to be maintained in a visual (non-phonological) format. In a third, 'No Load'

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

condition, participants performed only the priming task. The load type was a between-subjects variable, so participants received exclusively trials with phonological or non-phonological load (or no load) throughout the experiment.

### Figure 2

*A Single Trial of the Load Task Paradigm (Zhang & Samuel, 2018)*



*Note.* The load task was *either* phonological *or* non-phonological load for a single participant.

There were two main findings of this study: 1) phonological and non-phonological load reduce priming; and 2) phonological load interferes with priming more than non-phonological priming. These findings suggest that initial lexical access is relatively automatic, but maintaining multiple lexical candidates in competition requires cognitive resources – specifically, resources tied to phonological processing.

Zhang & Samuel (2015) proposed that lexical access can be a resource-demanding process and specifically relies on phonological resources. Under optimal conditions, non-target alternative lexical candidates are activated and retain enough activation over time to spread activation to semantically related words. This aligns with models of spoken word recognition

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

(e.g., Marslen-Wilson, 1987; McClelland & Elman, 1986) which posit that multiple lexical candidates are activated during word recognition and compete for selection. However, the maintenance of the activation of non-target words appears to depend on phonological resources. When phonological resources are not available, e.g., when a participant is keeping a string of letters in mind, the ability of alternative candidates to remain activated diminishes, truncating the competition process that normally unfolds over the course of word recognition.

These findings have important implications. At a theoretical level, they suggest that while initial stimulus-based activation of lexical items may be automatic, the subsequent maintenance of activation during selection requires phonological resources. If so, the characterization of lexical access as fully automatic must be revised: while the initial activation of lexical candidates may proceed without cognitive resources, the maintenance of multiple competing candidates does not. Secondly, they have important methodological implications: the empirical data we use to develop these theories are likely strongly affected by the particular properties of the task that is used. These properties will have to be intentionally considered/explicitly manipulated rather than being taken for granted. Lastly, they suggest that lexical access proceeds differently in different contexts. In quiet or otherwise “optimal” conditions, lexical access takes place according to the timecourse depicted in Figure 1. However, in more typical contexts (e.g., in an environment with ambient noise, in a conversational context where the listener is formulating their own response while listening) lexical access may proceed very differently. The larger implications of these findings suggest that our understanding of language processing, especially in real-world contexts, must account for the additional cognitive demands and environmental factors that shape lexical access. In environments where listeners are multitasking – such as holding a conversation while driving or responding in a noisy setting – lexical access may rely

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

more heavily on phonological and cognitive resources, slowing down or altering the timecourse of activation. This highlights the need for models of language processing to incorporate flexibility in how lexical access is understood, particularly under varying cognitive and environmental constraints.

One important question that remains is where these load effects originate. When cognitive load disrupts lexical processing, does this occur because load interferes with language-internal mechanisms, or because it taxes domain-general resources that language processing also depends upon? This distinction has profound implications for how we model language processing and how we understand the relationship between language and cognition more broadly.

### **Localizing the source of load effects**

A fundamental question in psycholinguistics concerns the locus of cognitive constraints on language processing. The studies reviewed above demonstrate that spoken word recognition can be disrupted when cognitive resources are taxed – whether by background noise, concurrent tasks, or working memory load. But what is the source of this disruption? Does interference arise from properties inherent to linguistic system itself, or do they reflect the influence of domain general, cognitive resources that operate external to language? This distinction – between lexicon-internal and domain-general mechanisms – has important implications for how we conceptualize the relationship between language and cognition more broadly.

The lexicon-internal perspective posits that certain processing limits are intrinsic to the language system's architecture. In this view, interference emerges from the properties of the linguistic representations and their connections themselves – for instance, the activation

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

dynamics of (competing) lexical units, their decay rates, etc. Such mechanisms are ‘lexicon-internal’ in that they are internal to the language processing system. For example, in production, speakers can monitor their speech *before* articulation – too fast to rely on external auditory feedback alone (Levelt, 1989). Self-correction can occur rapidly seen during silent speech articulation and when overt speech is masked by noise, suggesting a monitoring system operating within the production architecture itself (Jacobs et al., 2020; Oppenheim & Dell, 2008). This illustrates that at least some control-like functions – such as error detection and correction – may be carried out by mechanisms internal to the language system rather than by domain-general executive resources.

In contrast, the domain-general perspective attribute processing limitations to broader cognitive resources (e.g., attentional capacity, working memory, executive control ability that manage interference and resolve competition) that are recruited by language tasks. Under this view, language processing difficulties arise when these shared resources are taxed either by the demands of the linguistic task itself or by concurrent cognitive load.

Recent work by Ness and colleagues (2023) suggests that cognitive control processes recruited to resolve conflict in non-linguistic tasks (e.g., Stroop, flanker) are also engaged when listeners must resolve ambiguity or conflict in sentence comprehension. Kuperberg and Jaeger (2016) posit that during incremental language comprehension, listeners and readers predict probabilistically at multiple levels of representation, using higher level inferences to pre-reactivate information at lower representational levels. They propose that comprehension can be understood within a multi representational hierarchical generative framework (a type of framework proposed to explain other aspects of complex cognition) where the listener’s goal is to infer the speaker’s intended message by incrementally updating hypotheses as new

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

information becomes available. In this view, the brain continually generates top-down predictions based on interpretations of prior context, and these predictions facilitate processing of incoming words through mechanisms like predictive coding, which is conceptualized as a general-purpose algorithm for approximating Bayesian inference (Clark, 2013).<sup>8</sup>

This dichotomy of load localization resonates with long-standing debates and psycholinguistics and philosophy regarding modularity versus interactivity and language processing. Modular accounts, following Fodor (1983), propose that language processing operates in informationally encapsulated modules with limited access to information from other cognitive systems. Interactive accounts by contrast, assume that linguistic processing is immediately continuously influenced by all available information, including domain general constraints from attention, memory and world knowledge (McClelland & Rumelhart, 1986). While the modularity debate has traditionally focused on the flow of information during processing, the lexicon-internal/domain-general framework refocus attention on the nature and origin of the mechanisms that constrain processing under load.

Critically these two types of mechanisms may produce different signatures in empirical data. Lexicon-internal limitations may manifest as specific effects on particular linguistic processes, i.e., for instance, modulating the time course of lexical competition by altering the activation dynamics among competitors. If load operates as a lexicon-internal constraint within the lexical system, we might expect to see selective changes of how activation spreads among lexical candidates: perhaps the target word maintains relatively stable activation while

---

<sup>8</sup> While predictive coding may represent a common computational principle across cognitive domains, this does not necessarily entail that these domains share processing resources. A stronger claim – and one central to the present study – is that the mechanisms underlying lexical competition specifically recruit domain-general resources, such that taxing these resources through concurrent cognitive load should affect how competitors are activated or resolved.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

competitors show altered dynamics (reduced lateral inhibition leading to sustained competitor activation), or the system shows changes in temporal parameters like decay rates or selection thresholds. Domain-general limitations, by contrast, might produce more global effects on task performance – an overall reduction in activation levels across all lexical candidates due to reduced processing resources, the introduction of noise that affects targets and competitors alike, or general slowing or inability of processing. Such domain-general effects would not selectively target lexical competition mechanisms but would instead reflect constraints imposed by limited attentional or working memory resources that language processing must compete for alongside other cognitive demands.

Distinguishing between these possibilities requires converging evidence from multiple methodologies. The present studies leverage physiological measures, computational modeling, and behavioral measures to investigate how phonological load impacts the lexical system across modalities, and whether it operates as a lexicon-internal constraint or as an domain-general limitation on domain-general resources.

### **Purpose of Dissertation**

This dissertation integrates these lines of research to investigate how phonological load influences lexical processing in both perception and production. While previous work has established that cognitive load can modulate lexical competition, fundamental questions remain about the mechanisms underlying these effects. I addressed these questions using three complementary approaches. First, I used eye tracking data during spoken word recognition to examine whether phonological load affects the temporal dynamics of lexical competition as measured through the visual world paradigm. Second, I employed computational modeling with

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

TRACE to explore whether load acts as a lexicon-internal constraint or as domain-general resource limitation that affects processing more globally (e.g., by modulating activation or decay parameters). Finally, I extend our investigation to speech production to test whether load affects generalize across modalities, and whether the pattern of results is more consistent with lexicon-internal or domain-general mechanisms. Together these studies provide a mechanistic account of how cognitive interfaces with lexical access and selection with implications remodels of spoken language processing under resource constraints.

**Eye can't believe it's not butter: Phonological load and visual word recognition**

While Zhang and Samuel's (2015, 2018) results clearly and importantly demonstrate the effects of load on lexical activation, they arguably don't provide enough information to articulate a revised theory of lexical access. It's unclear, for example, whether phonological load interferes with all words that become active during comprehension or solely non-target competitors (e.g., when hearing /pik/, are *peak* and *peel* both disrupted, or just *peel*?). It is also unclear at what point phonological load interferes with competitors after their initial activation. It's also unclear whether this interference is constant (e.g., competitors lose activation at a consistent pace); logically speaking it could be non-linear (e.g., a sigmoidal drop off of activation). The extent of the interference is also unclear; while the interference is enough to render priming non-significant, it's not clear whether the activation of competitors is brought entirely to 0.

In Study 1, I adapted McMurray, Farris-Trimble, and Rigler's (2017) visual world paradigm (VWP) study to collect fine-grained information about how phonological load affects lexical access. The key adaptation of their study is the introduction of a between-subjects comparison: Participants either completed the VWP task under Sham load conditions (selecting a red letter) or while performing a letter recognition phonological load task. This design will conceptually replicate the load manipulation used in Zhang and Samuel (2018) with the VWP paradigm and data.

I hypothesized that participants in the Sham load condition should elicit similar eye fixation patterns as the original study. Figure 1 was taken directly from their study and can be used as a benchmark: Both the Target & Competitors (Cohort, Rhyme, and Unrelated) were

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

activated early on, and as more input was received, Target gains prominence while Competitors lose activation.

In order to predict what would happen under phonological load, I translated Zhang & Samuel's (2018) results to the VWP. From a mechanistic perspective, the non-significant priming effects that they observed can be most straightforwardly understood as the primes having reduced activation in the Phon load vs. Sham load condition. In the VWP, this would be indexed by fewer fixations to the competitors (Cohort and Rhyme) and faster declines in their fixations over time, relative to the Sham condition. Zhang and Samuel do not report a difference in priming when the carrier and prime were one and the same (e.g., when *nap* is presented auditorily instead of *napkin*), suggesting that a reduction of fixations in the Phon condition may only be observed in competitor words, not the target (e.g., when hearing *peek*, there may be fewer fixations to *peel* and *beak* but not to *peek* itself).

Using both the fixation patterns and other behavioral measures (e.g., response times, accuracy), Study 1 overarchingly gave us insight into how the activation dynamics for Target words and Competitors unfold over time, and whether these dynamics differ for Competitors with varying degrees of similarity (e.g., Cohort, Rhyme, Unrelated) to the Target word.

### **Method**

#### ***Participants***

The McMurray et al. (2017) study recruited 18 participants for each of their participant groups. Forty participants were recruited for this study from the Tufts University SONA Pool. Four participants' data were excluded from all analyses due to not completing the study accurately and to completion. This left 36 participants – 18 in the Sham load condition and 18 in

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

the Phon load condition for all analyses. All were native English speakers with no (self-reported) history of learning or language disability, and no auditory impairments. They all also provided informed consent following the guidelines of the SBER Institutional Review Board (IRB) and were compensated for their time with research credit. They all also had to consent to having their eyes be tracked throughout the course of the study. A participant could not participate in this study if they took part in Study 3. Lastly, each participant only sat in one study session which took approximately 90 minutes.

### *Design*

The design for both Studies 1 and 3 is as follows: The experiment consists of both a primary and secondary task. The primary task was the main language task I was interested in (in Study 1: the VWP task; in Study 3: the word production task). The secondary task was the load task. There were two types of load conditions, and participants got randomly assigned to either one of them. This dual task paradigm was implemented to resemble the Zhang and Samuel (2018) load manipulation; each trial of the experiment consisted of one trial of the primary task and one trial of the secondary task interweaved with each other (see Figure 2). Although Zhang and Samuel had No load and Phon load comparison groups, I opted to using a Sham load condition in place of the No load – this was primarily to keep the pacing of both experimental conditions the same. The pacing of linguistic tasks has been found to affect priming and other linguistic processing effects (e.g., priming studies with varying stimulus onset asynchronies (SOAs), such as Branigan et al., 2000; Tooley, Konopka, & Watson, 2018), suggesting differences in processing timing (particularly in production) could introduce confounds beyond the load manipulation itself.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

For the Sham load condition, a trial proceeded as follows: Participants first saw the stimulus presentation of the load task – this was four black letters arranged on screen so that each letter was in its own quadrant (see Figure 3). Then, the primary task (VWP) began. After the participant responded to the primary trial, or if they failed to respond within 3000 ms, the second part of the load task automatically commenced. For the Sham load task, participants saw the same four letters from the beginning of the trial on screen and had to select the letter that was now colored red (the remaining letters (3/4) maintained their black color).

For the Phon load condition, a trial will proceed as follows: Participants first saw the stimulus presentation of the load task (see Figure 3). Then, the primary task (VWP) began. After the participant responded to the primary trial, or if they failed to respond within 3000 ms, the second part of the load task automatically commenced. For the Phon load task, participants saw one letter at the center of the screen and had to respond YES or NO to whether that letter was in the sequence of four letters they saw previously in the stimulus presentation at the start of the trial.

### ***Materials***

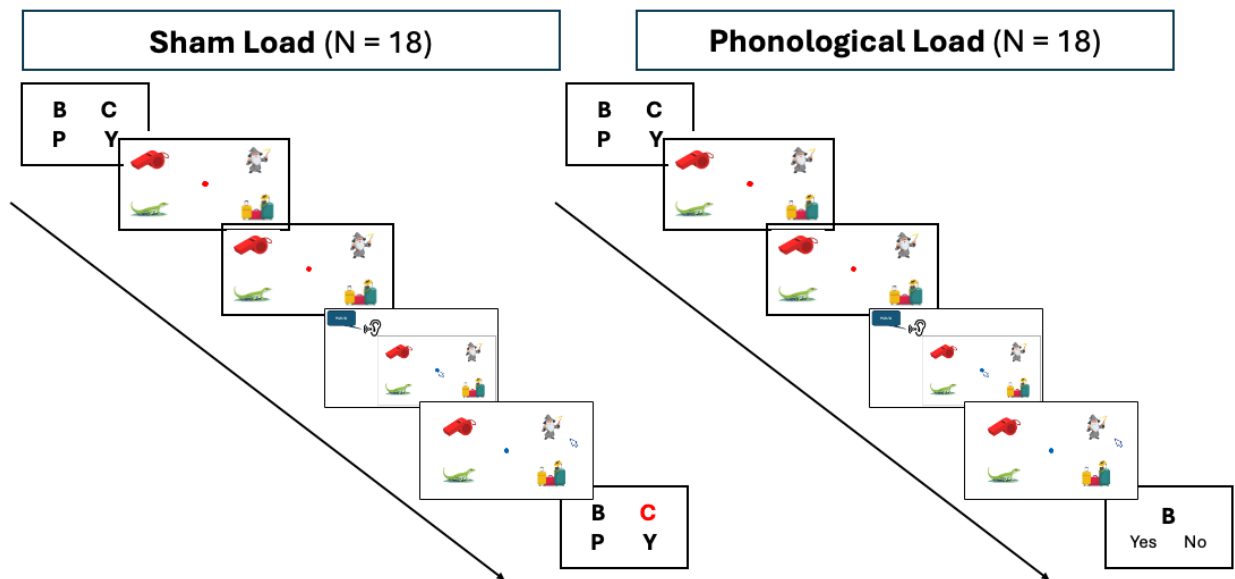
For the spoken word recognition task, 60 sets of four words were used. Each set contained a target (e.g., *rocket*), an onset/Cohort competitor (*robin*), an offset/Rhyme competitor (*pocket*), and an Unrelated word (*castle*). There were 30 monosyllabic sets and 30 disyllabic sets. These words and images of them were obtained from Bob McMurray, are easily picturable, and were piloted beforehand to ensure that they were readily understood. This stimuli list included the words used in McMurray et al. (2017) in addition to other words that have been used in other

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

studies in their laboratory. The auditory stimuli were also obtained from McMurray and were recorded from a monolingual female American English speaker in a natural cadence.

**Figure 3**

*A Single Trial of Study 1 by Load Condition*



*Note.* After selecting the center fixation dot, the auditory cue was played. In this trial, that auditory cue was “wizard.”

Each word in a set was a Target, which served to control any frequency, density, imageability, etc. effects between the Target and Competitors. As in the original study, there were four different trial-types made from a single set of four words (see Table 1). When the base word (*rocket*) is the target stimulus, there was a Cohort (*robin*), Rhyme (*pocket*), and Unrelated (*castle*) item (a ‘TCRU’ trial). When *robin* was the Target, there was a Cohort (*rocket*) and two Unrelated words (*pocket* and *castle*) (‘TCUU’ trial). When *pocket* was the Target, there was a

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Rhyme (*rocket*) and two Unrelated words (*robin* and *castle*) ('TRUU' trial). Lastly, when *castle* was the Target, all three competitors were Unrelated (*rocket*, *robin*, and *pocket*) (TUUU trial).

**Table 1**

*Trial types for Study 1*

<b>Trial Type</b>	<b>Target</b>	<b>Cohort / Unrelated</b>	<b>Rhyme / Unrelated</b>	<b>Unrelated</b>
<b>TCRU</b>	rocket	robin	pocket	castle
<b>TCUU</b>	robin	rocket	pocket	castle
<b>TRUU</b>	pocket	robin	rocket	castle
<b>TUUU</b>	castle	rocket	robin	pocket

*Note.* Each word in a set could serve as the target.

Each set appeared in the experiment 8 times. This comes from there being 4 different trial types which were presented 2 times across the course of the experiment. With 60 sets of words, this yielded 480 trials (60 sets x 8 repetitions of each set).

Participants sat in a dark, semi-sound-attenuated room, in front of a table with their head supported by a chin rest. The experimental program was written in PsychoPy (Peirce et al., 2022). Eye position was monitored and recorded at 1000 Hz using an infrared eye tracking system (Eyelink, SR Research, Ottawa, Canada).

### ***Procedure***

The experiment proceeded as follows: The eye tracker was first sanitized, adjusted, and calibrated for each new participant. Participants first saw the load task stimulus – four black

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

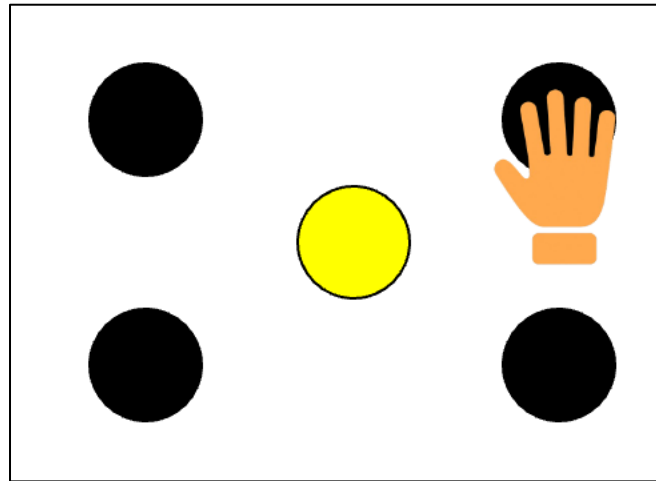
letters appearing in each quadrant – on the monitor. Once that passed, the primary task began. Four images appeared on screen, each within its own quadrant of the monitor. At the onset of each trial, the pictures were displayed with a small red dot in the center. After 500 ms, the dot turned blue, and the participant was instructed to fixate on the dot in order to initiate the auditory cue. Then, after hearing the cue, participants selected the picture that matched the word the auditory cue corresponded to by clicking on the button corresponding to that picture's position on screen (See Figure 4). This is different from McMurray and colleagues' (2017) protocol – in their study, participants used a computer mouse to click on the picture. After the participant responded to the primary task, or if they failed to respond within 3000 ms, the second part of the load task commenced. A single trial consisted of an iteration of both the primary and secondary tasks.

The secondary task was inspired by the Zhang & Samuel (2018) letter recognition task. If participants were in the Sham load condition, they saw a fixation cross appear on screen for 500 ms, followed by the same four letters they saw at the start of the trial. They were asked to select the red letter using the buttons corresponding to each letter's position on screen.

Alternatively, if participants were in the Phon load condition, they saw a letter (e.g., *B*) on screen and were asked to decide whether this letter was in the string of letters they saw at the start of the trial by clicking on the lower two buttons, corresponding to YES (bottom right) and NO (bottom left). The next trial began 1000 ms after the response, or 3000 ms after stimulus onset if the participant did not respond.

**Figure 4**

*Response Mechanism in Study 1*



*Note:* Participants in Study 1 responded using 5 buttons – 4 black buttons and 1 yellow button (coined “Gothic Daisy” by the Multisensory Perception Laboratory at Tufts University).

Participants were instructed to keep their hands in quick reach of the four black buttons – the yellow button was used only to proceed from the instructions, break times, and eye tracker calibration screens. The black buttons were used to select the image corresponding to the position on the computer monitor. The bottom two black buttons had a dual function – in the Phon load task, they acted as YES (bottom right) and NO (bottom left) responses.

## Results

### *Data Preprocessing*

There were 8,640 trials of behavioral data each for the Sham and Phon load conditions (18 participants x 480 trials = 8640 for each condition). Trials with missing and/or incorrect responses across the load and VWP tasks were removed from further analysis (7% of data).<sup>9</sup>

---

<sup>9</sup> There were 264 trials out of 8640 (3%) of Sham load data with missing/incorrect data in either the VWP or Load task. There were 980 (11%) of 8640 trials in the Phon load conditions that were missing/incorrect.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

There were over 30 million eye fixation data points recorded across both load conditions. Eye-tracking data were preprocessed following standard procedures for Visual World Paradigm studies (McMurray et al., 2010).<sup>10</sup>

Fixations were analyzed within four regions of interest (ROIs): target, the competitors, and the unrelated distractor. The time window for analysis spanned from the onset of the auditory cue – this allowed us to time lock trial onset (0 ms) to the stimulus. Each participants' data was analyzed within a window of time bounded by the onset of the auditory stimulus and their average reaction time (RT) calculated across all trials, as eye movements after the manual response reflect post-decisional processes rather than lexical competition dynamics. This resulted in individual time series of varying lengths depending on each participant's response speed. To account for trial-to-trial variability when aggregating data, trials with RTs faster than the participant's mean RT were extended by carrying forward their final fixation ROI through the mean RT. Conversely, trials with RTs slower than the mean were truncated at the mean RT, holding their fixation constant at the value observed at the time of response. For each trial, I calculated the proportion of time spent fixating each object (Target, Cohort, Rhyme, Unrelated). This procedure resulted in standardized time series for each participant while preserving the dynamics of lexical competition up to the point of response (Farris-Trimble & McMurray, 2013; McMurray et al., 2010).

Trials in which 1) participants did not respond to either task, 2) clicked on an incorrect object, or 3) answered incorrectly on the load task were excluded from analysis. This was because I only wanted to analyze trials where participants were clearly attending to both tasks and performing correctly.

---

<sup>10</sup> Portions of the data preprocessing were assisted by AI tools (ChatGPT and Claude.ai). All conceptual decisions, statistical choices, and interpretations were made by us.

*Analysis*

**Load task analyses.** Descriptive statistics were calculated for both the average load task response time (RT) and load task accuracy (see Table 2). Participants demonstrated higher accuracy in the Sham condition ( $M = 99.5\%$ ,  $SD = 0.5$ ) compared to the Phon condition ( $M = 91.2\%$ ,  $SD = 5.2$ ). Similarly, reaction times were faster in the Sham condition ( $M = 460$  ms,  $SD = 600$  ms) than in the Phon condition ( $M = 890$  ms,  $SD = 100$  ms). These values are consistent with Zhang and Samuel (2018), who reported Phon load task accuracy and RT to be  $M = 0.89$ ,  $SD = 0.10$ ; and  $M = 842$  ms,  $SD = 154$ , respectively.

**Table 2**

*Study 1 Average RT<sup>11</sup> and Accuracy in Primary and Secondary Tasks*

	Secondary (Load) Task		Primary Task	
	Mean Accuracy $\pm$ SD (%)	Avg RT $\pm$ SD (ms)	Mean Accuracy $\pm$ SD (%)	Avg RT $\pm$ SD (ms)
Sham	99.5 $\pm$ 0.5	460 $\pm$ 600	98.3 $\pm$ 1.3	1050 $\pm$ 100
Phon	91.2 $\pm$ 5.2	890 $\pm$ 100	98.7 $\pm$ 0.9 <sup>12</sup>	1130 $\pm$ 100

Load task performance was analyzed using generalized linear mixed-effects models (GLMMs) for accuracy and linear mixed-effects models (LMMs) for reaction time. Accuracy was modeled using a logistic mixed effect model. Reaction times were log-transformed to meet assumptions of normality and analyzed only for correct responses. Both models included load

<sup>11</sup> RT = reaction time for correct responses in load and VWP task responses.

<sup>12</sup> Zhang & Samuel (2018) reported mean accuracy for their primary priming task at  $M = 0.93$ ,  $SD = 0.1$ . This is in comparison to their non-phonological load task manipulation, which yielded priming task accuracy at  $M = 0.92$ ,  $SD = 0.01$ .

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

condition (Phon vs. Sham) as a fixed effect and participant as a random intercept to account for individual differences. Models were fit using the lme4 package (Bates et al., 2015) in R (Version 4.2.3; R Core Team, 2024), with p-values for the LMM obtained using Satterthwaite's approximation via the lmerTest package (Kuznetsova et al., 2017).

The logistic model revealed a significant effect of load condition on load task accuracy, with participants showing significantly higher accuracy in the Sham condition compared to the Phon condition ( $\beta = 3.03$ ,  $SE = 0.26$ ,  $z = 11.48$ ,  $p < .001$ ). A model examining reaction times for correct responses showed that participants were significantly faster in the Sham load task compared to the Phon load task ( $\beta = -0.63$ ,  $SE = 0.04$ ,  $t(34.0) = -14.29$ ,  $p < .001$ ). These findings indicate that the Phon load task was more cognitively demanding than the Sham task, as evidenced by both reduced accuracy and slower response times.

**VWP Analyses.** I also examined the effects of trial type and load condition on VWP task accuracy. Accuracy was modeled using a logistic mixed effect model. The model revealed a significant main effect of trial type, with TCRU trials showing significantly lower accuracy compared to the TUUU baseline ( $\beta = -0.97$ ,  $SE = 0.25$ ,  $z = -3.89$ ,  $p < .001$ ). This indicates that participants were less likely to respond correctly to TCRU trials compared to TUUU trials. TCUU and TRUU trial types also showed numerically lower accuracy than TUUU, though these differences did not reach statistical significance (TCUU:  $\beta = -0.32$ ,  $SE = 0.28$ ,  $z = -1.13$ ,  $p = .258$ ; TRUU:  $\beta = -0.44$ ,  $SE = 0.27$ ,  $z = -1.63$ ,  $p = .103$ ). This pattern aligns with lexical competition accounts of spoken word recognition. TCRU trials – containing both cohort and rhyme competitors – showed the greatest accuracy cost, likely because competition occurred throughout word processing (early from cohort overlap, late from rhyme overlap). Single-

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

competitor conditions (TCUU and TRUU) showed intermediate effects, suggesting cumulative interference when multiple competitor types are present.

Critically, there was no significant main effect of load condition on VWP accuracy ( $\beta = 0.65$ ,  $SE = 0.42$ ,  $z = 1.55$ ,  $p = .121$ ), indicating that participants performed similarly on the VWP task regardless of whether they were under Phon or Sham load. Additionally, none of the trial type  $\times$  load condition interactions reached significance (all  $p > .197$ ), suggesting that the effect of Competitor type on accuracy did not differ between load conditions. These findings indicate that phonological working memory load did not significantly impair participants' ability to identify target words in the presence of lexical competitors.

A linear mixed effects model examining log-transformed reaction times for correct responses revealed significant effects of both trial type and load condition. Models included load condition (Phon vs. Sham) as a fixed effect and participant as a random intercept to account for individual differences. All competitor trial types showed significantly slower RTs compared to the TUUU baseline: TCRU ( $\beta = 0.07$ ,  $SE = 0.01$ ,  $t(16787) = 8.39$ ,  $p < .001$ ), TCUU ( $\beta = 0.09$ ,  $SE = 0.01$ ,  $t(16787) = 11.37$ ,  $p < .001$ ), and TRUU ( $\beta = 0.05$ ,  $SE = 0.01$ ,  $t(16787) = 6.85$ ,  $p < .001$ ). This corresponds to slower responses when lexical competitors were present compared to trials with only unrelated items.

Critically, there was a significant main effect of load condition ( $\beta = 0.09$ ,  $SE = 0.04$ ,  $t(36) = 2.25$ ,  $p = .030$ ), indicating that participants responded approximately 9% slower under Phon load compared to Sham load. The trial type  $\times$  load condition interactions were not significant (all  $p > .174$ ), suggesting that the slowing effect of phonological load was consistent across all trial types and did not differentially impact lexical competition resolution. These findings indicate

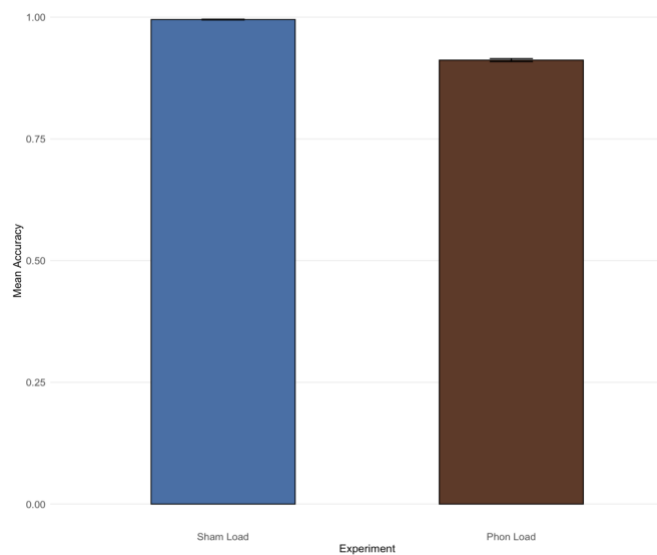
## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

that phonological working memory load globally slowed response times and did not selectively interfere with processing when phonologically similar competitors were present.

Together, the results suggest that phonological working memory load produced general cognitive slowing, as evidenced by the general slowing of RTs under load. However, this interference was not sufficient to compromise the accuracy of lexical selection. While participants were slower under Phon load, they maintained comparable accuracy levels and showed similar patterns of lexical competition effects (i.e., difficulty with Cohort competitors) across both load conditions.

### Figure 5

*Study 1 - Mean accuracy across load conditions*



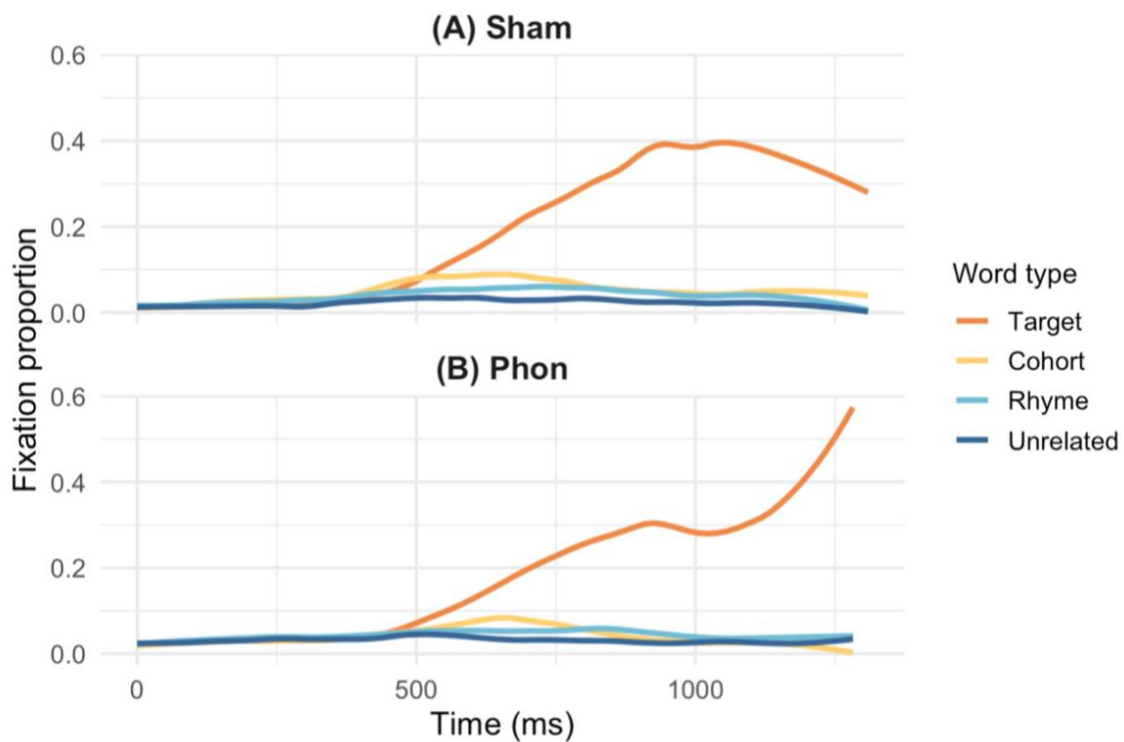
*Note:* Mean accuracy on the load task by condition. Error bars represent standard error of the mean. Participants demonstrated significantly higher accuracy in the Sham condition compared to the Phon condition ( $p < .001$ ).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Fixation Analyses.** Figure 6 displays fixation proportions for each word type under Sham (A) and Phon (B) load conditions in the TCRU trial type. In both conditions, target fixations show a similar pattern, rising gradually from approximately 500 ms onward. Under Sham load, target fixation proportions peak around 1000-1100 ms at approximately 40% (Fig. 6A). Under Phon load, they show a comparable rise ultimately, reaching approximately 60% (Fig. 6B). Interpretation of fixation patterns near the mean RT boundary should be made with caution, since these time points reflect fewer observations due to the truncation and extension procedure. Cohort, rhyme, and unrelated competitors show minimal activation across both conditions, remaining near or below 10% fixation proportion throughout the time course.

**Figure 6A & B**

*Proportion of fixations over time (TCRU trials)*



## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Notably, fixation proportions to targets remain relatively modest compared to prior VWP studies (McMurray et al., 2010, 2017), likely reflecting the fact that fixations in the current study were not guided by a mouse-based response mechanism (see Discussion for further elaboration). There was also substantial variation in the fixation patterns, with some participants' peak target fixations reaching  $\sim 0.7$ , while others only got to 0.2 (see Appendix for visualization). This could be attributed to the lack of visually guided responses (because of the button response), and participants' engaging in different strategies to maintain high performances across both tasks.

To quantify differences between stimulus types across load, I used a non-linear curve fitting technique developed in prior studies (Farris-Trimble & McMurray, 2013; McMurray et al., 2010; Mirman et al., 2008; Rigler et al., 2015). This approach addresses challenges inherent to analyzing eye-tracking data. First, there is considerable inter-subject variability in reaction times, peak fixation proportions, and the overall time course of lexical processing; fitting curves to individual participants' data and extracting parameters provides a principled way to characterize each subject's processing dynamics. Second, averaging across participants can obscure or even artifactually create effects—for instance, variability in the timing of fixation curves could average to yield what appears to be a difference in slope (McMurray et al., 2010). By obtaining parameters from individually fitted curves, this approach avoids such averaging artifacts. Finally, extracting curve parameters allows for a direct comparison of theoretically meaningful parameters – such as the timing of competitor resolution or the rate of target activation – across load conditions.

I divided the data for each competitor analysis by trial type: Target analyses used data in which the Target competitor was present (i.e., TCRU, TCUU, TRUU trials); Cohort analyses used data in which the Cohort competitor was present (i.e., TCRU and TCUU trials); Rhyme

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

analyses used data in which the Rhyme competitor was present (i.e., TCRU and TRUU trials); and Unrelated analyses used trials in which an Unrelated item was present (i.e., TCRU, TCUU, and TRUU trials). TUUU trials were not included in these analyses because trials with no phonological competitors may behave quite differently to trials with some (McMurray et al. 2017).

*Target Fixation Analyses.* Instead of just comparing raw fixation proportions at each time bin, curve fitting summarizes each participant's fixation trajectory with parameters (timing, peak, slope, etc.). To characterize the time course of target fixation, I fit a four-parameter logistic function to each participant's data:

### Equation 1

$$P(\text{target}) = \frac{p - b}{1 + \exp\left(4 \times \frac{s}{p - b} \times (c - t)\right)}$$

where  $t$  is time in milliseconds,  $b$  is the baseline fixation proportion,  $p$  is the peak (asymptotic) fixation proportion,  $c$  is the crossover time (inflection point at which fixations reach the halfway point between baseline and peak) and  $s$  is the slope parameter reflecting the steepness of the rise.

To improve model convergence, I calculated participant-specific starting values for each parameter:

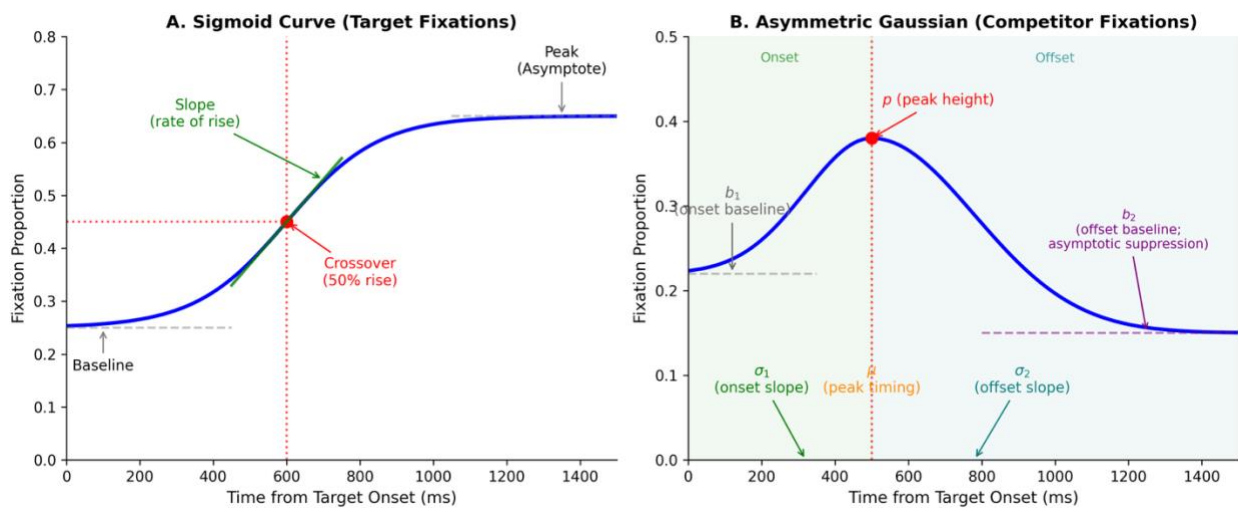
- Baseline ( $b$ ): Mean target fixation proportion in the first 200 ms
- Peak ( $p$ ): The maximum of either (1) 90% of the observed maximum fixation proportion, or (2) the mean fixation proportion in the final 20% of the time window

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Crossover ( $c$ ): The timepoint at which fixations reached the halfway point between baseline and peak
- Slope ( $s$ ): Fixed at 0.002 for all participants, as derived slope estimates proved unstable and led to a lack of convergence, probably due to the relationship between slope and the ( $p-b$ ) term in the denominator.

**Figure 7**

*Illustration of curve fitting parameters*<sup>13</sup>



Note: (A) Sigmoid curves were fit to target fixation proportions, yielding parameters for baseline, peak, crossover, and slope. (B) Asymmetric Gaussian curves were fit to competitor fixation proportions, with separate parameters for the onset phase ( $b_1$ ,  $\sigma_1$ ) and offset phase ( $b_2$ ,  $\sigma_2$ ), as well as  $\mu$  and  $\rho$ .

<sup>13</sup> Figure created with assistance from Claude (Anthropic, 2025). I input the formulas and parameters to the LLM and asked for a figure that illustrates what the parameters control in both types of curves.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Models were fit using the Levenberg-Marquardt algorithm (nlsLM function from the minpack.lm package in R: Elzhov et al., 2023; Moré, 1978). I assessed model fit using  $R^2$  (which measures the proportion of the variance explained by the model) and root mean squared error (RMSE) (which measures the average magnitude of prediction errors). Participants were excluded from the target fixation analyses if the model failed to converge, or if  $R^2 < 0.65$  (poor fit quality). I then compared Phon and Sham groups on these key curve parameters using linear regressions.

The four-parameter logistic function provided excellent fits to individual participant data (mean  $R^2 = .93$ ,  $SD = .08$ ; mean RMSE = .029,  $SD = .01$ ). After applying the exclusion criteria, 33 participants were retained for analysis ( $N_{\text{Phon}} = 16$ ,  $N_{\text{Sham}} = 17$ ; 3 participants removed).

**Table 3**

*Results of analysis of logistic curve fits on target fixations*

	M (SD)		$t(32)$	$p$
	Sham	Phon		
Crossover	691 (76.6)	695 (62.6)	-0.184	.855
Slope	0.001 (0.0005)	0.001 (0.0007)	0.027	0.979
Peak	0.425 (0.203)	0.391 (0.216)	0.473	0.64

Table 3 presents the mean (and standard deviation) parameter estimates by load group, as well as the results of the linear regression testing for a significant difference as a function of group. Linear models revealed no significant differences between Phon and Sham groups on any

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

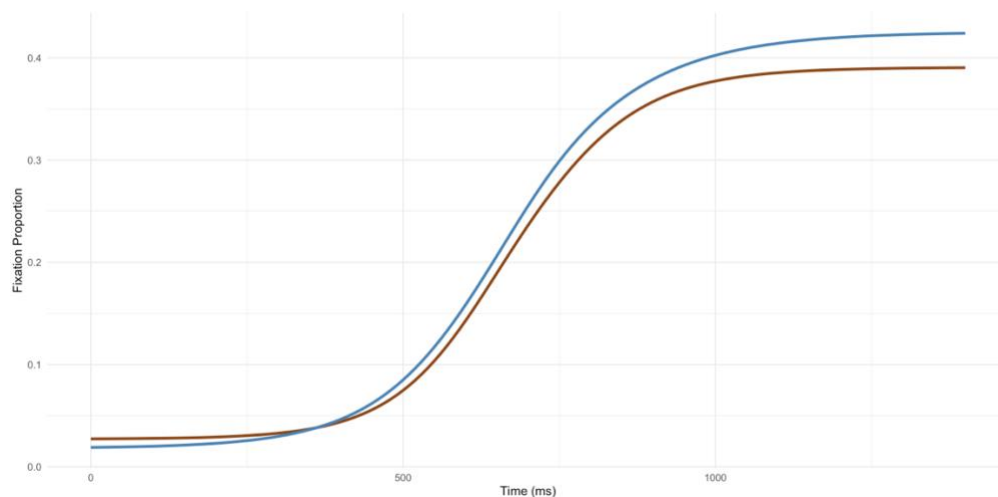
parameter. These results suggest that phonological working memory load did not affect the speed, efficiency, or magnitude of target word recognition.

To quantify the evidence for the null hypothesis, I computed Bayes factors using the BayesFactor package in R (Morey et al., 2018). Bayes factors were interpreted using the classification scheme proposed by Lee and Wagenmakers (2013), where  $BF_{01}$  values between 1-3 indicate anecdotal evidence, 3-10 indicate moderate evidence, and  $>10$  indicate strong evidence for the null hypothesis.

All three parameters showed anecdotal-to-moderate evidence for the null (crossover:  $BF_{01} = 2.97$ ; slope:  $BF_{01} = 3.01$ ; peak:  $BF_{01} = 2.76$ ). These results suggest the data are approximately three times more likely under the null than the alternative hypothesis, though this level of evidence is not strong enough to conclusively support the absence of an effect.

### Figure 8

#### *Average Fitted Target Curves*



*Note.* Sham load = blue; Phonological load = maroon

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

*Competitor Fixations Analyses.* Competitor fixations (Cohort, Rhyme, and Unrelated) were analyzed separately using an asymmetric Gaussian function:

### Equation 2

$$P(t) = \begin{cases} \exp\left(\frac{(t-\mu)^2}{-2\sigma_1^2}\right) (p - b_1) + b_1 & \text{if } t \leq \mu \\ \exp\left(\frac{(t-\mu)^2}{-2\sigma_2^2}\right) (p - b_2) + b_2 & \text{if } t > \mu \end{cases}$$

where  $\mu$  is the time of peak competitor fixation,  $p$  is the peak height,  $\sigma_1$  controls the onset slope (rise to peak),  $b_1$  is the onset baseline,  $\sigma_2$  controls the offset slope (decline from peak), and  $b_2$  is the offset baseline (asymptotic suppression level).

For each competitor analysis, I compared the Phon and Sham groups on peak time ( $\mu$ ), peak height ( $p$ ), onset slope ( $\sigma_1$ ), onset baseline ( $b_1$ ), offset slope ( $\sigma_2$ ), and offset baseline ( $b_2$ ) using regressions. The asymmetric Gaussian function provided good fits to Cohort fixation data (mean  $R^2 = .907$ ,  $SD = .079$ ; mean  $RMSE = .007$ ,  $SD = .003$ ). After applying the exclusion criteria, 32 participants were retained for analysis (14 Phon, 18 Sham; 4 removed). Group comparisons revealed a significant difference in onset baseline  $b_1$ : Phon  $M = .029$  ( $SD = .017$ ), Sham  $M = .016$  ( $SD = .017$ ),  $t(31) = -2.17$ ,  $p < .05$ , indicating that participants in the Phon condition showed more early Cohort competitor fixations. This is probably due to noise in the early time window - McMurray approximated  $b_1$  to be zero, as participants have not yet heard the auditory stimulus and should not be able to anticipate which competitor types will appear. Since this pattern does not persist across other competitor conditions, it likely reflects random noise rather than a meaningful effect and was not looked into further. No other parameters differed significantly between groups (see Table 4 for all competitor parameter comparisons).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Bayesian analyses were conducted to quantify the evidence regarding the effect of cognitive load on competitor fixation dynamics. For  $\mu$  (peak timing), peak proportion, and  $\sigma_1$  (rise rate), Bayes factors provided anecdotal evidence for the null hypothesis ( $\mu$ :  $BF_{01} = 2.94$ ; peak:  $BF_{01} = 2.84$ ;  $\sigma_1$ :  $BF_{01} = 2.94$ ), suggesting the data were approximately three times more likely under the null than the alternative. In contrast,  $b_1$  (rise baseline) showed anecdotal evidence for the alternative hypothesis ( $BF_{10} = 1.91$ ), again, probably due to noise in the early time window.

Rhyme competitor fixations were also well-fitted by the asymmetric Gaussian function (mean  $R^2 = .872$ ,  $SD = .089$ ; mean RMSE = .006,  $SD = .002$ ). After applying the exclusion criteria, 27 participants were retained for analysis (12 Phon, 15 Sham). No significant differences emerged between load groups on any parameter (all  $p > .05$ ).

Bayesian analysis of rhyme competitor fixation dynamics revealed a consistent pattern of anecdotal-to-inconclusive evidence for the null hypothesis across all parameters.  $\mu$  (peak timing:  $BF_{01} = 2.73$ ), peak proportion ( $BF_{01} = 2.52$ ),  $b_1$  (rise baseline:  $BF_{01} = 2.10$ ), and  $b_2$  (fall baseline:  $BF_{01} = 2.24$ ) all showed anecdotal evidence for the null.  $\sigma_1$  (rise rate:  $BF_{01} = 1.48$ ) and  $\sigma_2$  (fall rate:  $BF_{01} = 1.30$ ) were essentially inconclusive, providing no meaningful evidence in either direction.

Unrelated item fixations showed good model fits (mean  $R^2 = .898$ ,  $SD = .082$ ; mean RMSE = .006,  $SD = .003$ ). After applying the exclusion criteria, 29 participants were retained for analysis (15 Phon, 14 Sham). Consistent with Rhyme analyses, no significant group differences were observed for any parameter (all  $p > .05$ ).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Table 4***Results of analysis of asymmetric Gaussian parameters on competitor fixations*

		M (SD)		<i>t</i> (31)	<i>p</i>
		Sham	Phon		
Cohort	Peak time	652 (123)	647 (91.8)	.108	.915
	Peak	.097 (.071)	.104 (.052)	-3.29	.744
	Onset slope	153 (104)	147 (112)	.134	.894
	Offset slope	343 (366)	179 (237)	1.45	.157
	Onset baseline	.016 (.017)	.029 (.017)	-2.17	.038*
	Offset baseline	.056 (.116)	.028 (.017)	.895	.378
Rhyme	Peak time	653 (108)	664 (127)	-.231	.819
	Peak	.074 (.045)	.081 (.024)	-.512	.613
	Onset slope	163 (122)	271 (289)	-1.31	.201
	Offset slope	292 (307)	157 (112)	1.48	.160
	Onset baseline	.020 (.018)	.026 (.018)	-.869	.393
	Offset baseline	.027 (.030)	.036 (.026)	-.767	.450
Unrelated	Peak time	639 (152)	593 (158)	.795	.434
	Peak	.100 (.072)	.097 (.039)	.153	.88
	Onset slope	234 (219)	166 (158)	.969	.341
	Offset slope	153 (91.6)	225 (161)	-1.471	.153
	Onset baseline	.036 (.035)	.037 (.039)	-.101	.920
	Offset baseline	.038 (.029)	.040 (.031)	-.157	.876

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Bayesian analysis of unrelated distractor fixation dynamics showed anecdotal evidence for the null hypothesis across most parameters.  $\mu$  (peak timing:  $BF_{01} = 2.26$ ), peak proportion ( $BF_{01} = 2.84$ ),  $b_1$  (rise baseline:  $BF_{01} = 2.86$ ),  $\sigma_1$  (rise rate:  $BF_{01} = 2.02$ ),  $b_2$  (fall baseline:  $BF_{01} = 2.84$ ), and  $\sigma_2$  (fall rate:  $BF_{01} = 1.29$ ) all showed inconclusive to anecdotal evidence for the null.

**Exploratory Participant Grouping Analyses.** Given the substantial variability observed in fixation patterns (See Appendix A4), I conducted an exploratory analysis to examine whether participants could be meaningfully partitioned based on their target fixation behavior. I classified participants based on whether their peak target fixations reached a threshold of .25 (i.e., 25% of fixations directed toward the target at peak).<sup>14</sup> This resulted in 24 participants classified as “above-or-at-chance target” (peak  $\geq 0.25$ ) and 12 participants classified as “below-chance target” (peak  $< 0.25$ ).

I first examined whether participants with below-chance target fixations showed qualitatively different patterns of lexical engagement. I calculated peak fixation proportions for target, cohort, rhyme, and unrelated items for each participant, then computed target-to-competitor ratios and a "competition index" (target peak divided by average competitor peak, where average competitor = [cohort + rhyme + unrelated] / 3).

Comparing the 12 participants with below-chance target fixations to the 24 above-chance participants revealed significant differences in the competition index. Below-chance participants showed substantially lower competition indices ( $M = 2.22$ ) compared to above-chance participants ( $M = 3.42$ ),  $t(24.06) = 3.79$ ,  $p < .001$ . The target-to-cohort ratio showed a similar

---

<sup>14</sup> This threshold was chosen to represent whether participants' target peak fixations exceeded chance level (25% of fixations towards 1 out of the 4 images on screen).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

numerical pattern but did not reach statistical significance (below-chance:  $M = 3.04$ ; above-chance:  $M = 3.81$ ;  $t(16.83) = 1.90$ ,  $p = .075$ ).

Critically, these differences reflected not just lower ratios but fundamentally different fixation patterns. Below-chance participants showed substantially lower overall target looking ( $M = 0.18$ ) compared to above-chance participants ( $M = 0.56$ ), with relatively undifferentiated fixation proportions between targets and competitors. This pattern suggests that these participants may not have been engaging with the task, resulting in relatively random fixation pattern rather than the stimulus-driven eye movements that the VWP depends on.

While these differences could partly reflect floor effects (as lower overall fixations constrain how much smaller competitor fixations can become), the pattern suggests a qualitative difference in task engagement rather than simply noisier versions of the same underlying process. Participants with below-chance target fixations appeared to distribute their attention more equally across all four items, inconsistent with word recognition processes that should drive selective attention toward the referent.

All subsequent eye-tracking analyses were conducted on the retained sample of 23 participants who exceeded chance-level target fixations. To examine whether phonological load affected the temporal dynamics of lexical competition, I fit mathematical functions to individual participants' fixation curves following McMurray et al. (2010). Target fixations were modeled using four-parameter logistic functions, while competitor fixations (cohort, rhyme, and unrelated) were modeled using asymmetric Gaussian functions. These functions yielded excellent fits across all conditions (Target: mean  $R^2 = 0.96$ ; Cohort: mean  $R^2 = 0.93$ ; Rhyme: mean  $R^2 = 0.89$ ; Unrelated: mean  $R^2 = 0.91$ ).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

For target fixations, no curve parameters differed significantly between load conditions: crossover time (time to 50% fixation; Phon:  $M = 681$  ms, Sham:  $M = 673$  ms,  $p = .773$ ), rise slope ( $p = .730$ ), or peak fixation height ( $p = .777$ ). These findings indicate that phonological load did not affect the speed or asymptote of target word recognition. For cohort competitor fixations, no parameters showed significant differences: peak timing ( $\mu$ :  $p = .346$ ), peak height ( $p$ :  $p = .950$ ), onset slope ( $\sigma_1$ :  $p = .864$ ), onset baseline ( $b_1$ :  $p = .186$ ), offset slope ( $\sigma_2$ :  $p = .096$ ), or offset baseline ( $b_2$ :  $p = .695$ ). Similarly, rhyme competitor fixations showed no significant load effects across any temporal parameters (all  $ps > .25$ ). Unrelated competitor fixations also showed no load-related differences in activation dynamics.

The time course of target activation and competitor suppression remained comparable between Phon and Sham conditions, suggesting that phonological load produced general response slowing (as observed in the RT analyses) without disrupting the core mechanisms of lexical access and competition. This pattern held robustly among participants demonstrating clear lexical engagement (i.e., above-chance target fixations), indicating that individual differences in task engagement did not obscure or moderate the absence of load effects on lexical competition dynamics.

These results are consistent with the possibility that phonological load did not substantially alter the timing, magnitude, or resolution of lexical competition. However, Bayes factors provided only anecdotal-to-moderate evidence for the null ( $BF_{01} \approx 3$ ), indicating that the data are inconclusive: while there was no indication of load effects on competition dynamics, there was also insufficient evidence to conclusively rule out such effects. The absence of clear numerical differences between conditions, combined with the modest Bayes factors, likely reflects substantial variability across participants rather than a systematic null effect. Thus, while

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

these findings are suggestive that core lexical competition processes may be robust to phonological load, stronger conclusions await replication with larger samples or reduced individual variability.

**Permutation Clustering Analyses.** To complement the curve-fitting analyses and provide an approach commonly used in eye fixation literature to identify potential time windows of load-related differences, I conducted cluster-based permutation tests on Competitor fixation time courses (Allopenna et al., 1998; Badde et al., 2020). This nonparametric approach identifies continuous time windows where groups differ significantly while controlling for multiple comparisons across the time points in eye-tracking data (Geller, Winn, Mahr, & Mirman, 2020). Unlike curve-fitting methods that summarize fixation patterns using predefined mathematical formulas, cluster-based permutation tests directly compare observed fixations at each time point, making them well-suited for detecting unanticipated temporal patterns of group differences. The analysis proceeded as follows:

- (1) Independent samples t-tests were computed between Phon and Sham conditions at each time point
- (2) Temporally continuous clusters of time points showing uncorrected significance ( $p < .05$ ) were identified
- (3) Cluster statistics were computed by summing absolute t-values within each cluster
- (4) Group labels were randomly permuted 1,000 times at the participant level, and steps 1-3 were repeated for each permutation to generate a distribution of maximum cluster masses
- (5) Observed cluster masses were compared to the 95<sup>th</sup> percentile of the permutation distribution to determine statistical significance.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

For cohort competitors, seven clusters of time points showed uncorrected significance ( $p < .05$ ), with the largest cluster spanning 350-440 ms (cluster mass = 22.1). However, none of these clusters exceeded the significance threshold (95<sup>th</sup> percentile = 573.08). For rhyme competitors, two clusters showed significance (the largest at 470-480 ms with cluster mass = 4.19), but neither survived correction (95<sup>th</sup> percentile = 409.38). For unrelated competitors, no clusters reached significance.

These results indicate that Phon load did not produce sustained differences in Competitor fixations at any point in the time course. Even when adopting a liberal threshold, any apparent differences were small and temporally isolated. This converges with the curve-fitting results in demonstrating that Phon load did not systematically alter the temporal dynamics of lexical competition, whether characterized through curve-fitting parameters (i.e., peak timing, magnitude) or through individual comparisons of fixation proportions.

### **Discussion**

The purpose of Study 1 was to investigate how phonological load affects speech recognition seen via eye fixations in a visual world paradigm (VWP) task. By comparing Sham load and Phon load conditions, I aimed to determine whether cognitive load disrupts lexical access. I examined whether maintaining phonological information in working memory would interfere with the processing of phonologically similar spoken words.

### *Summary of Main Findings*

Across all analyses – including behavioral performance measures, curve-fitting of fixation time courses, and cluster-based permutation tests – I found no evidence that phonological working memory load disrupted the core mechanisms of lexical access and competition. While Phon load increased response times uniformly across all trial types, it did not appear to affect the speed, magnitude, or resolution of competitor activation as measured through eye-tracking. This contrasts with Zhang and Samuel's (2017) findings, which found that load disrupted the automatic activation of embedded words and their ability to prime target words.

### *Hypothesis evaluation*

The present study tested two primary hypotheses regarding how phonological working memory load would affect lexical competition dynamics in the visual world paradigm (VWP). First, I hypothesized that participants in the Sham load condition would replicate the basic pattern of lexical competition observed in prior VWP studies: Early activation of both targets and phonologically similar competitors (cohort and rhyme), followed by increasing target prominence and competitor suppression as more auditory information disambiguates the target word. This hypothesis was supported. In the Sham condition, target fixations rose steadily from auditory onset and reached peak prominence, while cohort and rhyme competitors showed activation that declined over time. Unrelated competitors remained near baseline throughout, confirming that phonological similarity drove the observed competition effects. This basic pattern is consistent with interactive activation accounts of spoken word recognition (McClelland & Elman, 1986) and demonstrates that our VWP implementation successfully captured normal lexical competition processes under minimal cognitive load.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Second, and critically, I hypothesized that phonological working memory load would disrupt this competition pattern by reducing competitor activation. Drawing on Zhang and Samuel's (2018) finding that phonological load reduces phonological priming effects, I predicted that load would diminish the initial activation of phonologically similar competitors. While the precise mechanism remains an open question – whether competitors receive less bottom-up activation, are more strongly inhibited, or are less effectively maintained – the key prediction is that cohort and rhyme competitors should show reduced activation under load, manifesting as fewer fixations to these items and faster declines in their fixation proportions relative to the Sham condition. Critically, I predicted this reduction would be specific to phonologically similar competitors (cohort and rhyme) rather than targets themselves, consistent with Zhang and Samuel's finding that load disrupts competition between similar-sounding words but not the activation of fully matching targets. Thus, upon hearing a word like "peek," I expected Phon load to produce fewer fixations to competitors like "peel" and "beak" compared to Sham load, but comparable fixations to the target "peek" itself.

This second hypothesis was not supported. I found no evidence that phonological load reduced competitor activation or altered the time course of lexical competition. Cohort and rhyme competitors showed statistically equivalent patterns of activation and suppression under Phon and Sham load conditions, with no significant differences in peak timing, peak magnitude, onset characteristics, or offset characteristics. Similarly, target activation proceeded with equivalent speed and strength across load conditions.

### *Interpretation of results*

There are a number of different ways the results could be interpreted. First, let's explore what the implications would be if we assume that the results are valid as observed and imply that phonological load does not influence lexical access in the VWP task. This finding would challenge shared resource models, which predict that phonological working memory and lexical processing compete for overlapping cognitive resources (Baddeley, 2003). Under such accounts, concurrent phonological load should disrupt lexical competition processes, particularly when phonologically similar competitors must be distinguished and resolved. My results suggest instead that lexical access and phonological working memory may rely on partially separable mechanisms, at least for the component processes reflected in competitor activation time courses.

One possible interpretation is that lexical competition operates largely through automatic, bottom-up activation spreading that is relatively insulated from executive control processes recruited by the secondary load task. According to interactive activation models of word recognition (McClelland & Elman, 1986), competition emerges naturally from mutual inhibition among co-activated lexical representations, without requiring explicit executive intervention. If this competition is primarily driven by automatic phonological-to-lexical mapping processes, it may proceed efficiently even when controlled phonological maintenance processes (i.e., rehearsal in working memory) are engaged elsewhere. One should note that if we were to adopt this stance, there would be a seeming disconnect between the present results and those of Zhang & Samuel (2018) that would need to be resolved.

Another possible interpretation is that the results that were obtained in this experiment are valid but the lack of an effect on lexical activation is the result of aspects of the task or participant strategies. Specifically, participants may have adopted task strategies that minimized

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

interference between the two tasks. It is possible, for example, that participants may have adopted a strategy for phonological load task that mitigated most of its effects on lexical processing (e.g., shifting active rehearsal to the parts of the trial that were the least necessary for the VWP task to minimize competition). The finding that load produced general RT slowing without affecting lexical competition dynamics is consistent with strategic time-sharing: participants may have allocated processing time flexibly across the two tasks, accepting slower overall performance to preserve the quality of lexical processing. This interpretation aligns with resource-allocation theories proposing that dual-task interference reflects strategic control over resource distribution rather than structural capacity limitations (Navon & Gopher, 1979). The fact that Sham vs. Phon load was a between-subjects variable may have contributed to this effect. Participants may have had the time and opportunity to adopt a strategy that allowed them to optimally perform each task.

The role of strategic processing in dual-task contexts deserves particular attention given emerging evidence that participants actively modulate their information-gathering strategies based on task demands. Recent work demonstrates that cognitive load substantially alters how individuals sample information from the environment, with increased load leading to more selective and efficiency-driven processing strategies (Liu et al., 2022; Walter & Bex, 2021). Under dual-task conditions, participants may strategically limit their visual information gathering – directing fewer fixations to task-relevant stimuli and relying more heavily on peripheral processing or auditory information – to preserve resources for the concurrent working memory task. This could explain why I observed overall lower fixation proportions compared to typical VWP studies: participants may have adopted a more resource-conserving strategy that emphasized auditory-driven lexical access over visually guided exploration. These findings

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

suggest that lexical access and phonological working memory may rely on partially separable cognitive mechanisms, or that listeners can strategically allocate resources to preserve lexical processing quality at the expense of overall processing speed.

Shifting perspectives, I would like to now consider the possibility that phonological load does in fact disrupt lexical access to the extent that it could be observed in the VWP task. The simplest possibility along these lines is that the phonological load task was not difficult enough to disrupt lexical activation.

I would like to spend some time examining another possibility along these lines – that the phonological load task used in this experiment is in fact be sufficient to measurably disrupt lexical access but that other aspects of the VWP task used in this experiment made it not possible to do so. One important difference in the design of this experiment is that participants used a button pad to select the image in the VWP task rather than using a mouse to click on the on-screen image corresponding to the target. I decided to bypass mouse usage because I anticipated that using eye fixation to initiate the auditory cue (via fixating on the central point instead of clicking on it with a mouse) followed by button presses to respond (instead of moving a mouse cursor to an image and clicking on it) would minimize confounds in the eye tracking data. In the standard visual world paradigm, fixations early in the trial (e.g., before the response) are known to be strongly related to lexical activation. For example, in experiments where participants are asked to specifically not look at the image corresponding to the auditory word, participants still look at it a substantial amount of time, suggesting that an important amount of visual attention is coupled to lexical activation (Laurinavichyute et al., 2024). This coupling is very likely enhanced by the fact that participants must look at the screen to respond – participants must guide the mouse to the correct image in order to click on it. The fact that in my experiment, participants

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

were able to select the target image using a button press very likely suppressed and altered their looking patterns. In retrospect, it is apparent that in the extreme, it would be possible for participants to successfully carry out the VWP task without looking at the screen during the active part of the trial. For example, participants could have memorized the 4 images and their locations (e.g., in Figure 9, *wizard* – upper right) during the preview. Once the auditory stimulus played, all they would have to do is recall which quadrant held that picture and push the appropriate button – neither of which necessitating on-screen fixations. While it is clear that on average participants did not use this strategy, there is evidence that some may have and that the availability of this response strategy may have led to atypical fixation patterns. Several participants' fixation data (particularly those that did not converge to the model or had lower  $R^2$  values) barely diverged from baseline throughout the whole timecourse, confirming this theory that participants could memorize images and their locations before responding to abide by the instructions to respond as quickly as possible.<sup>15</sup>

This choice may have allowed participants greater flexibility in strategic resource allocation. Research comparing eye-tracking and mouse-tracking methodologies has shown that response methods requiring continuous visual-motor integration (e.g., mouse-clicking on images) enforce tighter coupling between perceptual processing and action execution, whereas spatially separated responses (e.g., button presses) allow for more independent control over the timing and allocation of attentional resources (Laurinavichyute et al., 2024). In my button-press paradigm, participants were not required to maintain continuous visual monitoring of potential targets to prepare motor responses, potentially enabling them to dynamically shift resources between visual monitoring and phonological maintenance based on momentary task demands. This flexible

---

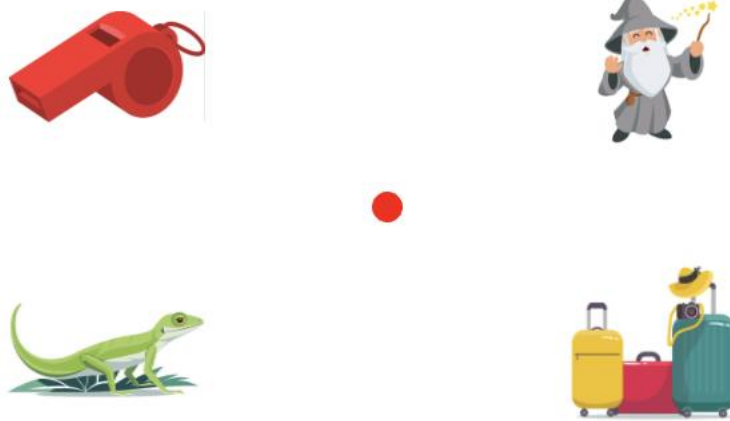
<sup>15</sup> Those participants with low fixation values had high accuracy on the VWP paradigm, ruling out the possibility that they were just not engaged in the experiment.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

allocation strategy could have allowed participants to preserve the core lexical competition processes which may occur relatively automatically once auditory input is processed while managing both visual signals and working memory. Switching to a regular mouse click should be explored in the future.

### Figure 9

*Pre-scanning screen in the Visual World Paradigm*



*Note: From right to left, top to bottom: whistle (Cohort), wizard (Target), lizard (Rhyme), and baggage (Unrelated).*

This may explain why overall fixation proportions to target items in Study 1 were different than those typically observed in visual world paradigm studies (e.g., McMurray et al., 2010, 2017). Target fixations were overall lower than in other studies, reaching approximately 40-60% at peak rather than 70-90% as commonly reported. Our fixations were also more transient – target fixations did not asymptote or plateau to the same degree as it does in previous studies.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

In conclusion, Study 1 demonstrated that phonological working memory load produces general cognitive slowing without disrupting the temporal dynamics of lexical access within the visual world paradigm. The substantial individual variability obtained in the experiment raises questions about whether these data, which made use of the button-press paradigm, optimally captured lexical competition dynamics. These methodological concerns combined with the theoretically unexpected null findings motivated Study 2. I turned to computational simulations to investigate our theoretical framework more directly.

**The proof is in the processing: Simulating load with jTRACE**

Study 2 aimed to use computational simulation to provide constraints on how phonological load could interfere with lexical access. Continuing the thrust of speech recognition research from Study 1, I introduced phonological load to lexical processing in the TRACE model (McClelland & Elman, 1986). Specifically, I examined how lexical processing occurred when load was assumed to act within the lexical system (similar to a competing word), and domain generally (by consuming resources that support lexical processing but not otherwise interacting with the lexical system). For the purposes of this work, I assumed that the null results of load in Study 1 are in fact a true reflection of how lexical processing occurred. I then examined whether domain-general or lexicon-internal implementations of load best align with these behavioral results.

TRACE is an interactive activation model of spoken word recognition originally developed by McClelland & Elman (1986). It simulates how listeners recognize spoken words and highlights the importance of interactive and dynamic processes in speech perception. Simulations with the TRACE model capture the major trends of fixations to the target and its competitor over time. It consists of three levels of processing units: the *feature* level, representing acoustic features of speech sounds that have been detected; the *phoneme* level, which simulates how phonemes are processed; and the *word* level, where lexical activation and selection is assumed to take place.

To simulate how an acoustic signal is processed, first the signal's corresponding features are activated at the feature level, which in turn activate phonemes at the phoneme level, which then activate words at the word level. In a nutshell, words that contain the activated phonemes

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

are themselves activated, creating a set of candidate words. Within each level, activated units compete with each other through lateral inhibition – units within the same level inhibit each other, which helps to select the most likely candidate for the speech signal. At any time during perception, the activation level of a lexical candidate is determined by the bottom-up activation received from the speech input and the lateral inhibition received from other activated candidates. Activation also flows top-down: activated lexical candidates send feedback activation to their constituent phonemes and potentially features, reinforcing phoneme representations consistent with active words. Under optimal circumstances, the candidate that is most similar to the speech signal usually has the strongest activation and sends out the strongest inhibition to other candidates and therefore will win the competition.

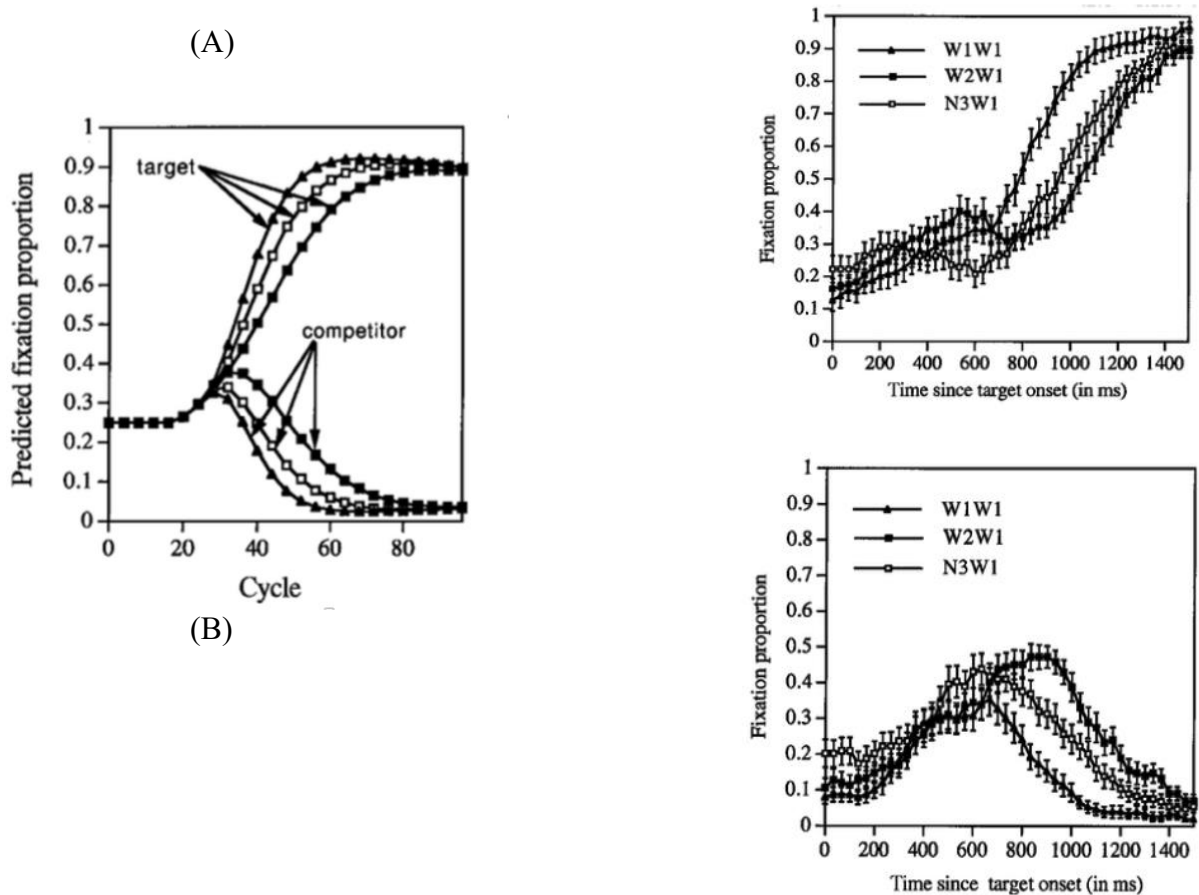
The TRACE model has been used over the years to simulate and test whether lexical access involves such features as lexical competition and word → phoneme feedback activation (e.g., Dahan, Magnuson, Tanenhaus, and Hogan, 2001). Dahan and colleagues (2001) used the TRACE model to simulate their eye-movement data taken from a VWP task. Using a publicly available TRACE implementation and the standard parameter set reported in McClelland and Elman (1986), they used the linking hypothesis that raw word activations in the model may be converted to fixation properties through Luce's choice rule (Luce, 1959). They found that their TRACE simulations mirrored important aspects of their eye-movement data, when Target words had Cohort competitors present in the display, both the model and human participants showed delayed and reduced target fixation during early processing, with fixations to Cohorts rising and then falling as the acoustic input disambiguated between the two candidates (see Figure 10). This shows that TRACE can specify the representations and processes involved in spoken word

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

recognition at a level of detail specific enough to make precise, quantitative temporal predictions about lexical access.

**Figure 10**

*Eye fixation proportions and TRACE predicted fixation proportions from Dahan et al. (2001)*



*Note.* (A) Proportion of fixations to the target picture (W1) (upper panel) and the competitor picture (W2) (lower panel) over time, for each experimental condition (W1W1, W2W1, N3W1). (B) TRACE predicted fixations for the target (W1) and the competitor (W2) over cycles, for each experimental condition (W1W1, W2W1, N3W1).

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

I used TRACE to explore different possibilities for how phonological load affects lexical access. I assumed that if the effects of phonological load on lexical access in speech perception are as indicated by the results of Study 1 (i.e., minimal to no effects of phonological load on lexical competition), then I should see manipulations of load in TRACE to lead to a similar pattern. I compared two implementations – one lexicon-internal, one domain-general – to determine 1) what these effects each produces and 2) which is more consistent with my Study 1 findings, thereby constraining the theoretical interpretation of how phonological load operates. I was able to explore what it means for load to interfere from within the lexical system itself (e.g., if the load functions as a competitor word during access), and what it means for it to interfere from outside of the system (e.g., via sapping of cognitive resources).

By examining the effects of both cognitive load on the activation and suppression of lexical candidates, this study contributes to a more comprehensive model of lexical processing that reflects real-world complexities.

### **Method**

#### ***Materials***

I used jTRACE, a Java-based reimplementation of TRACE (Strauss, Harris, & Magnuson, 2007), which uses the same underlying architecture as TRACE. For this and subsequent simulations, I created a lexicon consisting of four disyllabic monomorphemic words that reflect the types of words used in a TCRU Visual World Paradigm competitor set: "rocket" (/rak<sup>^</sup>t/; target), "rocker" (/rak<sup>^</sup>r/; cohort competitor), "pocket" (/pak<sup>^</sup>t/; rhyme competitor), and "bubble" (/b<sup>^</sup>b<sup>^</sup>l/; unrelated control). These items were selected to match the stimulus structure used in Study 1, allowing for direct comparisons between model predictions and the eye-tracking

data. Response probabilities are analogous to fixation proportions in visual world paradigms (VWPs).

### Procedures & Results

#### *Baseline*

I first established that jTRACE could capture the basic dynamics of lexical activation in our four-word competitor set. The baseline simulation (TRACE default parameters) produced the characteristic time course of spoken word recognition observed in visual world paradigm studies (Figure 11). At the onset of the simulation, all four lexical candidates – target (*rocket*), cohort competitor (*rocker*), rhyme competitor (*pocket*), and unrelated word (*bubble*) – began at equal predicted response proportions (0.25), reflecting the initial ambiguity before acoustic information becomes available.

As the presentation of the first few phonemes was simulated, the activation of the target and cohort competitor rose together, reflecting the fact that they both begin with these phonemes (*/rak/-*). This early co-activation is the hallmark of cohort competition in continuous speech processing: listeners can't distinguish between words that share onset phonemes yet. During this period, the rhyme and unrelated competitors were rapidly suppressed toward floor, as they received inhibition from the target and cohort competitor as they initially became active.

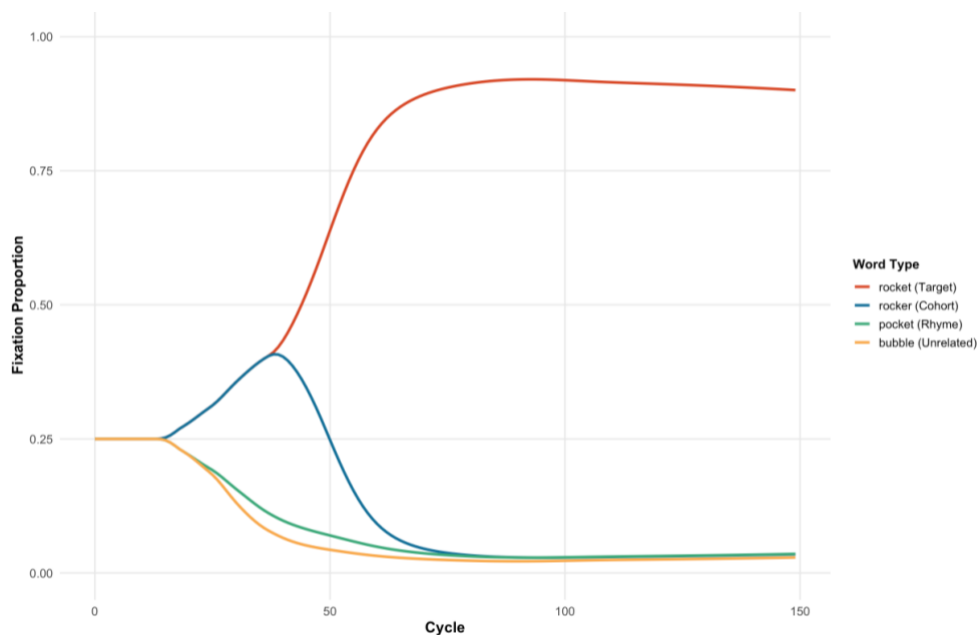
The cohort competitor reached its peak activation around cycle 40, at which point the acoustic input began to disambiguate between the two candidates. As the input phonemes began to diverge from those of the cohort competitor, activation diverged: the target continued to rise while the cohort competitor no longer received bottom-up activation and became suppressed through lateral inhibition. The target showed a clean sigmoidal rise to asymptote, reaching near-ceiling predicted response proportions (~1.0) and maintaining this level through the remainder of

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

the trial. This asymptotic pattern indicates complete lexical resolution – the model had accumulated sufficient evidence to unambiguously identify the target word. The default parameters also allowed the model to capture a secondary empirical finding, which is that rhyme competitors also become activated (diverging from unrelated competitors; Strauss et al., 2007).

**Figure 11**

*Baseline Target Predicted Response Probabilities*



The baseline simulation thus demonstrates that jTRACE, under default parameters, produces the fundamental dynamics of lexical competition and resolution: early ambiguity among candidates that overlap with the input, followed by later resolution in favor of the target. This establishes a foundation for evaluating how manipulations of cognitive load – whether through external resource depletion or lexicon-internal competitor pre-activation – alter these recognition dynamics.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Having established that the baseline TRACE model captures key aspects of lexical activation dynamics, I next investigated whether the model could account for the phonological load patterns observed in Study 1, assuming they were accurate. Specifically, I tested two competing hypotheses about the mechanism by which phonological load affects lexical access, via internal interference or through external interference.

### *Domain-general source of load*

Phonological load can affect lexical competition through reducing the amount of activation available to the lexical system for. Under this account, the resources used for lexical access are 1) finite and 2) shared with other cognitive processes (outside of the lexical system), including those responsible for working memory maintenance. The addition of phonological load would consume a portion of these shared resources, reducing the amount available for lexical processing. This resource depletion might not affect which specific lexical representations are activated (as in the lexicon-internal account) but rather would reduce the overall efficiency or strength of lexical activation processes globally.

In jTRACE, one way of implementing this domain-general reduction of resources is by reducing the value of the  $\beta$  parameter as given in the equation governing unit activation below. The beta parameter is a constant that controls the overall amount of activation that each unit has. In their simulations, Mirman et al. (2008) varied this parameter to simulate how the lexical system might function in cases of high and low attention.<sup>16</sup>

---

<sup>16</sup> Mirman and colleagues (2008) assumed that the normal state of affairs in the network is the state of high attention, and that low attention would represent relatively less attention. This is why they implemented  $\beta$  as a negative modification; units have more activation when  $\beta$  is small, and low activation when  $\beta$  is larger. I followed their work and implemented  $\beta$  in the same way.

**Equation 3**

$$net_{input} = \left( \left[ \sum a_p \times W_{p \rightarrow l} \right] + \left[ \sum a_l \times W_{l' \rightarrow l} \right] \right) - \beta$$

They used  $\beta$  manipulation to simulate attentional modulation of lexical influences on speech perception. Decreasing  $\beta$  reduces the amount of activation available to all of the word representations, simulating a reduction in cognitive resources (e.g., attention, working memory) allocated to lexical processing without changing the competitive structure among specific words.<sup>17</sup>

Critically, this domain-general manipulation affects all lexical units equally. This implementation of the external account thus predicts interference across the board for all stimuli in a VWP experiment. As I will detail shortly, this contrasts with the lexicon-internal account, which predicts differential interference based on the phonological similarity of the load items and target/competitor words.

**Parameter selection.** To observe the effects of implementing load using the beta parameter, I tested five conditions: baseline ( $\beta = 0.0$ , TRACE default), and three reduced-resource conditions ( $\beta = 0.05, 0.1, 0.15$ ). These values were selected to span the range of mild resource reduction to severe resource limitation that were used in Mirman et al. (2008).

**Qualitative interpretation of curves.** The results of these simulations are presented in Figure 12 (see Figure 11 as reference for the baseline condition of  $\beta = 0.0$ ). For  $\beta = 0.05$ , the general pattern and time course of target recognition remained qualitatively similar to the baseline: the Target word (rocket) showed a clean sigmoidal rise to asymptote, the cohort

---

<sup>17</sup> Unlike decay rate, which affects how quickly activation dissipates over time, the lexical bias manipulation affects the initial strength and sustained level of lexical activation.

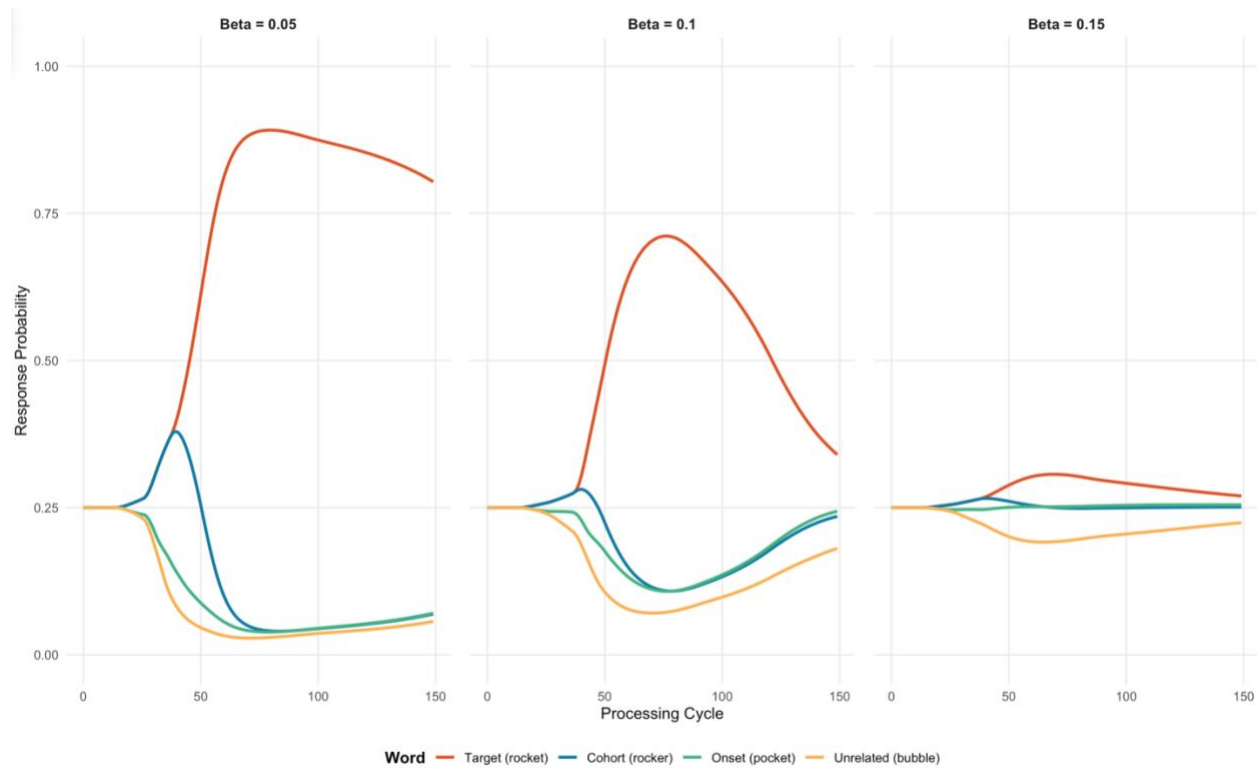
## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

competitor (rocker) exhibited a characteristic early peak followed by suppression, and rhyme and unrelated competitors remained at floor. The primary difference from baseline was a modest reduction in the rate and peak of target activation – precisely the pattern observed in our behavioral data comparing Phon to Sham conditions.

At  $\beta = 0.10$ , recognition dynamics began to show qualitative changes. Target and Cohort activation were reduced compared to  $\beta = 0.05$ , but the Rhyme and Unrelated items showed *higher* response probabilities (i.e., less suppression). This reflects how global decay disproportionately weakens the dominant competitors, reducing the lateral inhibition they send out to weaker items. For Rhyme and Unrelated competitors, this reduction in inhibition outweighs their own decay, producing a net increase in activation. The target still ultimately dominated, indicating that recognition could still be achieved.

At  $\beta = 0.15$ , recognition dynamics were fundamentally altered. The target word failed to achieve clear dominance over competitors, with predicted response proportions remaining near chance levels (approximately 0.25-0.35) throughout much of the time course. Rather than the characteristic sigmoidal rise to asymptote, activation remained flat and unresolved. This pattern suggests a qualitative breakdown in lexical access – the system could no longer accumulate sufficient evidence to resolve competition and identify the target word.

Based on this qualitative analysis I identified  $\beta = 0.10$  as the approximate boundary at which domain-general resource depletion began to produce qualitatively different recognition dynamics, with  $\beta = 0.15$  representing a clear breakdown of normal processing. The  $\beta = 0.05$  condition best approximated the subtle quantitative differences observed between our behavioral Phon and Sham conditions, suggesting that phonological load produces a modest reduction in lexical processing resources rather than a fundamental destruction of recognition mechanisms.

**Figure 12***TRACE Lexical Activation by Domain-General Load ( $\beta$ )*

*Note:* Higher beta levels simulate increased perceptual/input noise.

### ***Lexicon-internal source of load***

In the phonological load condition of the visual world paradigm (VWP), participants maintain a set of letters in working memory while performing the primary task. One mechanism through which load could affect lexical processing is if participants maintained these letters as phonological representations that directly compete with the lexical representations activated during spoken word recognition. Under this account, lexical units corresponding to letter names are activated in order to encode the letters given in the load task. These units would then compete with words activated by the auditory stimulus during the VWP task. For example, in a VWP trial where participants are given the letters the letters C, T, G, D in the load task, they would activate

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

lexical nodes in the language system corresponding to their phonological names (/si/, /ti/, /gi/, /di/). Then during the VWP, participants would hear /pik/ ("peek") as the stimulus while choosing between "peek," "peel," and "beak". The lexical representations of these words would become activated through the normal process of speech recognition. If we make the assumption that lexical nodes corresponding to letter names have the same lateral inhibitory connections as regular words, we would expect the activation of these additional lexical items (the letter names) would impact the activation - and consequently the time course of activation - of target and competitor words relative to the Sham condition.

In jTRACE, this lexicon-internal mechanism can be implemented by pre-activating letter name lexical units and then simulating the auditory presentation of the target word. Spoken word recognition would thus simultaneously be a function of the stimulus-driven processing and the preactivated load representations. The model's lateral inhibition mechanism would then allow these representations to compete, altering the time course and magnitude of Target and Competitor activation relative to baseline. This scenario represents a situation where the influence of phonological load on lexical activation occurs entirely through the intrinsic mechanisms of lexical processing – specifically, through competition among co-activated lexical representations.

For the lexicon-internal simulation I added letter names /si/, /ti/, /gi/, /di/ (corresponding to C, T, G, and D) as additional items in the simulated lexicon. I operationalized internal memory load by pre-activating these lexical units at the start of each simulation. TRACE implements this through the priming mechanism, which provides a one-time boost to the resting activation of specified lexical items before the simulation begins (Strauss, Harris, Magnuson, 2007).<sup>18</sup>

---

<sup>18</sup> This mechanism is functionally identical to the frequency-based resting level manipulation validated by Dahan, Magnuson, and Tanenhaus (2001), who use resting activation to model word frequency effects on lexical access. In

**Parameter selection.** The Priming (rest) parameter, which scales the values in the priming column of the lexicon, was systematically varied to determine values that produced effects comparable to those observed in the behavioral data. By elevating the resting activation of memory load items, these items begin the simulation in a partially activated state, enabling them to exert inhibitory influence on the target word even when they share no phonological overlap with the input. This effectively increases the "neighborhood density" or competition faced by the target, slowing recognition – a pattern consistent with the behavioral effects of working memory load on speech perception.

Based on the values documented in the jTRACE implementation (Strauss et al., 2007), I tested Priming (rest) values ranging from 0.03 to 0.9, with higher values representing greater pre-activation of load-relevant phonemes. The value of 0.06 served as a baseline reference point, as this was the scaling factor used by Dahan et al. (2001) to implement frequency-based resting activation effects. Other values were tested to identify parameters that best captured the magnitude of load patterns observed in the behavioral experiment.

**Qualitative interpretation of curves.** Visual inspection of the predicted response proportion curves revealed a striking pattern: recognition dynamics remained qualitatively unchanged across all priming levels (Figure 13). Regardless of the degree of phoneme pre-activation, the target word (rocket) showed the same characteristic sigmoidal rise to asymptote, reaching near-ceiling response proportions ( $\sim 0.90$ ) and maintaining this level through the end of the trial. The cohort competitor (rocker) exhibited its typical early peak followed by suppression, and the rhyme and unrelated competitors remained at floor throughout.

---

their implementation, log frequency values were scaled by a parameter (freq resting levels = 0.06) to produce resting activation boosts that successfully fit eye-tracking data on the time course of frequency effects during spoken word recognition.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

The most notable effect of increasing priming was a modest reduction in the cohort competitor's peak activation. At the lowest priming level (0.03), the cohort peaked at approximately 0.40; at the highest level (0.9), this peak was reduced to approximately 0.30. This pattern suggests that pre-activating phonemes may slightly dampen competitor activation, but critically, this effect did not propagate to target recognition dynamics in any meaningful way.

Even at extreme priming levels (0.9) – far exceeding any plausible level of phonological pre-activation from a four-item memory load – the fundamental architecture of lexical access remained intact. The target still dominated, competitors were still suppressed, and the overall time course was indistinguishable from baseline. This stands in stark contrast to the domain-general manipulation, where modest reductions in lexical bias ( $\beta = 0.10\text{--}0.15$ ) produced visible degradation of recognition dynamics.

The contrasting results of the domain-general and lexicon-internal simulations generate distinct predictions for behavioral data. Two key qualitative differences emerged. First, the locus of effects differed across simulation types: manipulating  $\beta$  (domain-general mechanism) produced changes across all simulated words, including both targets and competitors, whereas manipulating phoneme pre-activation (lexicon-internal mechanism) primarily affected competitors while leaving Target recognition largely intact. Second, the sensitivity to parameter manipulation differed drastically: recognition dynamics changed substantially over a very small range of  $\beta$  values, whereas they remained relatively invariant across large ranges of pre-activation parameter.

These patterns generate several testable predictions. Relative to a no-load control, if phonological load operates through an lexicon-internal mechanism, one should observe fewer fixations to Cohort competitors but not to Targets. If load operates through a domain-general

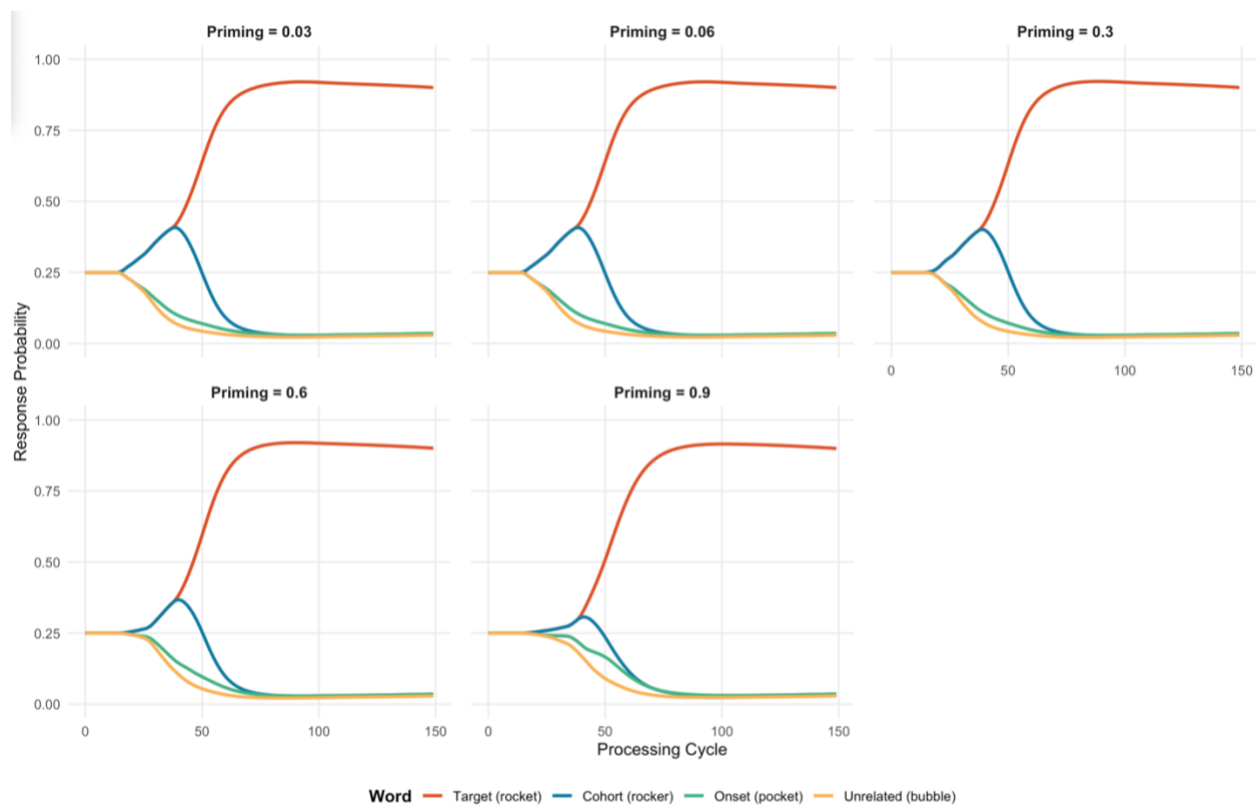
## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

mechanism, one should observe fewer fixations to both Targets and Cohort competitors.

Additionally, parametrically increasing load should produce a rapid decline in fixations if the mechanism is domain-general, but only a modest decline if lexicon-internal. Interestingly, both accounts converge on one prediction: fixations to rhyme competitors (and possibly unrelated competitors) should increase under load, as reduced activation of stronger competitors allows weaker alternatives to capture more looks.

**Figure 13**

*TRACE Lexical Activation by Resting Priming Level*



*Note:* High priming levels simulate increased working memory load.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

The behavioral results obtained in Study 1 revealed no significant difference in target recognition dynamics between Sham and Phon conditions. This pattern is consistent with the lexicon-internal simulations, in which phoneme pre-activation failed to disrupt lexical access regardless of the level of pre-activation. In contrast, the domain-general simulations predicted attenuation of recognition under load – a pattern not observed in the behavioral data. Although there are reasons to suspect that the Study 1’s null result may not be reliable (given the low Bayes’ Factors) it is reasonable to say that it is more compatible with a lexicon-internal account in which phonological load engages internal pre-activation mechanisms that do not meaningfully interfere with ongoing lexical processing. This interpretation also aligns with Zhang and Samuel’s (2018) findings. They observed that phonological load reduced priming for *ham* when hearing *hamster* but had essentially no effect on Target word recognition itself. If phonological load operates through internal pre-activation of specific phonemes, it would selectively disadvantage competitors sharing those phonemes while leaving Target processing intact – which is precisely the pattern that Zhang and Samuel observed.

### ***Exploratory analyses: Noise implementation***

In addition to primary load manipulations, I conducted exploratory simulations to examine how auditory noise affects lexical access. Zhang & Samuel (2015) observed that segment loss reduces phonological priming, raising the question of whether noise-induced reductions in priming operate through the same mechanism as load-induced effects. To address this, I manipulated the input signal in jTRACE by removing or replacing specific phonemes, simulating the degraded acoustic input that listeners experience under noisy conditions.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Under Zhang and Samuel's (2015) account, adding noise to the system reduces the bottom-up activation that both target and competitor words receive. In TRACE, this would manifest as a weakened priming effect: all lexical candidates would be less activated overall due to impoverished input, rather than any specific change in competition dynamics. This contrasts with the lexicon-internal load manipulation, which selectively pre-activates certain representations, and the domain-general manipulation, which globally reduces processing resources.

This exploratory analysis thus asks: Does noise impact lexical access through the same mechanism as phonological load? If noise and load produce similar patterns of recognition dynamics, this would suggest a shared mechanism – likely degraded bottom-up input. If they produce dissociable patterns, this would indicate that noise and load affect different components of lexical processing.

Related work by Mattys, Barden, and Samuel (2014) suggests that cognitive load primarily affects perceptual encoding rather than lexical processing itself. Using the phoneme restoration paradigm, they found that load increased reliance on lexical information, consistent with imprecise phonetic encoding under load rather than direct interference with lexical access. These findings motivate our exploration of whether noise – which directly degrades the input signal – produces patterns distinct from cognitive load manipulations.

**Methods.** To examine how degraded input affects lexical access, I implemented two noise manipulations in jTRACE following approaches used in auditory priming research. Both manipulations targeted the embedded word within a carrier word (e.g., "pie" within "pirate," "com" within "compact"), simulating conditions where portions of the speech signal are compromised.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

In the degraded condition, the target segment was replaced with acoustic noise, simulating environments where speech is masked but some signal remains (e.g., background noise, low-pass filtering). This manipulation reduces the clarity of phonetic information without eliminating it entirely, allowing partial bottom-up activation of the affected phonemes. In the deleted condition, the target segment was entirely removed from the input, simulating conditions of complete signal loss. This manipulation eliminates bottom-up support for the affected phonemes, forcing the system to rely on surrounding context and top-down lexical information.

I tested these manipulations on two-word types: carrier words (e.g., "compact," "pirate") and their embedded words (e.g., "pact," "pie"), allowing us to examine how noise affects both full lexical items and partial activations.

**Results.** Visual inspection of activation time courses revealed distinct patterns across noise conditions, with interesting differences between word-initial and word-medial embedded words (Figure 14).

For "compact," where the embedded word "pact" occurs at word offset, noise manipulations produced clear effects on carrier activation but minimal effects on embedded activation. All three conditions showed similar rise dynamics, with activation beginning around cycle 25 and reaching peak around cycle 75. However, the degraded condition showed slightly reduced peak activation compared to baseline, and both noise conditions showed continued decline after the peak rather than maintaining asymptote. This late decline suggests that degraded or missing input creates lingering uncertainty that affects sustained lexical commitment.

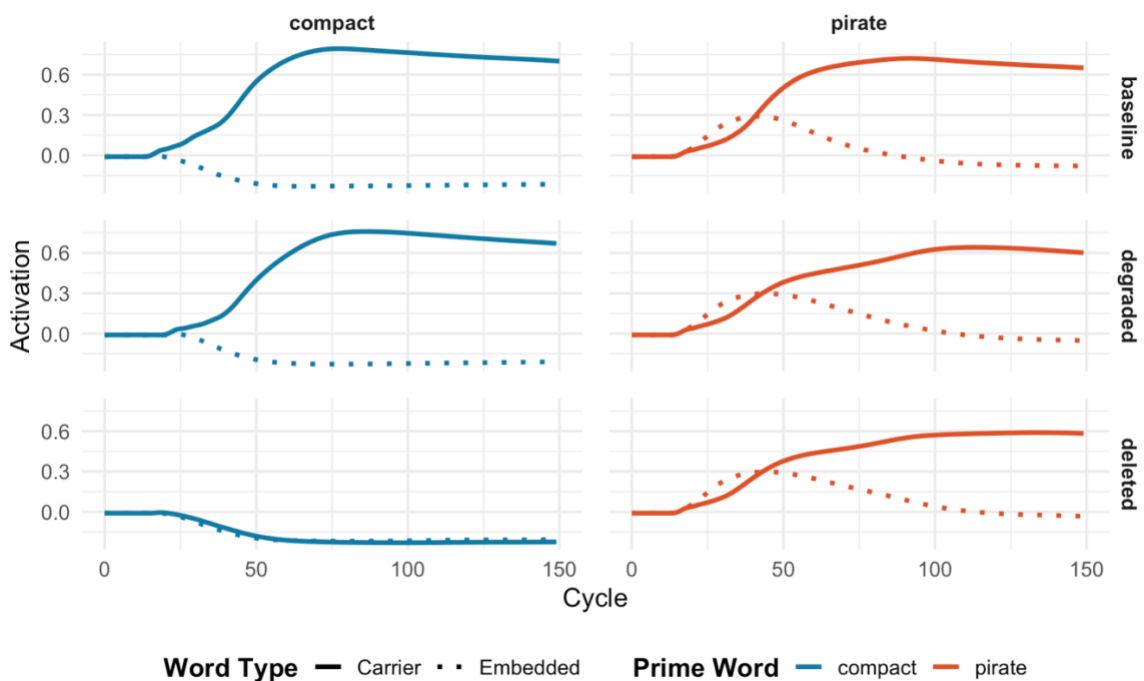
The embedded word "pact" showed strong suppression across all conditions, with activation dropping into negative values (indicating active inhibition). Notably, all three conditions produced nearly identical embedded suppression patterns; baseline, degraded, and

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

deleted curves overlapped almost completely. This suggests that for word-final embedded segments, the carrier word's lateral inhibition dominates regardless of input quality; by the time "pact" becomes relevant in the input, "compact" has already accumulated substantial activation and suppresses the embedded competitor.

**Figure 14**

*Activation Time Course Across Noise Conditions*



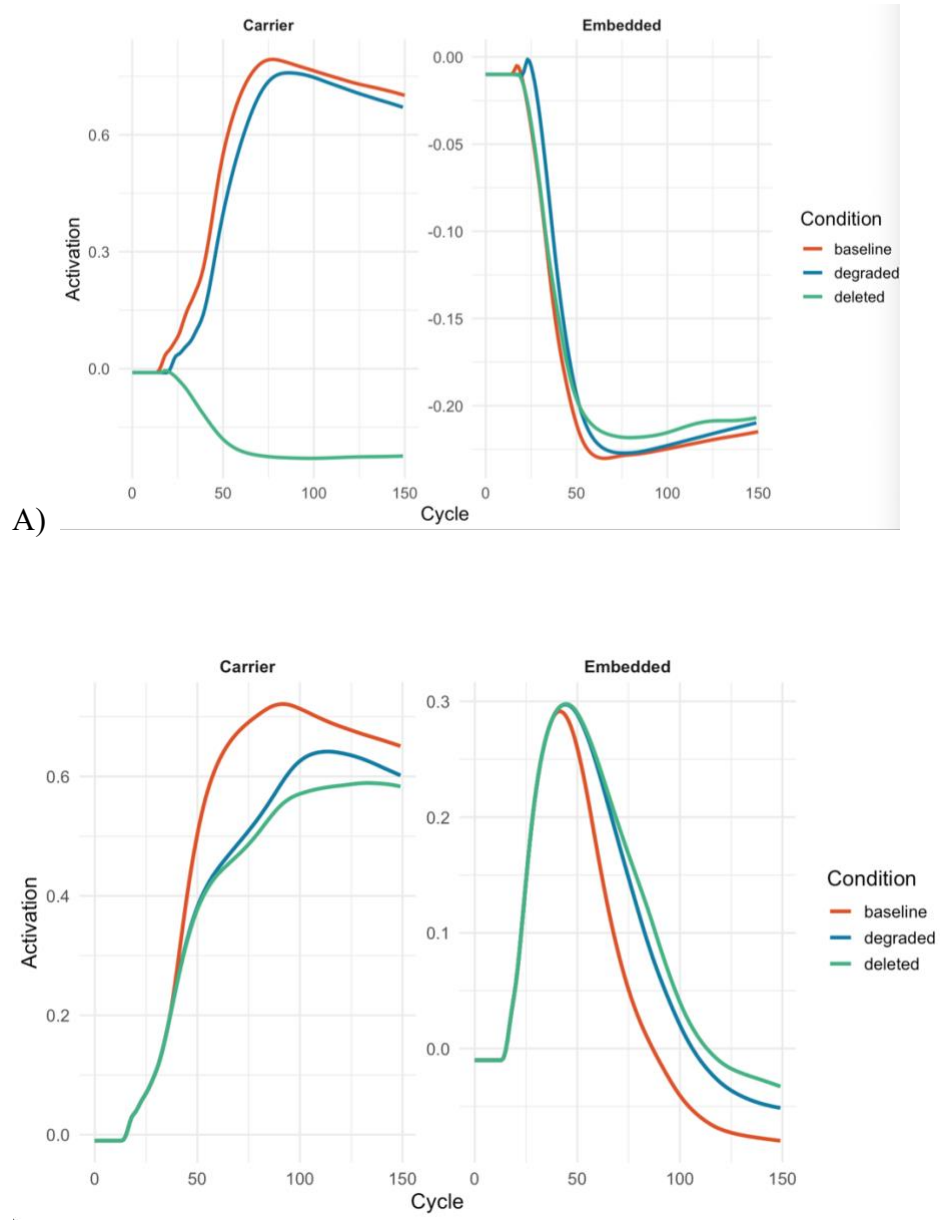
For "pirate," where the embedded word "pie" occurs at word onset, noise produced markedly different effects. Noise manipulations produced graded attenuation of carrier activation. Baseline reached peak activation around 0.70; degraded peaked lower at approximately 0.65; deleted peaked lowest at approximately 0.60. Additionally, the deleted condition showed a delayed rise, taking longer to reach peak activation. Unlike "compact," all

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

conditions maintained their activation levels after peaking rather than declining, suggesting that once the full word is heard, lexical commitment is sustained.

**Figure 15**

*Individual item curves under noise*



*Note: A) Compact and B) Pirate*

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Noise produced visible effects on the embedded word "pie." All conditions showed the characteristic rise-then-fall pattern of a temporarily activated competitor, peaking around cycle 50 at approximately 0.30 activation. However, the degraded and deleted conditions showed prolonged activation compared to baseline; the embedded word took longer to be fully suppressed under noise. This suggests that degraded input slows the resolution of competition between embedded and carrier words when the embedded word occurs at onset and receives early bottom-up support.

In summary: For "compact" (offset-embedded "pact") the carrier shows late instability under noise, but embedded suppression is unaffected. The carrier accumulates activation before "pact" becomes relevant, so lateral inhibition dominates regardless of input quality. For "pirate" (onset-embedded "pie"), the carrier shows graded attenuation under noise, and embedded word shows prolonged activation - because "pie" receives early bottom-up support at word onset, noise slows the resolution of carrier-embedded competition.

**Comparison to load localization.** Domain-general load ( $\beta$  reduction) produced global attenuation of all lexical activation - both targets and competitors were reduced proportionally, and at extreme levels, recognition dynamics broke down entirely. In contrast, noise produced selective effects: carrier words were relatively preserved while embedded words were disproportionately affected. This selectivity suggests that noise degrades specific bottom-up input rather than reducing global processing resources. The degraded condition, in particular, showed intact carrier recognition with reduced embedded activation—a pattern that cannot be explained by uniform resource depletion.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Lexicon-internal load (phoneme pre-activation) produced no meaningful change in recognition dynamics at any level tested. Noise, by contrast, produced clear and graded effects: degradation attenuated embedded word activation, and deletion nearly eliminated it. This dissociation indicates that noise and lexicon-internal load affect different aspects of processing. Lexicon-internal load increases resting activation of specific representations but does not alter the incoming signal; noise directly impoverishes the signal itself, reducing the bottom-up evidence available for lexical access.

If phonological load in our behavioral paradigm operates through a lexicon-internal mechanism (as suggested by the null behavioral effect), then load and noise should produce dissociable effects on lexical access. This is consistent with Mattys et al.'s (2014) proposal that cognitive load affects perceptual encoding rather than lexical processing: load may reduce encoding precision without eliminating bottom-up input, whereas noise directly removes or corrupts that input. The noise simulations thus provide a useful contrast case, demonstrating that jTRACE can capture input degradation effects while confirming that phonological load likely operates through a different pathway.

### **Discussion**

The primary aim of Study 2 was to develop proposals at the mechanistic level for how phonological load interferes with lexical access. Using the TRACE model (McClelland & Elman, 1986) to simulate lexical access advances the theory of spoken word recognition in natural communication settings and enhances our understanding of how environmental factors, such as noise, interact with cognitive load to influence lexical competition.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

I tested two mechanisms of localizing load by implementing both in jTRACE simulations and examining which best produces a pattern consistent with the empirical findings from Study 1. If the lexicon-internal mechanism is correct, we should observe that activating letter-name representations alongside stimulus-driven words produces changes in competition time course – specifically, altered competitor activation patterns. If the domain-general mechanism is correct, globally reducing activation resources should produce overall slowing that could manifest as delayed response times without necessarily altering the relative time course of target versus competitor activation.

Domain-general load produced graded attenuation of recognition, from subtle effects at modest reductions to complete disruption at severe reductions. Lexicon-internal load produced no change in recognition dynamics, even at extreme pre-activation levels—lateral inhibition resolved additional competition without cost to target recognition. Behavioral results showed no significant difference between Sham and Phon conditions, a pattern more consistent with lexicon-internal simulations than domain-general predictions. While this does not confirm the null, it suggests that phonological load engages pre-activation mechanisms that do not meaningfully interfere with lexical access - consistent with proposals that load affects perceptual encoding rather than lexical processing (Mattys et al., 2014).

Exploratory noise simulations produced a third, distinct pattern: selective impairment of representations dependent on degraded input, with effects varying by embedded word position. This three-way dissociation—global attenuation (domain-general), no change (lexicon-internal), selective impairment (noise)—suggests these mechanisms affect different aspects of lexical processing.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

The findings indicate that the lexical system is robust to increased competition from pre-activated representations. Qualitative disruption of recognition emerges only under severe resource constraints, suggesting that typical phonological load operates within a range that the system can accommodate without fundamental disruption to lexical access.

**Eating your words: Load effects on speech production**

In Study 3, I investigated whether phonological load interferes with language production, extending the central question from Study 1 into the domain of production. Study 1 used eye tracking in the visual world paradigm to ask whether phonological load disrupts lexical competition during spoken word recognition – specifically whether maintaining phonological information would reduce activation of cohort and rhyme competitors. Study 3 asks whether similar effects emerge when speakers are *producing* language: does phonological load interfere with the phonetic implementation of speech?

Theories of spoken production propose that phonological neighbors become active during processing (through feedback connections from phonemes to lexical units), resulting in lexical competition that is not unlike that of speech comprehension (Dell, 1986). Participants were instructed to produce spoken responses to visual stimuli while simultaneously maintaining either a Sham load (control condition with minimal cognitive demand) or a phonological load (rehearsing a four-letter phonological string) on each trial. After viewing each set of words, participants selected the item indicated by a visual cue and produced it aloud as quickly and clearly as possible. An example of the trial sequence is provided in Figure 16.

This study aimed to investigate how cognitive load interacts with lexicon-internal factors such as lexical competition (onset-differing neighbors, VC), and domain-general factors, such as contextual cues, in shaping voice onset time (VOT) during speech production. Based on prior research (Baese-Berk & Goldrick, 2009; Fricke et al., 2016), I created the following predictions:

**Hypothesis 1:** In the Sham load condition, words presented in the Context condition will elicit longer VOTs than words in a No Context and No Competitor condition ( $VOT_{\text{Context}} >$

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

$VOT_{NoContext} > VOT_{NoCompetitor}$ ), replicating prior findings that external context lengthens production timing.

**Hypothesis 2:** Under phonological load,  $VOT_{Context}$  and  $VOT_{NoContext}$  will be reduced, while  $VOT_{NoCompetitor}$  will remain relatively stable. Phonological load should reduce the activation of lexical items. The VOT effect observed by Baese-Berk and Goldrick (XXXX) was thought to arise from the co-activation of a competitor; therefore, when load reduces competitor activation, this should reduce overall VOT magnitude.

**Hypothesis 3:** Phonological load should also reduce the context effect. Cognitive load will interfere with the use of domain-general cues, leading to smaller differences in VOT between Context and No Context/No Competitor conditions. This would echo the reduced activation of non-target words observed in Zhang & Samuel's (2015, 2018) comprehension studies. Thus, we may see a similar pattern in production, where phonological load impairs the ability to produce the target word in a way that reflects competition with its phonological neighbors.

**Hypothesis 4:** Lexical competition measured through onset-differing neighbors (VC) will affect VOT in both Sham and phonological load conditions. This effect is expected to reflect internal, lexicon-internal mechanisms of production variation, which should remain stable even when cognitive resources are taxed by phonological load.

By directly comparing performance under Sham and phonological load conditions, Study 3 examined how cognitive resources shape the dynamics of speech production. These hypotheses also allowed us to test whether phonological load disrupts primarily external, context-driven mechanisms of speech planning or internal, lexically driven mechanisms of competition. This

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

study extended the framework established in Studies 1 and 2, where load effects on lexical activation were observed and modeled in comprehension.

### **Method**

#### ***Participants***

Thirty-six participants were recruited for this study from the Tufts University SONA Pool – 17 in the Sham load condition and 19 in the phonological load condition. (In comparison, the Baese-Berk (2009) study recruited 12 participants.) All were native English speakers with no (self-reported) history of learning or language disability, and no auditory impairments. They all also provided informed consent following the guidelines of the SBER Institutional Review Board (IRB) and were compensated for their time with research credit. They all also had to consent to having their spoken responses be audio recorded. A participant could not participate in this study if they took part in Study 1. Lastly, each participant only sat in one study session which took approximately 90 minutes.

#### ***Materials***

The stimuli used in this experiment were from the original study and were 36 pairs of monosyllabic words beginning with a voiceless /p/, /t/, or /k/ stop. These pairs were matched for initial consonant and vowel (e.g., *peek-peel*, *tab-tat*, *cob-cog*), sum segmental probability, sum biphone probability, and phoneme length.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Within each pair, one word had a minimal pair neighbor when the initial consonant was voiced (e.g., *peek* with neighbor *beak*),<sup>19</sup> while the other pair had no similar neighbor (*peel* and *\*beel*). The pairs were embedded in a list of monosyllabic fillers.

Trials were presented in three conditions, as in Baese-Berk and Goldrick (2009): a Context, No Context, and No Competitor condition. In the Context condition, a target, its minimal pair neighbor, and a filler word were presented on the computer screen; in the No Context condition, the target from the Context condition was presented without its minimal pair, and instead with two Unrelated fillers; lastly, in the No Competitor condition, words with no competitors were presented with two Unrelated fillers. Target trials were mixed with filler trials. The list of words was presented three times in a different random order to increase the number of tokens and reduce the influence of random variance within participants. In total, participants underwent 432 trials ([18 Context + 18 No Context + 36 No Competitor + 72 filler trials] x 3 = 432).

**Table 5**

*Trial conditions for Study 3*

<b>Trial Condition</b>			
<b>Context</b>	cod	god	yell
<b>No Context</b>	cod	lamp	yell
<b>No Competitor</b>	cop	lamp	yell

---

<sup>19</sup> While the stimuli for this experiment were constructed to include pairs of monosyllabic words with either a minimal pair neighbor or no similar neighbor, it is important to note that the analyses demonstrating the effects were primarily driven by the presence of onset-differing neighbors. Specifically, the non-minimal pair group contains fewer onset-differing neighbors than the minimal pair group, which supports the validity of the observed effects across these conditions.

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

The load tasks were the same as in Experiment 1, with the only exception being that the phonological load stimuli did not share any letters with the word pair to avoid any cross-priming effects.

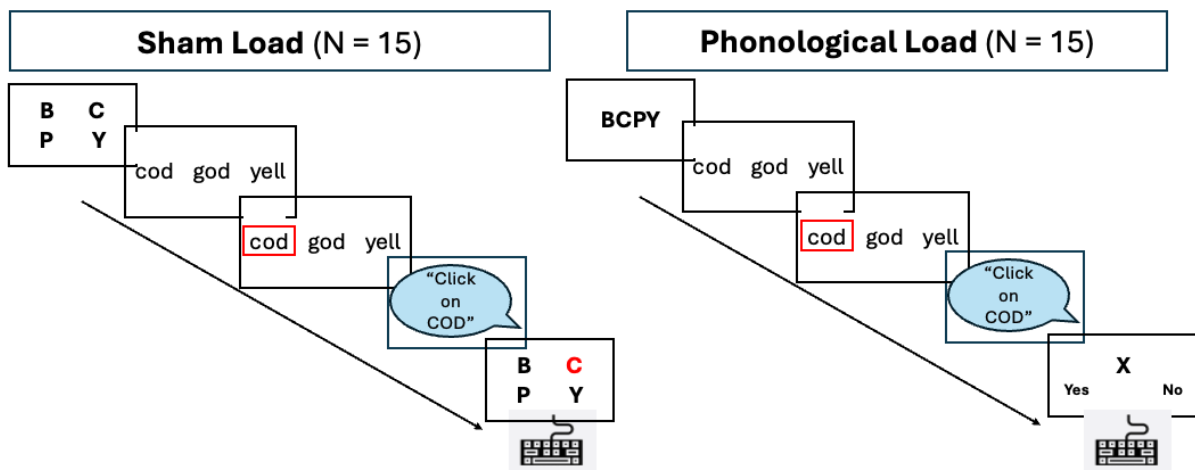
Participants sat in a soundproof audio booth. The experimental program was written in PsychoPy.

## Procedure

The procedure followed the same load task paradigm as Study 1, except with the Baese-Berk and Goldrick (2009) spoken production task as the primary task. Participants were randomly assigned to the Sham load or the phonological load group.

**Figure 16**

*A Single Trial of Study 3*



*Note:* The load task was *either* phonological *or* non-phonological load for a single participant.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

The experiment proceeded as follows: Participants first saw the load task stimulus on the monitor. Once that passed, the primary task began. Three words appeared simultaneously on both computer screens. After 1000 ms, a box appeared around one of the words on the participant's screen. Participants instructed a listener to 'Click on the *target word*' (where target word stood for the word in the box) out loud into the microphone. Prior to the experiment beginning, participants were instructed to speak clearly into the microphone, so that someone who listens to the audio at a later time would be able to follow their instructions. This was to mimic the participant-listener aspect of the original study, and to also ensure the participant was speaking naturally and clearly. The participant was recorded during the entire experiment. After 2000 ms, the screen automatically timed out, the primary task trial ended, and the secondary load task response screen showed depending on the participant condition. If participants were in the Sham load condition, they saw the same four letters they saw before that primary task trial began and were asked to choose the side (right or left) with the red letter on it ("Z" for left; "/" for right). Alternatively, if participants were in the phonological load condition, participants saw a letter and were asked to select Yes or No with the keyboard ("/" being yes; "Z" being no) to whether they had seen this item on screen before the primary trial began. Either way, the next trial began 1000 ms after the primary or secondary response. If the participant failed to respond within 3000 ms, the next trial began.

## Results

### *Data Preprocessing*

There were 7009 total data points (after removing filler trial data). Voice onset times were calculated using Dr. VOT (Shrem, Goldrick, Keshet, 2019). Dr. VOT is an automated script that

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

measures voice onset time (VOT) from recorded speech samples. It identifies the release burst and the onset of voicing using acoustic cues (e.g., amplitude and zero-crossing rate changes), allowing for consistent, high-precision measurements across tokens and speakers.

It was discovered that on typical trials, Dr. VOT worked well at isolating and calculating VOT measurements – however, after further investigation, it yielded a 5 ms as a default VOT assignment for trials in which:

- 1) Participants did not say the target word (timed out, or no response at all).
- 2) Participants did not say the whole phrase (“Click on the *target word*”). A majority of the anomalous VOT measurements came from this – in which the “the” was omitted or barely distinguishable.
- 3) Participants stuttered, repeated the phrase again, or introduced new words that Dr. VOT was not trained for (e.g., “Click on – click on the *target word*”).

This caused Dr. VOT to isolate the wrong parts of the audio file. Because of this, half of the data points (approximately 3000) were manually reviewed for quality control. These manual measurements were added to the data set. Trials in which VOT was 0 (no response or time out) were removed (N=7005). I excluded all VOTs 2.5 standard deviations away from an individual’s average (N= 6909; 1.4% of data removed). Tokens were averaged across the 3 repetitions.

**Table 6**

*Average VOT and Accuracy in Primary and Secondary Tasks*

Condition	Secondary (Load) Task	Primary Task
	Mean Accuracy $\pm$ SD	Avg VOT (ms) $\pm$ SD <sup>20</sup>

<sup>20</sup> VOTs were calculated across trials with accurate load task responses only.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Sham Load	1.0	80.8 ± 0.24
Phonological Load	0.79 ± .17	78.8 ± 0.21

### *Analyses*

Descriptive statistics were calculated for both the average VOT (across tokens) and load task accuracy (see Table 7). Because these two metrics were not correlated, I chose to keep data from participants who did not do as well on the load task (<60% accurate) for the overall analyses.

Instead of using a model to analyze the differences between minimal pairs and non-minimal pair trials as the original Baese-Berk and Goldrick study (2009) did, I modeled my analyses off of the Fricke, Baese-Berk, and Goldrick (2016) follow-up study, that found that using onset-differing neighborhood density (ND) better explained the results. This parameter was used to 1) verify that there was a main effect that the original study found and 2) to see if there was an interaction with phonological load.

All analyses were conducted using linear mixed effects models with the response variable being log-transformed VOT (in milliseconds). The model included fixed effects of load condition (two levels: Sham load, phonological load), context (three levels: Context, No Context, No Competitor), consonant (onset of target word, either /p/, /t/, or /k/), vowel height (high versus low vowels in the target word), and onset-differing ND (continuous predictor, the number of onset-differing neighbors for each target). Random intercepts were included for participant and target word.

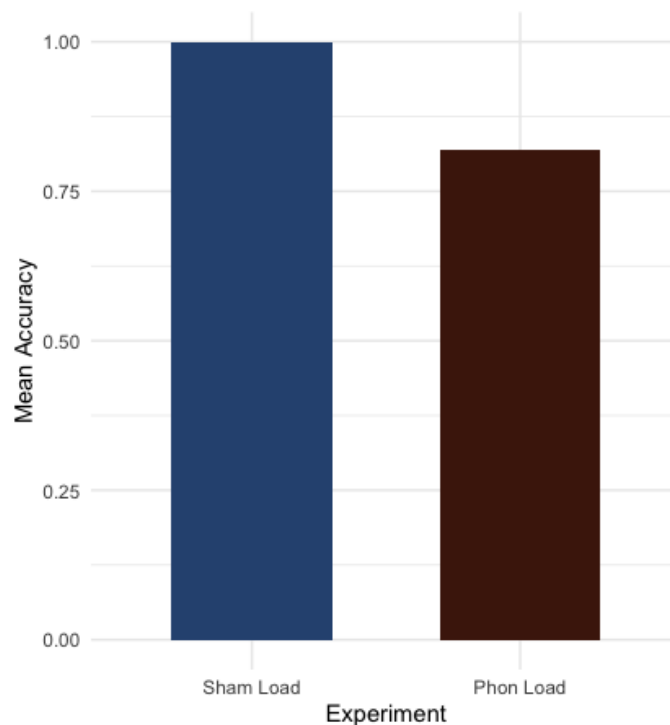
**Accuracy.** Accuracy in the production task was determined using VOT measurements – if VOTs were 0, this meant that there was no response or time out.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Participants were highly accurate in the load task across load conditions (see Figure 17). Accuracy in the Sham load task was at ceiling (100%), while accuracy in the phonological load task was slightly lower but remained high ( $M = .79$ ;  $SD = .17$ ). Because the Sham load task showed no variability, trial-level inferential comparisons could not be estimated.<sup>21</sup> Regardless, at face value this shows that the load manipulation worked – the phonological load task was more difficult to complete accurately than the Sham load task.

**Figure 17**

*Mean Accuracy across Load Conditions*



*Note:* Participants were less accurate in the phonological load task versus the Sham load task.

This shows that the phonological load task manipulation was successful.

---

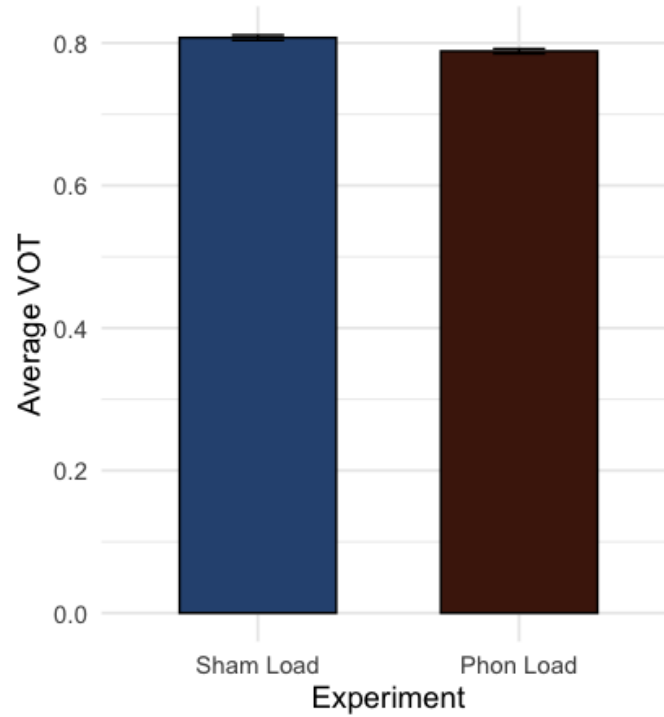
<sup>21</sup> Statistical models estimate differences in variability and covariance. If one group has no variance, the model can't estimate an effect because there's nothing to compare against.

**VOT Analysis 1: Effect of Context in Sham Load.** I predicted that in the Sham load condition, Context VOTs should be longer than No Context VOTs, which should also be longer than No Competitor VOTs ( $VOT_{Context} > VOT_{NoContext} > VOT_{NoCompetitor}$ ), consistent with Baese-Berk and Goldrick (2009). Linear mixed effects models indicated that  $VOT_{NoContext}$  did not differ significantly from  $VOT_{NoCompetitor}$  in the Sham load condition. Because the  $VOT_{NoContext}$  also trended opposite to predictions – showing slightly shorter VOTs at baseline than  $VOT_{NoCompetitor}$  – it was removed from further analyses. Words in the Context condition, however, had significantly longer VOTs than words in the No Competitor condition, suggesting that contextual information at baseline strengthens VOT, affirming part of Baese-Berk and Goldrick (2009)'s results and Hypothesis 1.

### **Figure 18**

*Average VOT across Load Conditions*

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION



*Note:* Average VOTs were slightly higher in the Sham load condition ( $80.8 \pm 0.24$  ms) than in the phonological load condition ( $78.8 \pm 0.21$  ms). This difference was not statistically significant.

**VOT Analysis 2: Effect of Phonological Load on Contextual Effects.** The average voice onset times (VOTs) across the two experimental conditions were slightly higher in the Sham load condition ( $80.8 \pm 0.24$  ms) compared to the phonological load condition ( $78.8 \pm 0.21$  ms). This difference, however, was not statistically significant – on average, VOTs in the phonological load were not significantly different from those in Sham load condition.

**Table 7**

*Average VOTs across Contextual Cue Type*

	Trial Context	Average VOT (ms) ± SE
Sham Load	No Competitor	$80.35 \pm 0.59$

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

	Context	82.08 ± 0.86
	No Context	80.3 ± 0.85
Phon load	No Competitor	78.64 ± 0.56
	Context	78.44 ± 0.82
	No Context	79.47 ± 0.84

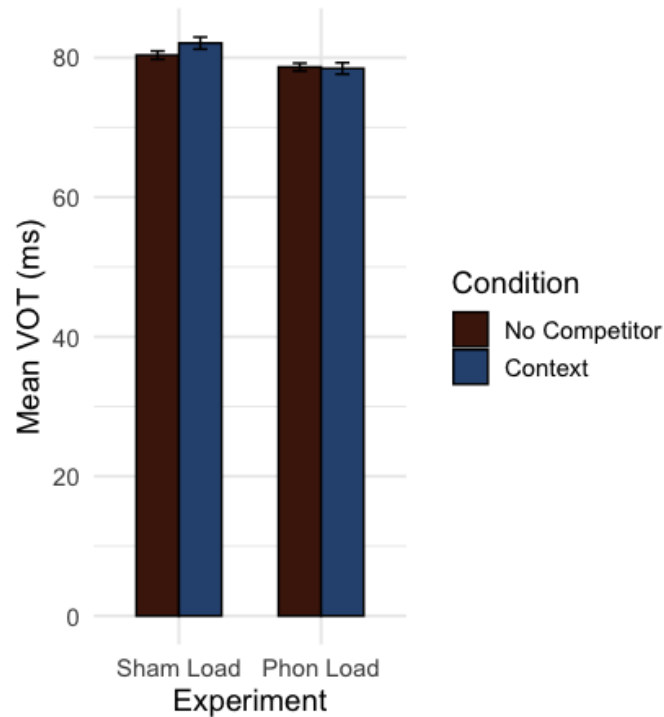
Across all contexts (Context, No Context, and No Competitor), VOTs were consistently shorter under phonological load compared to Sham load (see Table 8). These results align with Hypothesis 2 and demonstrate a clear effect of phonological load in reducing VOTs across all contextual conditions.

The interaction between phonological load and Context was significant. Under phonological load, the context effect ( $VOT_{Context} > VOT_{NoCompetitor}$ ) was reduced and actually reversed, indicating that cognitive load significantly moderates the effect of context on VOT. This provides evidence in support of Hypothesis 3.

### Figure 19

*Mean VOT: Context vs No Competitor Across Load Experiments*

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION



**VOT Analysis 3: Effect of Lexical Competition (VC).** To further investigate the contributions of different factors to VOT, I conducted model comparisons to assess the significance of VC (onset-differing neighbors) and context (Context, No Context, and No Competitor). Adding VC to a model that already contained context significantly improved model fit, and conversely, adding context to a model that already included VC was also significant. This indicates that both factors independently contribute to VOT variation.

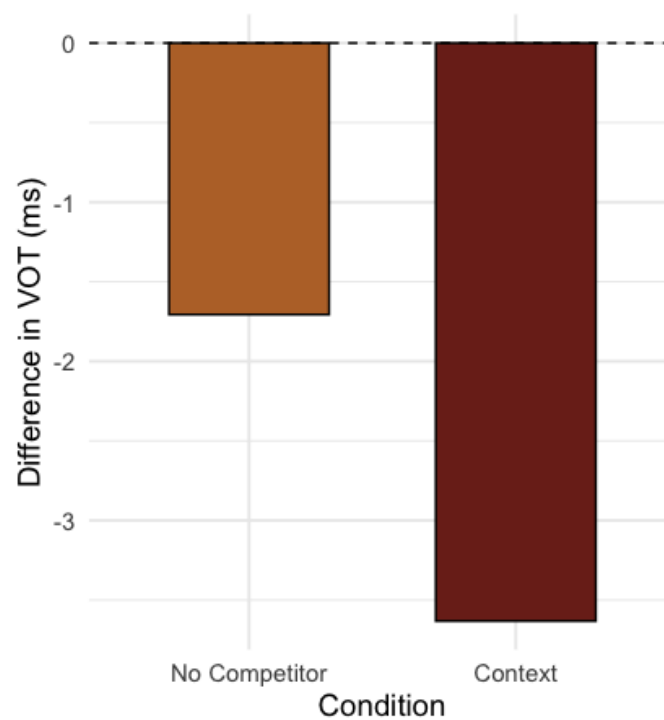
I then tested the interaction between VC and Load Condition (phonological versus Sham load). This interaction could help pinpoint the source of the VOT lengthening – whether it’s exogenous (context-dependent) or lexicon-internal (VC). The results showed that VC affected VOT similarly across both tasks: VOTs increased or decreased with VC in the same way, regardless of whether participants were under phonological load, providing evidence for Hypothesis 4. This is in contrast to the interaction effect discussed earlier between phonological

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

load and Context - Context interacted with Load Condition to affect VOTs, suggesting that the effects of Context differ under phonological load. Both interaction analyses together suggest two distinct sources of VOT lengthening – VC and context. The presence of these sources and their implications will be discussed in detail in the next section.

**Figure 20**

*Difference in Mean VOT (ms): Phon Load – No Load*



*Note:* The interaction between phonological load and Context was significant. Under phonological load, the context effect ( $VOT_{Context} > VOT_{NoCompetitor}$ ) was reduced and reversed, indicating that cognitive load moderates the influence of context on VOT.

**Other Fixed Effects.** Other fixed effects were also significant. Targets beginning with consonants /p/ and /k/ were significantly shorter than /t/, although the effect for /k/ was minimal

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

(estimate = -0.03). VOTs were shorter when the following vowel was low compared to high, consistent in direction and magnitude with Fricke et al. (2016). Finally, each onset-differing neighbor (VC) slightly but significantly lengthened VOT (estimate = 0.008), indicating that more onset-differing neighbors were associated with longer VOTs, again aligning with Fricke's findings.

### **Discussion**

The purpose of Study 3 was to investigate how phonological load affects voice onset time (VOT) in spoken word production. By comparing Sham load and phonological load conditions, I aimed to determine whether cognitive load disrupts internal lexical mechanisms, external contextual effects, or both.

#### ***Summary of Main Findings***

Across all analyses, I found that phonological load consistently reduced VOTs compared to Sham load. Contextual effects on VOT were evident at baseline: Words in the Context condition had longer VOT than words in the No Competitor condition, while the No Context condition did not differ significantly from the No Competitor and trended opposite to predictions. Importantly the interaction between context and load revealed that phonological load reduced and even reversed the context effect, suggesting that cognitive load selectively impacts the use of external cues. In contrast, VC affected VOT similarly across both load conditions, indicating that internal lexical competition mechanisms remain robust under load. Other fixed effects – consonant identity and vowel height – were consistent with prior findings in the literature (e.g., Fricke et al., 2016), further validating the model.

### *Contextual Effects and External Cues*

At baseline, Context produced significantly longer VOTs than the No Competitor condition, replicating the pattern reported in Baese-Berk and Goldrick (2009). This suggests that speakers utilize external contextual information to modify speech production. Under phonological load, the context effect was reduced, indicating that cognitive load interferes with the processing of external cues. This aligns with attentional accounts suggesting that load reduces available resources for attending to contextual information. These results indicate that domain-general, context-driven influences on speech production are vulnerable to cognitive constraints.

### *Lexicon-Internal Competition Mechanisms*

Onset-differing neighbors (VC) significantly contributed to VOT modulation independently of context, and the effect of VC was consistent across load conditions. This demonstrates that internal lexical competition – the activation of phonologically similar sounding neighbors – operates even under cognitive constraints. These results support a distinction between lexicon-internal mechanisms, which are resilient to cognitive load, and exogenous mechanisms, which are susceptible. This finding is consistent with theories of spoken production emphasizing parallel activation and competitive selection within the lexicon (Dell, 1986; Howard et al., 2006; Roelofs, 1992; Wheeldon & Monsell, 1994).

### *Interaction of Load with Context and Lexical Competition*

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

The differential effects of phonological load on contextual cues versus onset-differing neighbors (VC) provide insights into how cognitive load selectively affects speech production. While internal competition processes remain intact, phonological load disrupts the use of external cues (context), suggesting that load primarily affects top-down processing rather than bottom-up, lexically driven mechanisms.

This distinction mirrors broader findings in psycholinguistics demonstrating that different types of competition – internal versus external – are differentially sensitive to cognitive constraints. In line with Lowe and Buchwald's (2023) work, which distinguishes between prepotent and underdetermined forms of conflict in language production, the current results suggest that lexicon-internal competition (VC) resembles underdetermined conflict – an automatic, lexicon-internal processes that the system can resolve without extensive executive control. In contrast, exogeneous context effects resemble prepotent conflict, where external information competes with lexical activation and requires top-down control to resolve. Under phonological load, when resources are taxed, this control-dependent pathway is weakened, diminishing the influence of context on production.

This division between lexicon-internal and exogeneous mechanisms echoes other dual-route frameworks in cognitive science, such as models of reading that differentiate between automatic lexical access and controlled grapheme-to-grapheme conversion routes. Just as skilled readers can flexibly rely on one route or the other depending on task demands, speakers may draw on both automatic and lexicon-internal mechanisms and more controlled, context-driven processes during production. The present findings suggest that these dual pathways are not equally resilient under cognitive load: while the automatic route remains robust, the control-dependent route is more vulnerable when attentional resources are limited.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

All together, these findings support a model of speech production in which lexicon-internal mechanisms and domain-general contextual cues contribute independently to VOT, but cognitive load selectively disrupts external influences. This distinction highlights the resilience of internal, lexicon-internal processes even under cognitive load, while top-down guidance from context is compromised. These results have implications for understanding how speakers manage language in real-world, multitasking environments, illustrating that different sources of competition rely on distinct cognitive resources and are differentially sensitive to load.

**Plating up: Bringing it all together**

My dissertation project investigated how cognitive load influences lexical access across three complementary methodologies: eye-tracking in the visual world paradigm (VWP), computational simulation using the TRACE model, and voice onset time (VOT) analyses in a continuous production task. Together, these findings converge on a nuanced picture of the relationship between cognitive demands and lexical access – one in which different types of load selectively affect lexical processing, and in which listeners demonstrate resilience in recognizing and producing target words even under challenging conditions.

Study 1 examined the effects of cognitive load on spoken word recognition. I used eye-tracking to examine the time course of lexical competition under cognitive load. Here, I saw that phonological load did not differentially differ in its effect on eye fixations than the sham load condition.

Study 2 used TRACE simulations to explore how inhibition is instantiated within the lexical system by modeling both domain-general (degraded input via increased bias) and lexicon-internal load (increased lexical competition via pre-activated memory items). The simulations revealed a striking dissociation: domain-general load delayed and reduced target activation across the board, whereas lexicon-internal load left target recognition largely intact while suppressing competitor activation. This pattern suggests that the locus of load effects critically depends on whether load targets input processing or lexical competition mechanisms.

Study 3 explored the effects of cognitive load on spoken word production. By examining voice onset times (VOTs), I found that phonological load consistently reduced VOTs compared to sham load. I was able to dissociate internal and external load affects and found that internal

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

lexical mechanisms remain robust under load. These findings inform the maintenance of activation of lexical candidates, offering new insights into the dynamics of word retrieval under varying cognitive demands.

### **Distinct mechanisms of load effects**

A central contribution of this research is the demonstration that cognitive load is not a single construct - the distinction between domain-general and lexicon-internal load, operationalized both behaviorally and computationally, maps onto different effects on the speech processing system.

Domain-general load (whether implemented as increased beta parameters in TRACE, or context in spoken production) affects the quality of bottom-up information flowing into the lexical system. When input is degraded, all lexical candidates receive weaker and noisier activation, slowing the accumulation of evidence for any particular word. This is consistent with resource-based accounts suggesting that perceptual encoding requires attentional resources and that load reduces the signal-to-noise ratio of acoustic representations (Mattys et al., 2011).

Lexicon-internal load (implemented as a concurrent memory task requiring maintenance of phonological information, or varying onset-differing neighbors (VC)) operates through a different mechanism. Rather than degrading input, maintaining items in working memory appears to introduce additional activation at the lexical level, creating competition that must be resolved through lateral inhibition. Targets, which receive overwhelming acoustic evidence, are relatively immune to this increased competition, while cohort competitors – activated on the basis of partial overlap and thus more dependent on the dynamics of lexical competition – showed suppressed activation. This interpretation aligns with models of spoken word recognition

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

that emphasize the role of lateral inhibition in resolving competition among activated candidates (McClelland & Elman, 1986; Luce & Pisoni, 1998). When working memory load introduces additional active representations into the system, the inhibitory dynamics that normally resolve competition become more complex. The result is not a global slowdown but rather a selective reduction in the prominence of competitors that lack decisive bottom-up support.

Future research in phonological load can vary its level to characterize its effect on language processing. The present studies used fixed load levels (4 items) but perhaps this wasn't challenging enough. Does the effect plateau at some load level, or do effects continue to increase with 5, 6, or more items? Manipulations of load would address these questions and could reveal ceiling effects in the lexical system's capacity for managing concurrent demands.

### **The resilience of target recognition**

Across Studies 1 and 2, a consistent finding was the remarkable stability of target word processing. Target fixations in the visual world paradigm were largely preserved and simulated target activation showed minimal effects of lexicon-internal load. This resilience suggests that the language system prioritizes the recovery of intended targets, perhaps through mechanisms that protect goal-relevant processing from interference.

One possibility is that target recognition benefits from top-down expectations and contextual support that competitors do not receive. In typical language situations, targets are predicted by semantic context, syntactic structure, and communicative goals. This top-down support may buffer target processing against the effects of load while leaving competitors, which lack such support, vulnerable to interference.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Another possibility is that the temporal dynamics of spoken word processing naturally protect targets. Because targets receive continuous bottom-up support as the acoustic signal unfolds, they accumulate activation steadily over time. Competitors, in contrast, are activated transiently and must be suppressed as disambiguating information arrives. Under load, this suppression may be less efficient, but the fundamental advantage of targets remains intact.

### **Bridging perception and production**

The inclusion of both perception (Studies 1-2) and production (Study 3) in this project allowed us to have a comprehensive investigation of whether cognitive load effects can generalize across two modalities of language use. Speech perception and production have been shown to be linked, sharing phonological representations and neural substrates (Hickok & Poeppel, 2007). If cognitive load affects shared phonological resources, similar patterns might emerge across modalities.

The tasks employed in both experimental studies were also substantially different. Although I used a visual world paradigm (VWP) task in Study 1, Zhang and Samuel (2015) used an auditory lexical decision task for their load manipulation. The lexical decision task requires explicit binary judgment and may invoke different strategies than the implicit processing measured in the VWP. This raises the possibility that the null load effects I observed reflect not the absence of interference between phonological working memory and lexical processing, but rather specific characteristics of how the VWP captures lexical competition dynamics.

In lexical decision and naming tasks, the phonological properties of the stimulus are directly relevant to the response: participants must either make metalinguistic judgments about lexicality or produce phonological output. This direct task relevance may create stronger

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

interference with concurrent phonological working memory maintenance. In the VWP, by contrast, the task requires visual selection rather than phonological output or metalinguistic judgment, potentially reducing the degree to which phonological working memory and task performance compete for overlapping processing operations.

The VWP may be particularly resilient to working memory interference for several reasons. First, it provides visual context that could support lexical processing through multimodal integration, potentially compensating for resource limitations imposed by concurrent load. When participants can see potential referents, the visual information may scaffold lexical access in ways that reduce reliance on purely phonological processing. In contrast, lexical decision and naming tasks require participants to rely more exclusively on phonological and lexical representations without visual support, potentially making these tasks more vulnerable to phonological working memory interference.

Beyond these task-specific factors, however, the divergent findings across Studies 1 and 3 may reflect a more fundamental asymmetry between language comprehension and production. In comprehension, listeners do not need to achieve exclusive lexical access to succeed – multiple lexical candidates can remain partially activated without catastrophic consequences for understanding (Marslen-Wilson, 1987). ERP evidence increasingly suggests that phonological neighbors remain neurally active well after word recognition, producing measurable neural signatures even when they do not influence immediate behavioral responses (Desroches, et al., 2009; Dufour et al., 2012). In the VWP specifically, participants can rapidly transition from phonological processing to lexical-semantic representations: once the referent is activated, the phonology can effectively be "offloaded" in service of the ultimate goal: visual selection.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Production, in contrast, demands exclusivity. Speakers must ultimately commit to a single phonological form; lingering competitor activation risks producing the wrong word entirely. This selection pressure means that production necessarily maintains sustained engagement with phonological representations through to articulation. Critically, in Study 3, participants faced a dual phonological burden: they had to maintain phonological information for the concurrent load task while simultaneously sustaining phonological representations for speech planning and output. This convergence on the same representational level may create an unavoidable bottleneck. In comprehension, these demands appear more readily dissociable – listeners can progress toward meaning while phonological details fade – but production's requirement for phonological commitment leaves fewer opportunities for strategic compensation when phonological working memory is taxed.

The current findings, when considered alongside Zhang and Samuel's (2017) results, suggest that lexical access may be divided into early (frequency sensitivity, neighborhood effects) and later (measured by visual fixation patterns) stages. These stages may also differ in their vulnerability to cognitive load depending on whether language is being comprehended or produced. In comprehension, the goal is ultimately to arrive at meaning—and once a referent is sufficiently activated, phonological details can fade without consequence. Production, however, requires sustained commitment to a specific phonological form through to articulation; competitors cannot simply be left "floating" without risking overt errors. This fundamental asymmetry may explain why cognitive load effects emerged in production but not perception: comprehension affords flexibility to offload phonological processing, whereas production creates an unavoidable bottleneck when both the memory task and the primary task converge on phonological representations. Future research can further investigate how the VWP's extended

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

processing window and multimodal support enable strategic compensation that may be unavailable in production or in more constrained perceptual tasks, and whether the locus of load effects shifts predictably across the time course of lexical access in each modality.

### **Prioritizing goals**

The present experiments required participants to perform two tasks concurrently: maintaining phonological information in working memory while processing language (Studies 1 and 3). This dual-task situation creates potential for strategic trade-offs in resource allocation. Participants may prioritize one task over the other, and this prioritization could vary across individuals and across trials within individuals.

In the present studies, participants were instructed to attend to both tasks, but the natural priority of speech comprehension and production may have led to relative protection of language processing at the expense of memory maintenance. If participants strategically allocated resources to language tasks when processing demands were high, this could explain the resilience of target processing observed across studies. The cost of this prioritization would be reflected in memory task performance, which showed expected effects of load level.

This strategic allocation raises important questions about ecological validity. In real-world multitasking, individuals often must make explicit or implicit decisions about which task to prioritize. A driver engaged in conversation must balance attention to speech with attention to the road; a student taking notes during a lecture must balance comprehension with transcription. The present findings suggest that when faced with competing demands, language users may naturally protect core language processing - but this protection may come at a cost to other concurrent activities.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Individual differences in strategy and resource allocation may also contribute to the variability observed in the eye-tracking data. Some participants may have adopted strategies that emphasized speech processing, while others may have attempted to balance both tasks more evenly. This variability, rather than simply reflecting noise, may provide insight into the flexibility of the cognitive system in managing competing demands.

This variability was most notable in the eye-tracking data from Study 1. Some participants seemed to keep their gaze on the center fixation point, even when responding highly accurately to the VWP task. This is indicative of a different strategy being used to go through the experiment and could also be a sign of some participants' fields of vision being wider than others in order to keep all of the images peripherally in view without blatantly looking from one to the other.

Future research could manipulate task priority explicitly, instructing participants to emphasize either memory maintenance or speech processing. Such manipulations would reveal the extent to which the observed patterns reflect strategic allocation versus structural limitations of the cognitive system.

### **Conclusion**

The present research demonstrates that cognitive load affects spoken language processing in selective and specific ways. These findings have several important implications. First, they challenge the view that cognitive load necessarily impairs language processing, suggesting instead that the language system is remarkably resilient under concurrent cognitive demands. Second, they demonstrate that different sources of lexical competition rely on distinct cognitive resources, supporting models that distinguish multiple mechanisms of competition resolution.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Third, they highlight the importance of considering how language processing operates in naturalistic multitasking environments, where people routinely manage multiple cognitive demands while comprehending and producing language.

The convergence of findings across behavioral, computational, and production methodologies strengthens confidence in these conclusions. Eye movements during spoken word recognition, activation dynamics in TRACE, and temporal measures of speech production all point to a system that prioritizes target processing while allowing peripheral aspects of lexical activation to be modulated by available resources.

Together, the experiments and simulations reported here converge on a picture of spoken language processing as a robust, adaptive system that prioritizes recovery and production of communicatively relevant information even under challenging conditions. Understanding the limits and mechanisms of this resilience remains an important goal for future research, with implications for theory, application, and our understanding of how language functions in the complex cognitive ecology of everyday life.

**Appendix A. Study 1 Supplemental Information****Table A1***Study 1 stimuli grouped by competitor set*

Set #	Target	Cohort	Rhyme	Unrelated
1	bath	bass	path	couch
2	beak	beet	sneak	map
3	bear	base	pear	Jet
4	boot	boom	suit	Fox
5	cage	cave	gauge	Hip
6	coat	cone	vote	Ram
7	crown	crowd	drown	soup
8	dent	desk	tent	brush
9	gum	gut	drum	whale
10	hole	hose	goal	Cap
11	horn	horse	corn	Bib
12	lab	lamb	crab	Tire
13	lips	list	chips	tape
14	mouse	mouth	house	chain
15	mug	mud	pug	cool
16	night	knife	bite	jar
17	pick	pit	kick	deer
18	porch	pork	torch	milk

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

19	rip	rib	ship	dog
20	rose	robe	nose	pool
21	sick	sip	wick	door
22	sock	sod	dock	lap
23	throne	throat	crone	dish
24	trap	trash	nap	fist
25	tug	tub	rug	mic
26	type	tile	pipe	coach
27	wheat	wheel	seat	gem
28	wig	wind	fig	buzz
29	yarn	yard	barn	safe
30	zit	zip	sit	coal
31	batter	baggage	ladder	peacock
32	berry	barrel	fairy	rapids
33	carrot	carriage	parrot	tadpole
34	cavern	cashew	tavern	banner
35	coffee	coffin	toffee	knuckle
36	dollar	dolphin	collar	hammock
37	letter	lettuce	sweater	cannon
38	money	mother	honey	beagle
39	mountain	mousetrap	fountain	target
40	mustard	mustache	custard	penguin
41	paddle	package	saddle	monkey

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

42	pickle	picture	nickel	donkey
43	rocket	rocker	pocket	bubble
44	sandal	sandwich	candle	building
45	socket	soccer	locket	filling
46	tailor	table	sailor	candy
47	tower	towel	shower	hamster
48	turtle	turkey	hurdle	banjo
49	wizard	whistle	lizard	bottle
50	dragon	dragster	wagon	peanut
51	funnel	fungus	tunnel	blanket
52	hockey	hotdog	jockey	campus
53	windmill	window	treadmill	badger
54	robber	robin	bobber	necklace
55	magnet	magic	bonnet	camera
56	powder	power	chowder	billboard
57	pillow	pillar	willow	sunrise
58	beaver	beehive	cleaver	children
59	castle	cabin	tassel	water
60	butter	button	putter	camel

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Table A2***Load Task Average RT and Load Task Accuracy by Participant*

Participant-Level Summary by Load Condition				
Participant	Load Condition	Average RT (ms)	Load Task Accuracy (%)	
P12	Phon	988	79.4	
P14	Phon	1,193	92.7	
P16	Phon	1,062	87.7	
P18	Phon	1,115	89.3	
P22	Phon	1,042	97.3	
P23	Phon	911	95.5	
P25	Phon	1,204	96.4	
P26	Phon	1,015	92.9	
P39	Phon	1,048	91.2	
P46	Phon	1,234	95.6	
P48	Phon	1,163	84.7	
P49	Phon	1,207	92.9	
P66	Phon	1,131	93.0	
P67	Phon	1,223	81.5	
P70	Phon	1,024	93.9	
P72	Phon	1,301	94.9	
P92	Phon	1,253	94.5	
P95	Phon	1,199	87.7	
P11	Sham	882	99.0	
P30	Sham	1,054	99.8	
P34	Sham	1,102	100.0	
P35	Sham	1,044	100.0	

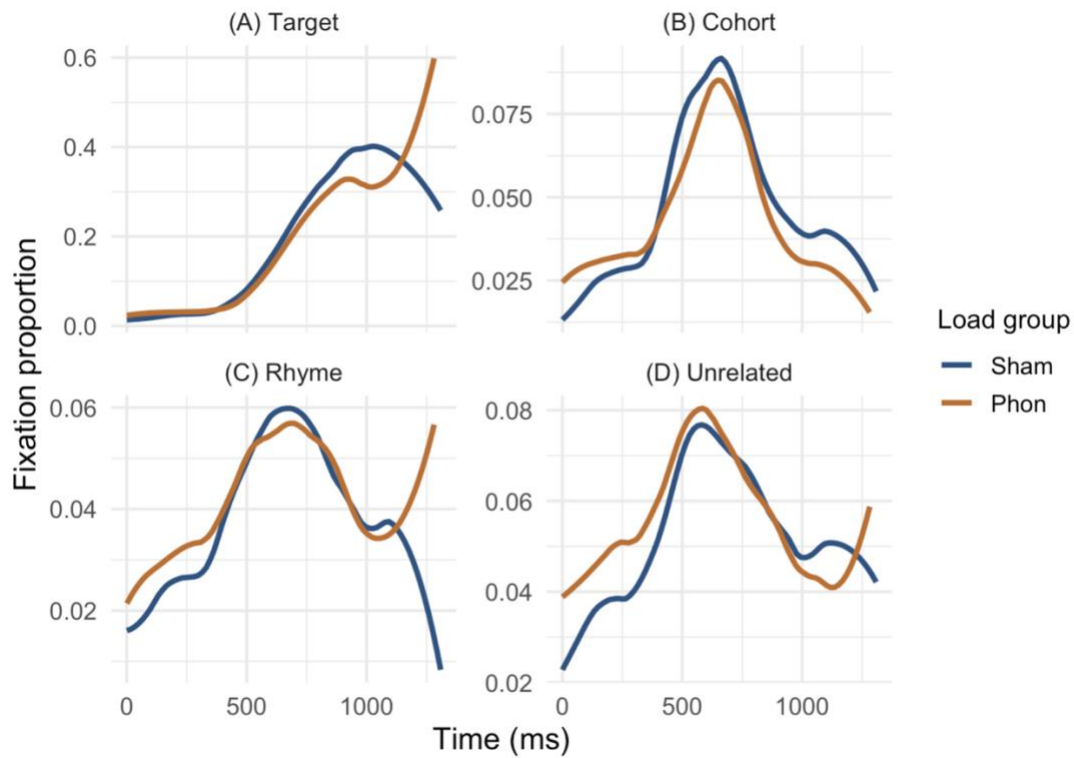
## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

P41	Sham	853	99.4
P42	Sham	943	99.6
P43	Sham	1,106	100.0
P45	Sham	1,217	99.6
P53	Sham	927	99.4
P58	Sham	1,314	99.8
P62	Sham	1,005	99.0
P75	Sham	1,084	99.8
P76	Sham	977	99.8
P78	Sham	1,323	99.8
P87	Sham	1,145	98.1
P88	Sham	808	99.8
P91	Sham	1,067	99.6
P98	Sham	1,051	98.9

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A1.**

*Proportion of fixations to each competitor type as a function of time and listener group*

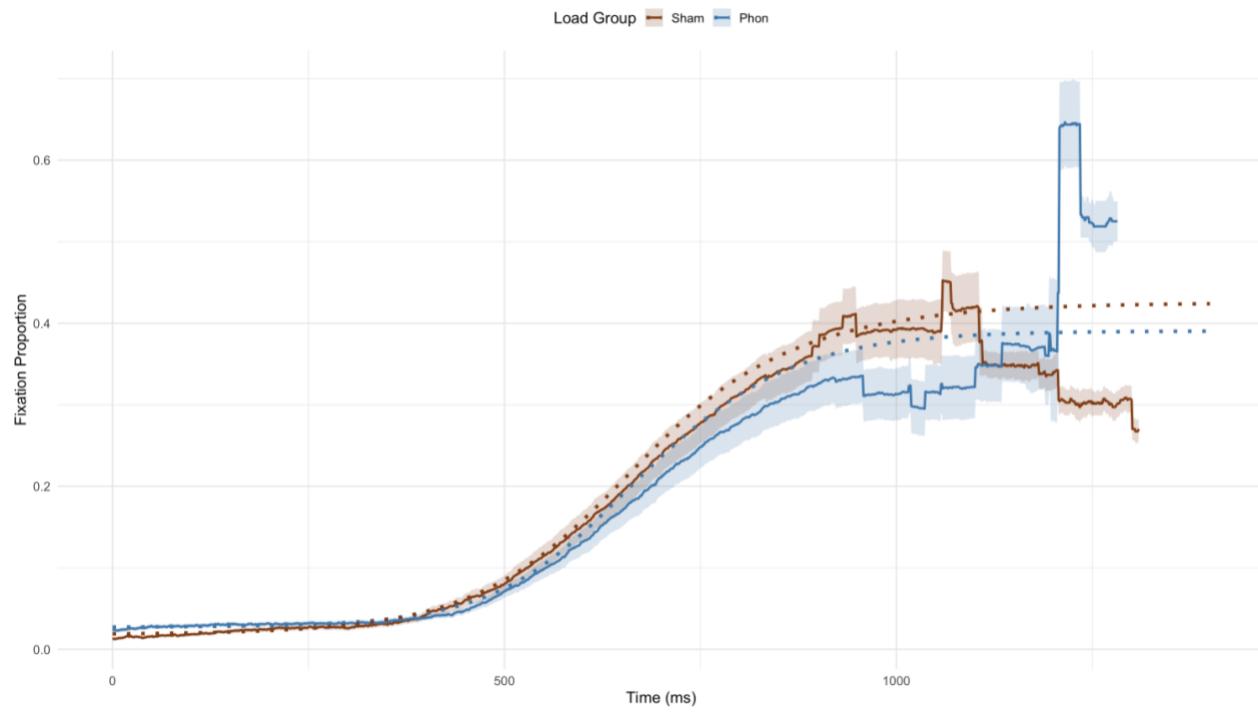


*Note.* (A) Target, averaged across TCRU, TCUU, and TRUU trials; (B) Cohort, averaged across TCRU and TCUU trials; (C) Rhyme, averaged across TCRU and TRUU trials; (D) Unrelated, averaged across TCRU, TCUU, TRUU trials.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A2.**

*Group average target fixations (observed vs. fitted curves)*

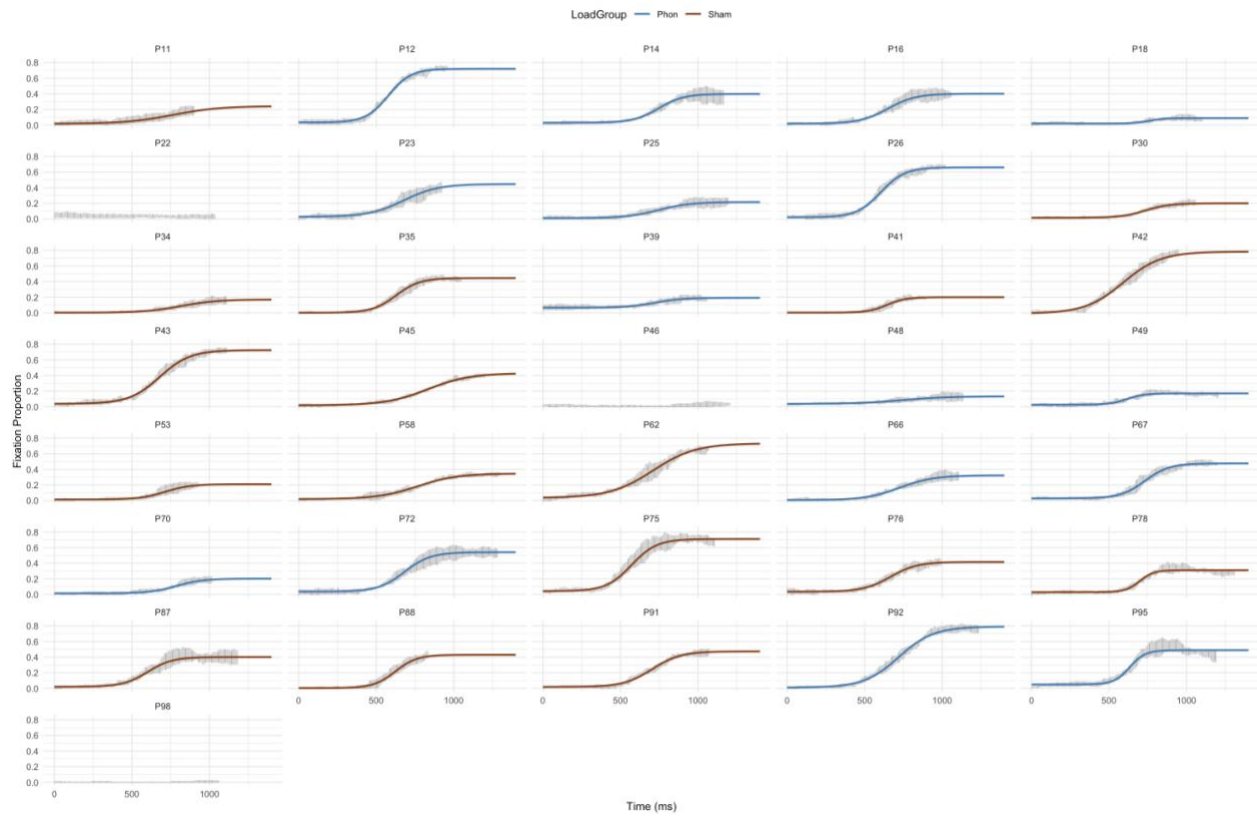


*Note:* Solid lines = observed data (+/- SE), Dashed lines = fitted curves

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A3.**

*Individual participant fits*

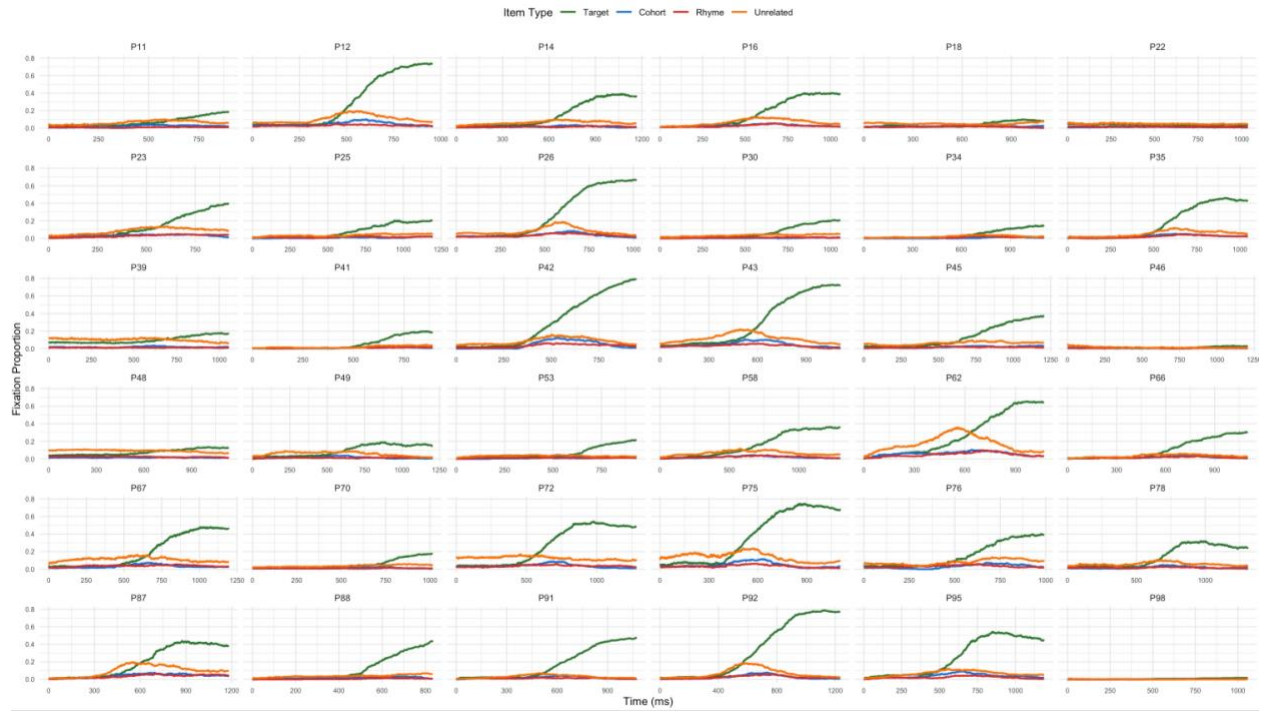


*Note:* Gray = observed data, Colored = fitted logistic curves

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A4.**

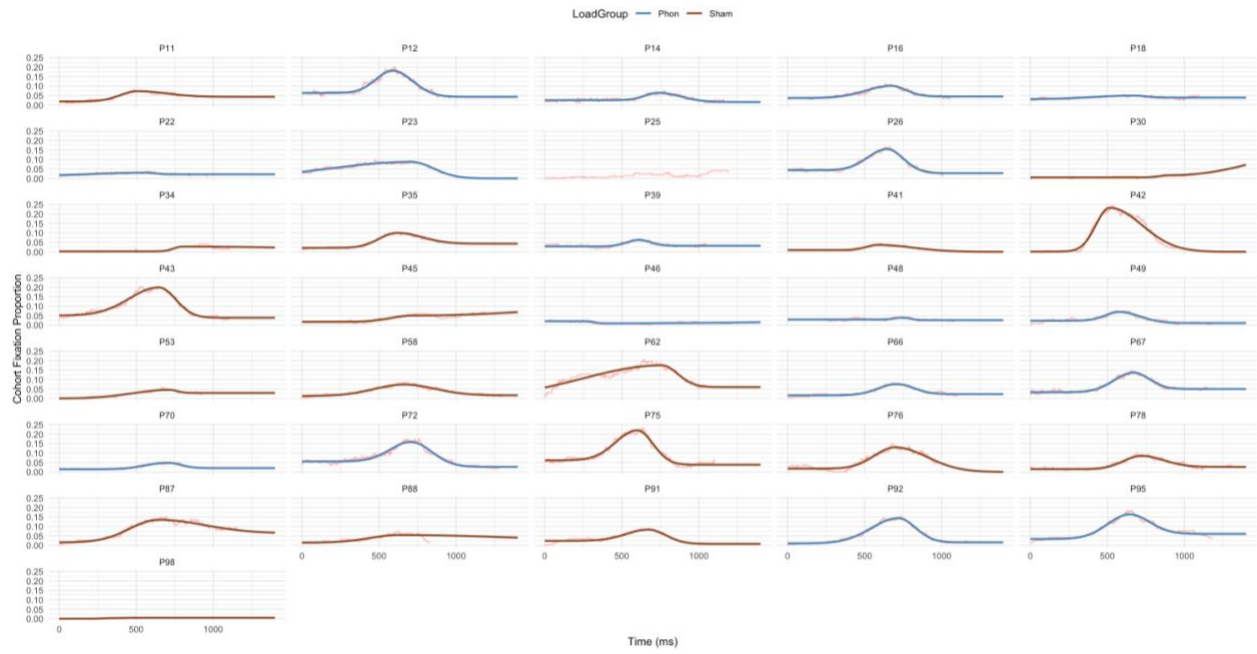
*Individual Target, Cohort, Rhyme, and Unrelated fixations*



# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A5.**

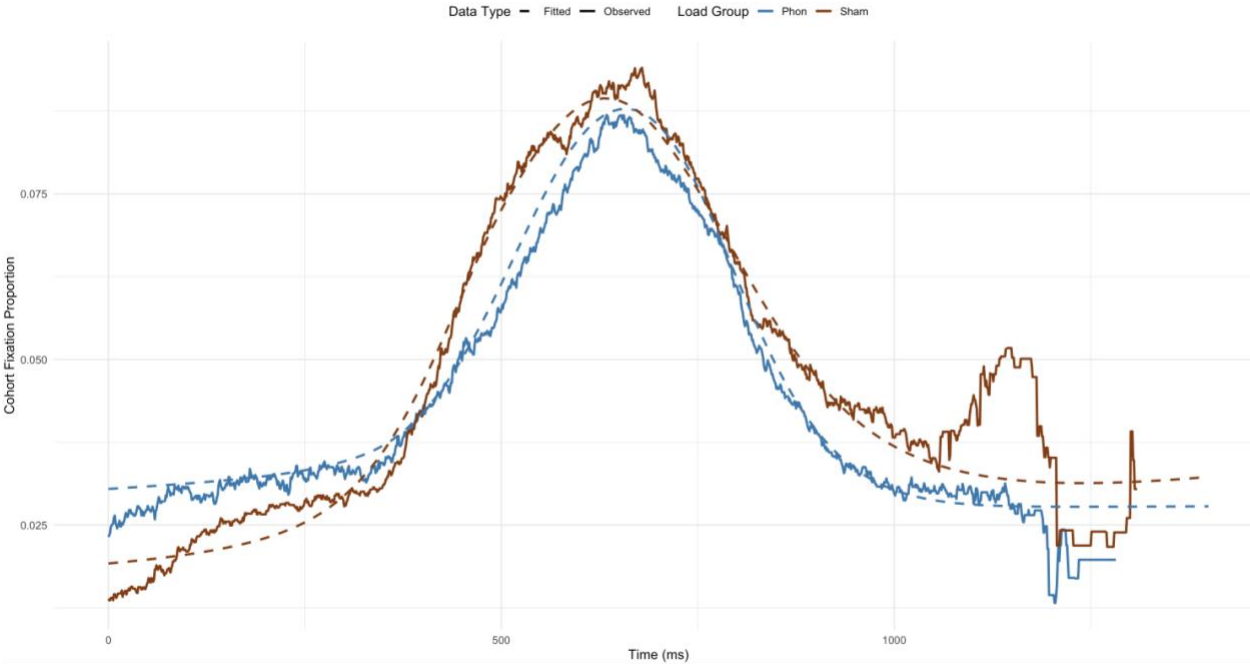
*Individual Cohort competitor fits*



*Note:* Gray = observed data, Colored = fitted logistic curves

Figure A6.

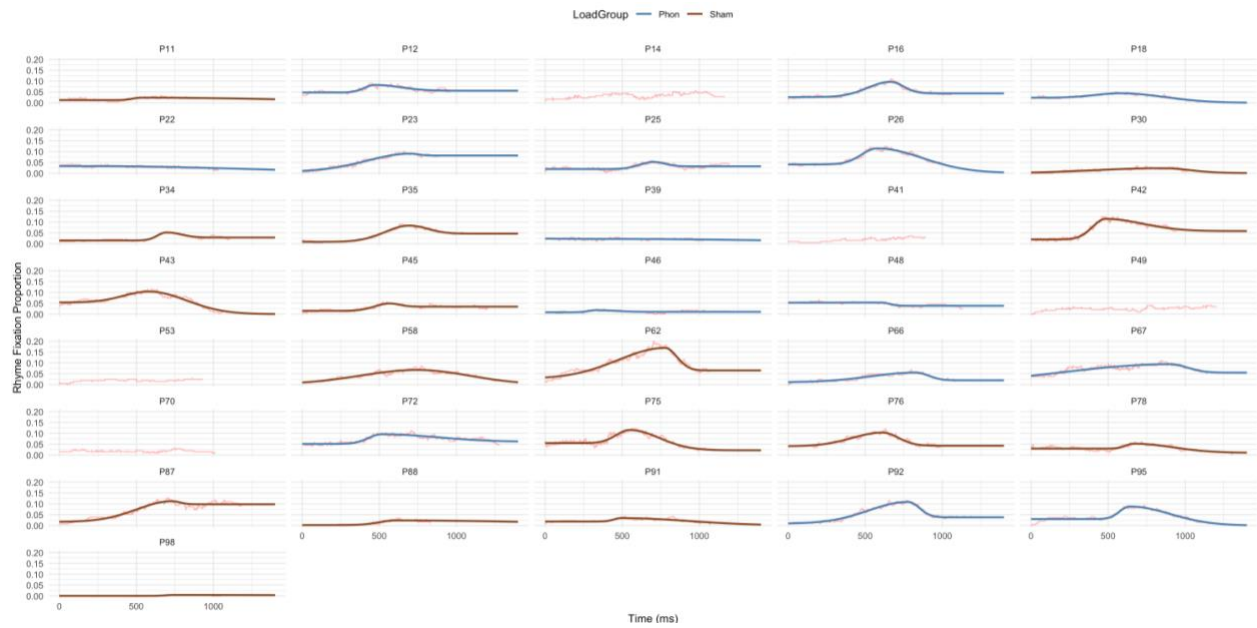
Group Cohort competitor fits



# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A7.**

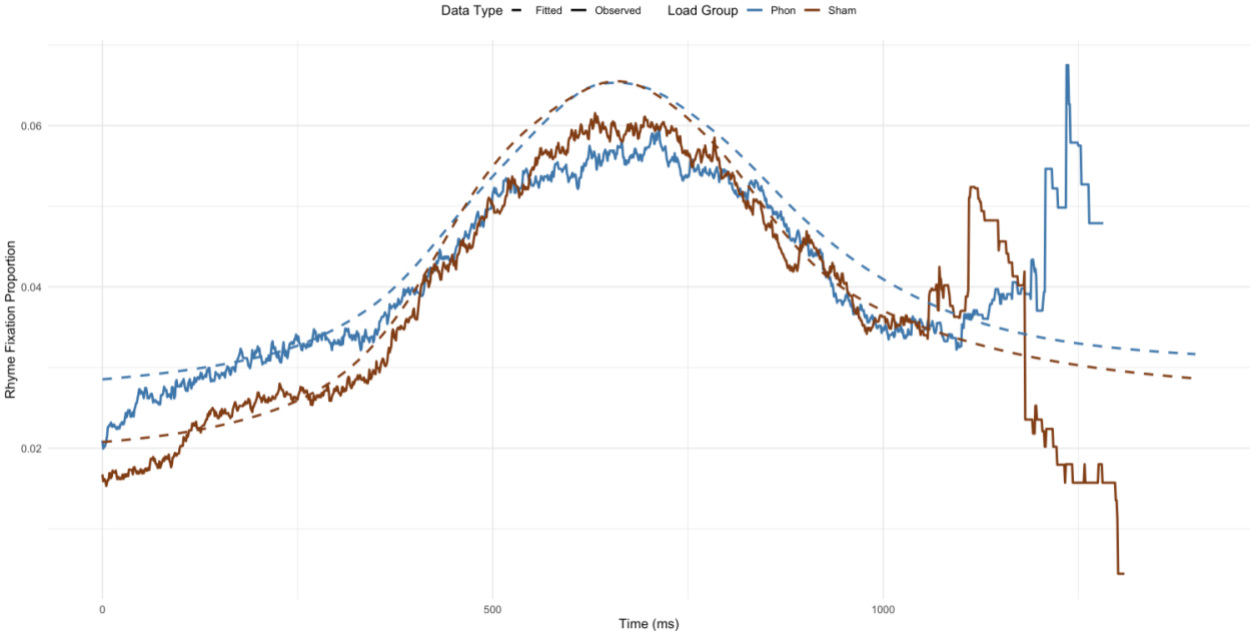
*Individual Rhyme competitor fits*



*Note:* Gray = observed data, Colored = fitted logistic curves

Figure A8.

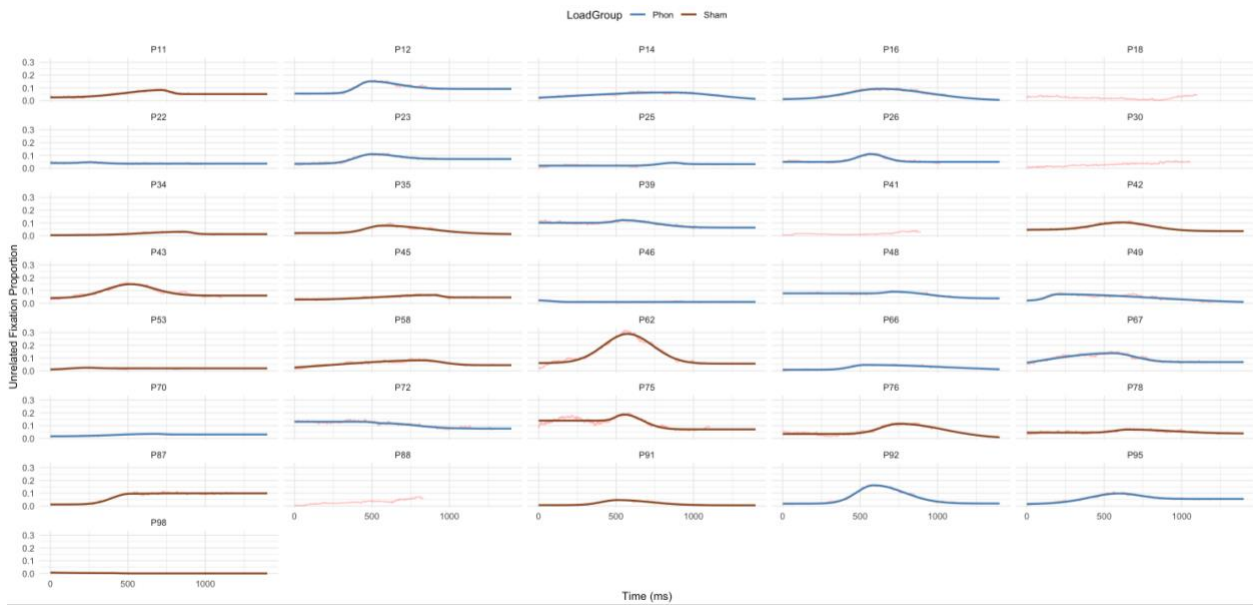
*Group Rhyme competitor fits*



# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A9.**

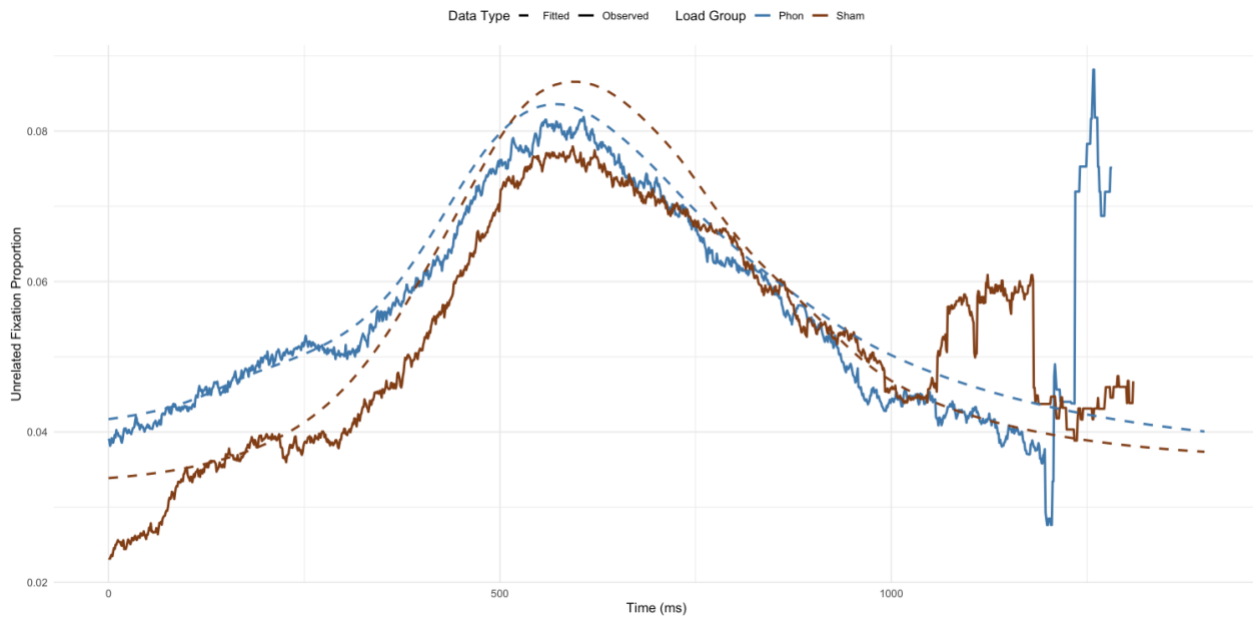
*Individual Unrelated competitor fits*



*Note:* Gray = observed data, Colored = fitted logistic curves

**Figure A10.**

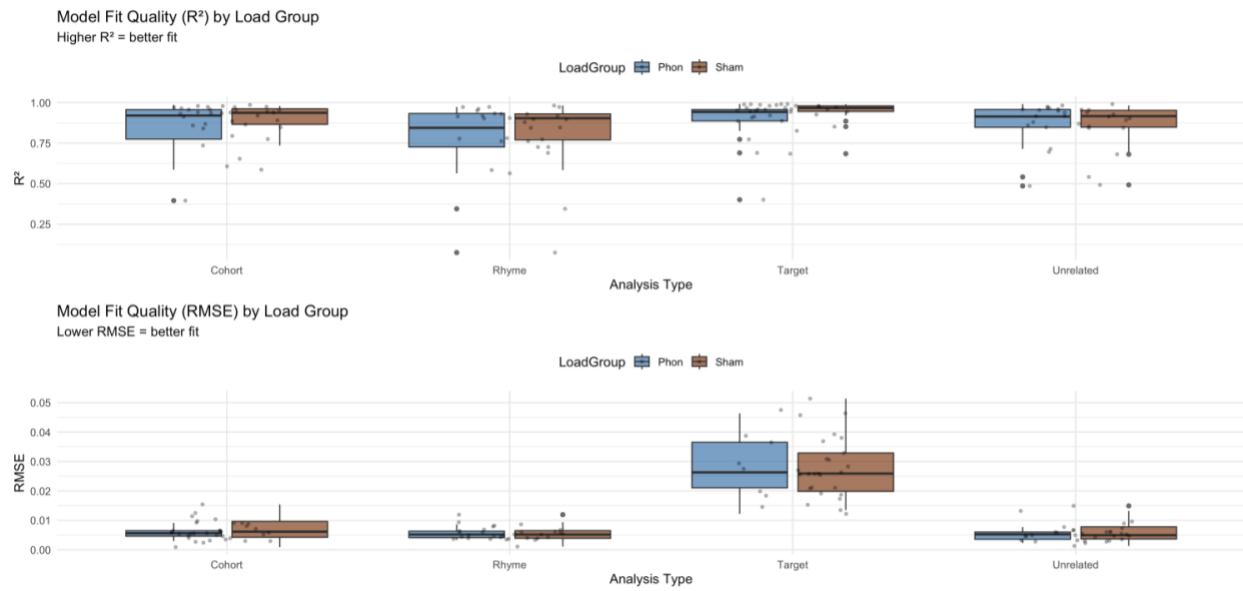
*Group Unrelated competitor fits*



# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure A11.**

*Model fit quality comparison across all analyses*



LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

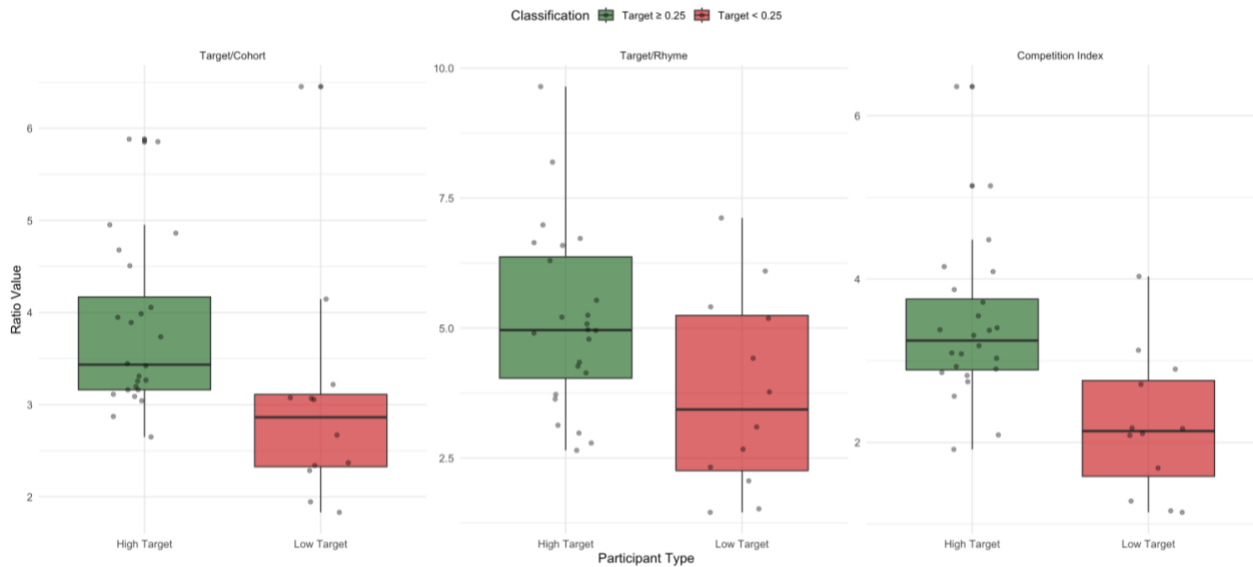
**Table A3**

*Summary table for model fit*

Analysis	LoadGroup	N	R2_mean	R2_SD	RMSE_mean	RMSE_SD
Cohort	Phon	17	0.838	0.174	0.00572	0.00193
Cohort	Sham	18	0.910	0.0696	0.000692	0.00389
Rhyme	Phon	15	0.760	0.256	0.00529	0.00146
Rhyme	Sham	16	0.856	0.111	0.00564	0.00265
Target	Phon	17	0.884	0.149	0.0280	0.0101
Target	Sham	17	0.939	0.0753	0.0282	0.0109
Unrelated	Phon	17	0.861	0.147	0.00502	0.00160
Unrelated	Sham	15	0.863	0.138	0.00625	0.00388

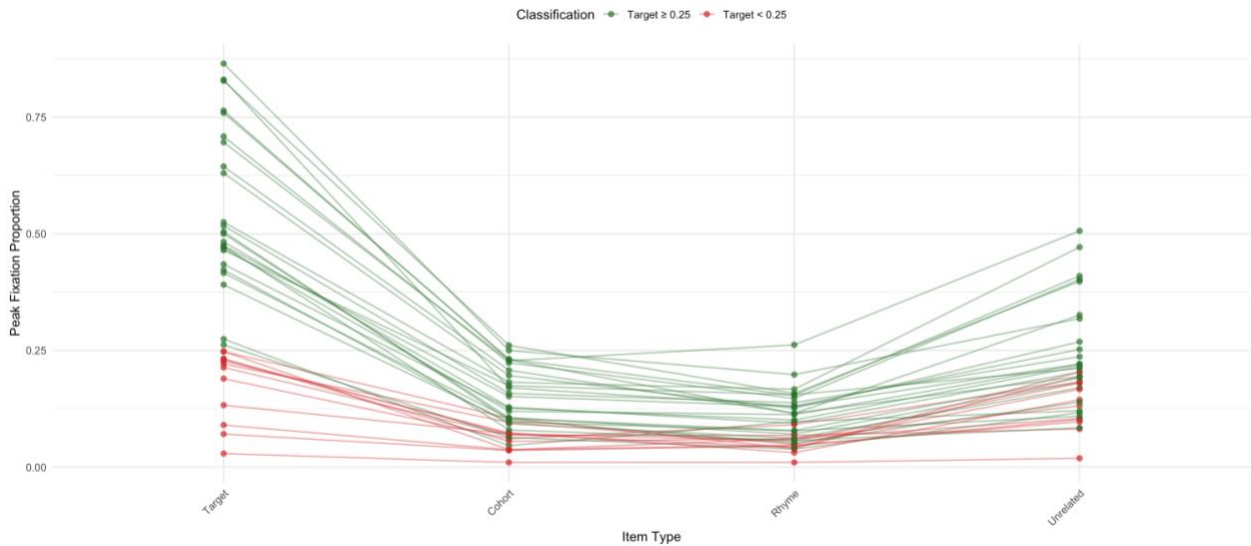
**Figure A12.A - Exploratory analyses visualizations**

*Target-Competitor ratios by target peak*



**Figure A12.B - Exploratory analyses visualizations**

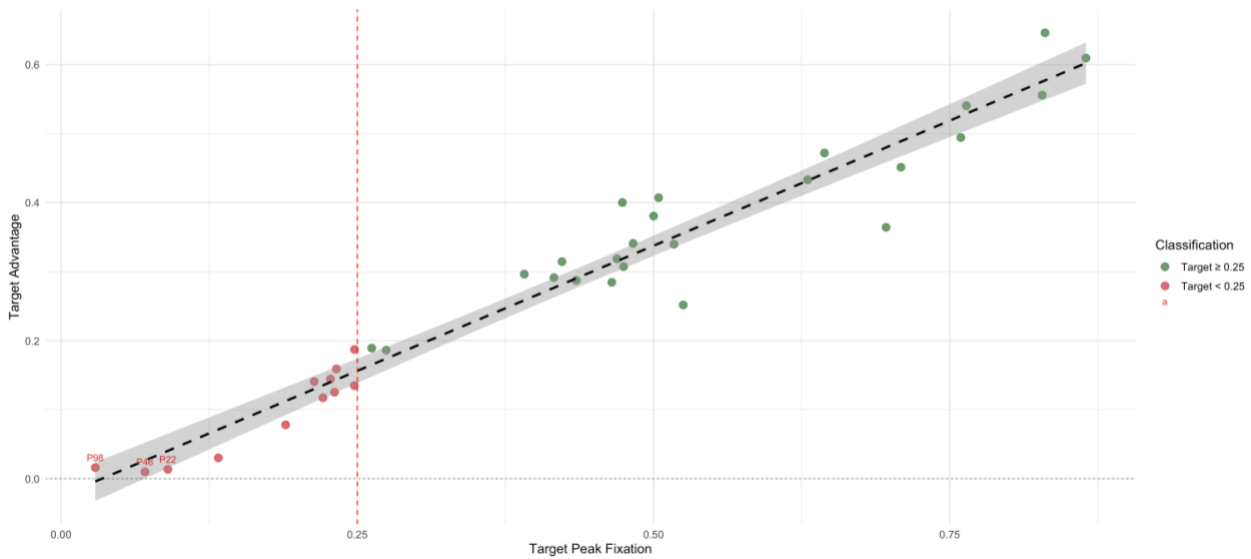
*Individual participant fixation profiles*



*Note:* Lines connect fixation peaks for each participant.

**Figure A12.C - Exploratory analyses visualizations**

*Target peak vs. Target Advantage*



*Note:* Target advantage = Target Peak – Average Competitor Peak ( $r = 0.97$ )

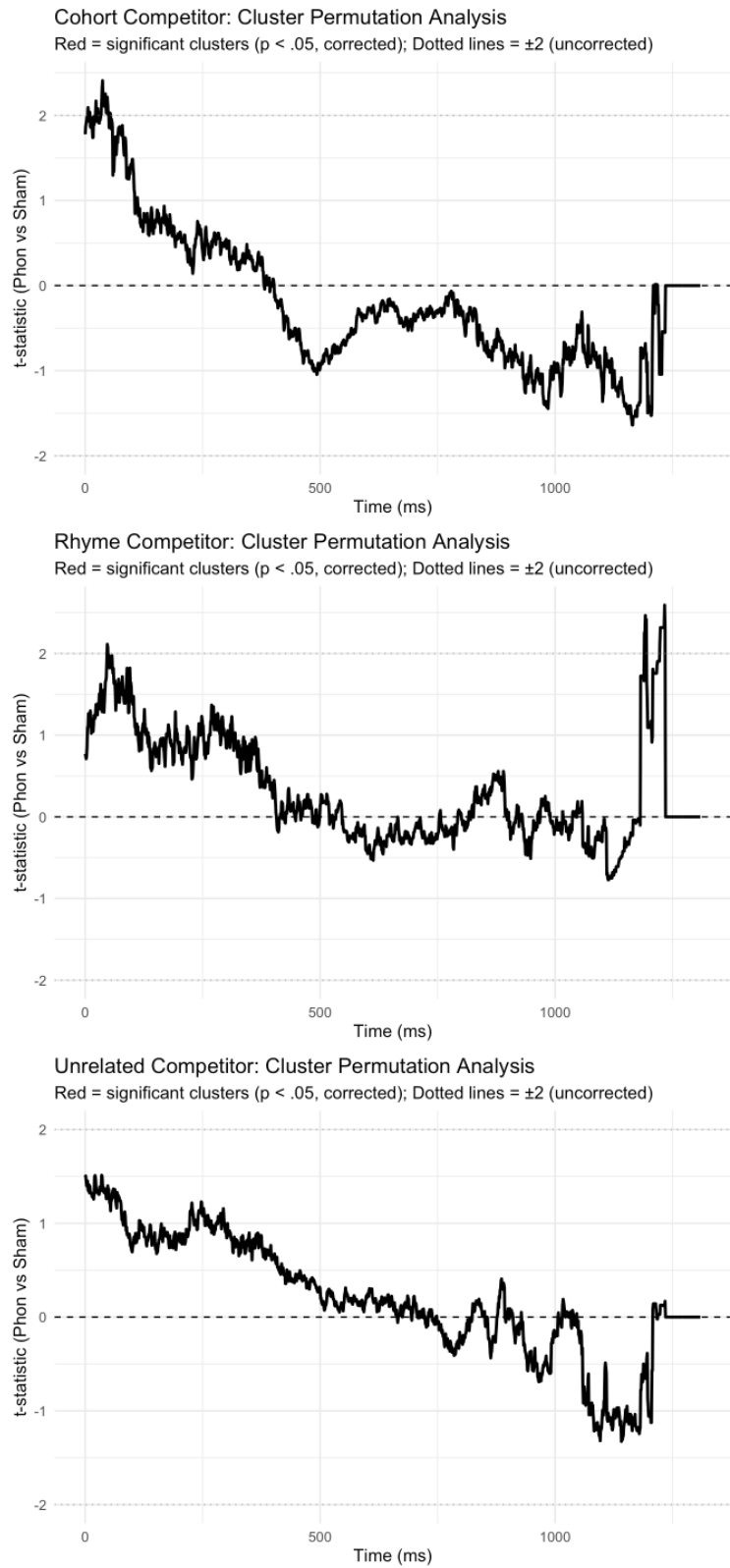
LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Table A4**

*Summary of Bayes Factors across fixation types and parameters*

<b>Parameter</b>	<b>Target</b>	<b>Cohort</b>	<b>Rhyme</b>	<b>Unrelated</b>
Crossover/Mu	2.97	2.94	2.73	2.26
Slope/Sigma <sub>1</sub>	3.01	2.94	1.48	2.02
Peak	2.76	2.84	2.52	2.84
b <sub>1</sub>	-----	0.52*	2.10	2.86
Sigma <sub>2</sub>	-----	1.34	1.30	1.29
b <sub>2</sub>	-----	2.18	2.24	2.84

Figure A13 – Cluster Permutation visuals



## Appendix B. Study 2 Supplemental Information

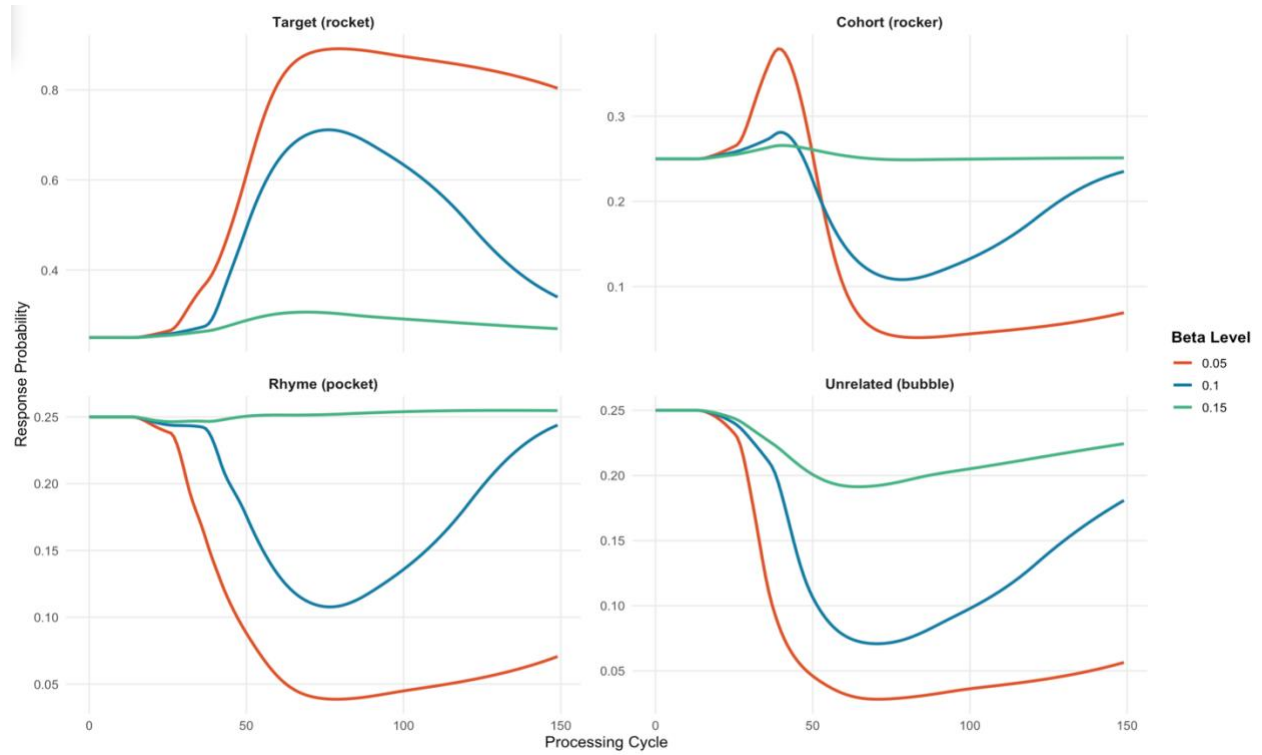
Table B1

*Summary of simulations*

Hypothesis/Condition	Description	jTRACE Parameter	jTRACE implementation	Simulated Effect
Baseline	No added cognitive load; normal lexical access without interference	Default parameters	Standard lexicon, default parameters, no additional load	Lexical activation proceeds typically
Load drains cognitive resources	Load acts as an external resource drain, reducing available activation for lexical processing	<i>Beta</i> (activation decays faster)	Increase <i>beta</i> parameter reduces resources available for activation	Load reduces ability to maintain activation across items.
Load adds lexical competitors	Load increases competition inside the lexicon by adding extra lexical items (e.g., letter names) that compete with targets	Additional input (e.g., letter names) given pre-activation	Add letter name nodes as lexical competitors and assign with priming levels to simulate pre-existing activation	Load competes within lexical system like real words

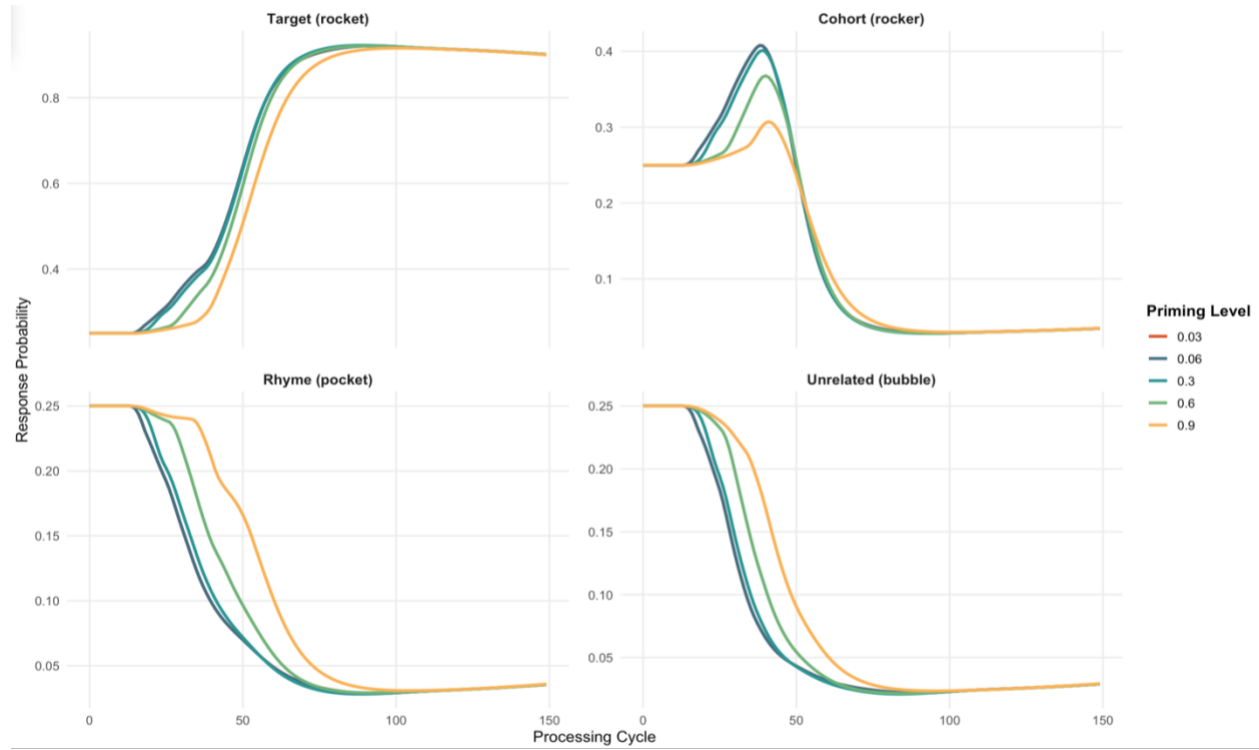
**Figure B1**

*Effect of domain-general load on response probability*



**Figure B2**

*Effect of resting priming on word activation*



## Appendix C. Study 3 Supplemental Information

Table C1

*Study 3 stimuli*

	<b>Minimal pair</b>	<b>No minimal pair</b>
<b>Labial stop</b>	peek	peel
	palm	pomp
	punk	pulp
	punch	pulse
	pun	pup
	pad	pal
	peat	peal
	pare	pep
	pig	pith
	poll	poach
	pox	posh
	putt	pub
<b>Alveolar stop</b>	tab	tat
	tan	tag
	teal	teat
	teem	teethe
	tick	tiff
	tuck	tuft

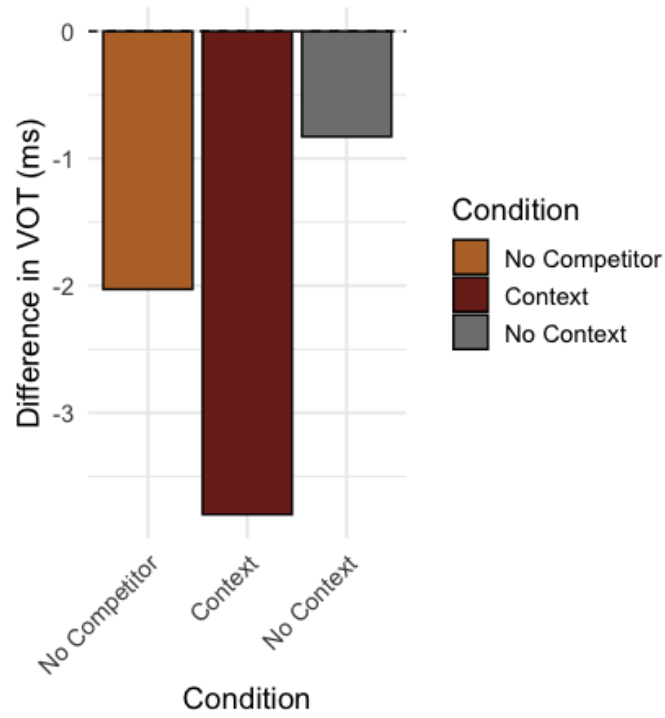
LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

	tense	tenth
	torque	torn
	tomb	tooth
	tame	taint
	tile	tights
	tote	toad
<b>Velar stop</b>	cob	cog
	cod	cop
	kilt	kin
	kit	kiln
	core	corn
	cuss	cub
	cuff	cud
	curl	curb
	coo	coot
	cab	cad
	cape	cake
	code	comb

# LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Figure C1**

*Difference in Mean VOT: Phon Load – No Load*



## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Table C2.***Sham Load Average VOT and Load Task Accuracy by Participant*

Participant ID	Average VOT (ms)	Accuracy (%)
1	83.74	100
2	79.33	100
3	63.83	100
4	98.81	100
5	55.62	100
6	56.37	100
7	62.99	100
8	97.62	100
9	69.61	100
10	112.51	100
11	82.78	100
12	89.26	100
13	72.50	100
14	76.02	100
15	69.77	100
16	107.44	100
17	88.91	100

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

**Table C3.***Phon Load Average VOT and Load Task Accuracy by Participant*

Participant ID	Average VOT (ms)	Accuracy (%)
1	89.41	70.9
2	79.41	53.8
3	76.32	83.8
4	78.15	41.0
5	86.22	91.7
6	67.40	47.8
7	69.06	86.4
8	75.87	93.4
9	62.51	86.9
10	78.46	91.2
11	90.26	92.6
12	100.16	89.9
13	98.55	88.2
14	57.11	89.7
15	85.66	85.3
16	65.70	91.1
17	69.43	74.9
18	87.36	94.4
19	80.58	51.5

**References**

- Alloppenna, 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.
- Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of verbal learning and verbal behavior*, 14(4), 382-391.
- Badde, S., Myers, C.F., Yuval-Greenberg, S., & Carrasco, M. (2020). Oculomotor freezing reflects tactile temporal expectation and aids tactile perception. *Nat Commun* 11, 3341.
- Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of cognitive neuroscience*, 4(3), 281-288.
- Baddeley A. (2003). Working memory: looking back and looking forward. *Nature reviews. Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and cognitive processes*, 24(4), 527-554.
- Bergen, B., Medeiros-Ward, N., Wheeler, K., Drews, F., & Strayer, D. (2013). The crosstalk hypothesis: why language interferes with driving. *Journal of experimental psychology: general*, 142(1), 119.
- Besner, D. & Roberts, M.A. (2003). Reading nonwords aloud: Results requiring change in the dual route cascaded model. *Psychonomic Bulletin & Review* 10, 398–404.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, 108(3), 624.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Bradlow, A.R. & Alexander, J.A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of Acoustical Society of America*, 121(4), 2339-2349.
- Branigan, H.P., Pickering, M.J., Stewart, A.J. *et al.* (2000). Syntactic priming in spoken production: Linguistic and temporal interference. *Memory & Cognition* 28, 1297–1302.
- Caselli, N. K., & Cohen-Goldberg, A. M. (2014). Lexical access in sign language: A computational model. *Frontiers in Psychology*, 5, 428.
- Caselli, N. K., Emmorey, K., & Cohen-Goldberg, A. M. (2021). The signed mental lexicon: Effects of phonological neighborhood density, iconicity, and childhood language experience. *Journal of Memory and Language*, 121, 104282.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (2022). Access to the internal lexicon. In *Attention and performance VI* (pp. 535-555). Routledge.
- Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive psychology*, 42(4), 317-367.
- 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.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, 93(3), 283.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production:

Lexical access and grammatical encoding. *Cognitive Science*, 23(4), 517-542.

Desroches, A. S., Newman, R. L., & Joanisse, M. F. (2009). Investigating the time course of

spoken word recognition: Electrophysiological evidence for the influences of

phonological similarity. *Journal of Cognitive Neuroscience*, 21(10), 1893–1906.

Dufour, S., Brunellière, A., & Frauenfelder, U. H. (2013). Tracking the time course of word-

frequency effects in auditory word recognition with event-related potentials. *Cognitive*

*Science*, 37(3), 489–507.

Eddine, S. N., Brothers, T., Wang, L., Spratling, M., & Kuperberg, G. R. (2024). A predictive

coding model of the N400. *Cognition*, 246, 105755.

Elzhov, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. (2023). *minpack.lm: R Interface to the*

*Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus*

*Support for Bounds*. R package version 1.2-4. [https://CRAN.R-](https://CRAN.R-project.org/package=minpack.lm)

[project.org/package=minpack.lm](https://CRAN.R-project.org/package=minpack.lm)

Eyelink, SR Research, Ottawa, Canada

Farris-Trimble, A., & McMurray, B. (2013). Test–retest reliability of eye tracking in the visual

world paradigm for the study of real-time spoken word recognition. *Journal of Speech,*

*Language, and Hearing Research*, 56(4), 1328-1345.

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time:

Evidence for autonomous lexical processing. *Quarterly Journal of Experimental*

*Psychology*, 33A(4), 465–495.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Freund, M., & Nozari, N. (2018). Is adaptive control in language production mediated by learning? *Cognition*, *176*, 107–130.
- Fricke, M., Baese-Berk, M. M., & Goldrick, M. (2016). Dimensions of similarity in the mental lexicon. *Language, Cognition and Neuroscience*, *31*(5), 639-645.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower, (Ed.), *The psychology of learning and motivation*. Volume 9. New York: Academic Press.
- Geller, J., Winn, M. B., Mahr, T., & Mirman, D. (2020). GazeR: A package for processing gaze position and pupil size data. *Behavior Research Methods*, *52*(5), 2232–2255.
- Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. <https://doi.org/10.1037/0096-3445.121.4.480>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews. Neuroscience*, *8*(5), 393–402. <https://doi.org/10.1038/nrn2113>
- Hussey, E. K. (2013). *The malleability of cognitive control and its effects on language skills*. University of Maryland, College Park.
- Hussey, E. K., & Novick, J. M. (2012). The benefits of executive control training and the implications for language processing. *Frontiers in psychology*, *3*, 158.
- Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464-482.
- Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological science*, *27*(4), 572-582.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Jacobs, C.L., Loucks, T.M., Watson, D.G., & Dell, G.S. (2020). Masking auditory feedback does not eliminate repetition reduction. *Language, cognition, and neuroscience*, 35(4), 485-497.
- James, W. (1890). *Principles of psychology* (Vol. 1). New York: Holt
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237.
- Khanna, M. M., & Boland, J. E. (2010). Children's use of language context in lexical ambiguity resolution. *Quarterly Journal of Experimental Psychology*, 63(1), 160-193.
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension?. *Language, cognition and neuroscience*, 31(1), 32-59.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, 82, 1-26.
- Laurinavichyute, A., Ziubanov, A., & Lopukhina, A. (2024). Eye-movement suppression in the visual world paradigm. *Open Mind*, 8, 1012-1036.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences*, 9(2), 75-82.
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current directions in psychological science*, 19(3), 143-148.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Levelt, W. J. M. (1989) *Speaking: From intention to articulation*. MIT Press.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22(1), 1-38.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Liu, J. C., Li, K. A., Yeh, S. L., & Chien, S. Y. (2022). Assessing Perceptual Load and Cognitive Load by Fixation-Related Information of Eye Movements. *Sensors (Basel, Switzerland)*, 22(3), 1187.
- Luce, R. D. (1959). *Individual choice behavior* (Vol. 4). New York: Wiley.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, 19(1), 1-36.
- Magnuson, J. S. (2019). Fixations in the visual world paradigm: Where, when, why?. *Journal of Cultural Cognitive Science*, 3(2), 113-139.
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 503.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Mattys, S. L., Barden, K., & Samuel, A. G. (2014). Extrinsic cognitive load impairs low-level speech perception. *Psychonomic bulletin & review*, 21, 748-754.
- Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of experimental psychology. General*, 134(4), 477-500.
- Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of memory and Language*, 65(2), 145-160.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive psychology*, 18(1), 1-86.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- McClelland, J.L. & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, 88(5), 375.
- McMurray, B. (2023). I'm not sure that curve means what you think it means: Toward a [more] realistic understanding of the role of eye-movement generation in the Visual World Paradigm. *Psychonomic bulletin & review*, 30(1), 102-146.
- McMurray, B., Farris-Trimble, A., & Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, 169, 147-164.
- McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive psychology*, 60(1), 1-39.
- 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.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24(1), 167-202.
- Mirman, D., McClelland, J. L., Holt, L. L., & Magnuson, J. S. (2008). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science*, 32(2), 398-417.
- Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In G. A. Watson (Ed.), *Lecture Notes in Mathematics 630: Numerical Analysis* (pp. 105–116). Springer-Verlag.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes factors for common designs* (R package version 0.9.12-4.2). <https://CRAN.R-project.org/package=BayesFactor>
- Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *21*(5), 578-594.
- Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*(3), 214–255.
- Ness, T., Langlois, V. J., Kim, A. E., & Novick, J. M. (2023). The State of Cognitive Control in Language Processing. *Perspectives on Psychological Science*, *20*(2), 219-240.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 263-281.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, *4*(10), 906-924.
- Nozari, N., & Novick, J. (2017). Monitoring and control in language production. *Current Directions in Psychological Science*, *26*(5), 403-410.
- Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & cognition*, *38*(8), 1147-1160.
- Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.
- Peirce, J., MacAskill, M., & Hirst, R. (2022). Building experiments in PsychoPy.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and brain sciences*, 36(4), 329-347.
- Pinet, S., & Nozari, N. (2018). “Twisting fingers”: The case for interactivity in typed language production. *Psychonomic Bulletin & Review*, 25, 1449-1457.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Erlbaum.
- R (Version 4.2.3; R Core Team, 2024)
- Rigler, H., Farris-Trimble, A., Greiner, L., Walker, J., Tomblin, J. B., & McMurray, B. (2015). The slow developmental time course of real-time spoken word recognition. *Developmental psychology*, 51(12), 1690.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42(1-3), 107-142.
- Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and language*, 54(2), 228-264.
- Shrem, Y., Goldrick, M., & Keshet, J. (2019). Dr. VOT: Measuring positive and negative voice onset time in the wild. *arXiv preprint arXiv:1910.13255*.
- Steinberg Lowe, M., & Buchwald, A. (2023). Role of cognitive control in resolving two types of conflict during spoken word production. *Language, cognition and neuroscience*, 38(8), 1082-1097.
- Strauss, T. J., Harris, H. D., & Magnuson, J. S. (2007). jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods*, 39(1), 19-30.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

- Strayer, D. L., & Drews, F. A. (2007). Multi-tasking in the automobile. *Attention: From theory to practice*, 121-133.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2), 257-285.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of experimental psychology: general*, 119(2), 176.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18(6), 645–659.
- 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.
- Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences*, 94(26), 14792-14797.
- Tooley, K. M., Konopka, A. E., & Watson, D. G. (2018). Assessing priming for prosodic representations: Speaking rate, intonational phrase boundaries, and pitch accenting. *Memory & cognition*, 46(4), 625-641.
- Walter, K., & Bex, P. (2021). Cognitive load influences oculomotor behavior in natural scenes. *Scientific reports*, 11(1), 12405. <https://doi.org/10.1038/s41598-021-91845-5>
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167(3917), 392-393.

## LOAD EFFECTS ON LEXICAL ACTIVATION AND COMPETITION

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of memory and language*, 33(3), 332-356.

Zhang, X., & Samuel, A. G. (2015). The activation of embedded words in spoken word recognition. *Journal of memory and language*, 79, 53-75.

Zhang, X., & Samuel, A. G. (2018). Is speech recognition automatic? Lexical competition, but not initial lexical access, requires cognitive resources. *Journal of Memory and Language*, 100, 32-50.