LISTENING IN NOISE AND DIVIDED ATTENTION

COMBINING LISTENING IN NOISE AND DIVIDED ATTENTION WITH PUPILLARY RESPONSE TO EXPLORE ATTENTIONAL RESOURCE USE

By STEFANIA CERISANO, B.Sc., M.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2022)

Hamilton, Ontario (Psychology, Neuroscience & Behaviour)

TITLE: Combining Listening in Noise and Divided Attention with Pupillary Response to Explore Attentional Resource Use

AUTHOR: Stefania Cerisano, B.Sc., M.Sc. (McMaster University)

SUPERVISORS: Dr. Bruce Milliken, Dr. Judith Shedden, Dr. Jennifer Campos

NUMBER OF PAGES: xvii, 195

LAY ABSTRACT

Listening to speech in a noisy environment is a cognitively difficult and effortful task. Attending to more than one task at a time is similarly demanding and effortful. These two kinds of tasks are assumed to use the same limited pool of cognitive resources that we have available to us. This thesis combines listening in noise with divided attention tasks to demonstrate this overlap in demands for cognitive resources using novel combinations of these kinds of tasks. Additionally, this thesis uses the pupillary response—a well-studied index of cognitive effort—to further examine the nature of these overlapping task demands. These studies found that the demands of these tasks do, in fact, overlap, and contribute evidence to the current literature supporting the underlying assumption that these two tasks, and the pupillary response as a measure of effort, are accessing the same pool of limited resources.

ABSTRACT

The concept of attention is complex and multifaceted and can be approached from many perspectives. One such perspective is of attention as a limited pool of resources. Kahneman's (1973) model of limited capacity provides a basis for understanding constraints on attention, including the costs of divided attention. In the same vein as Kahneman's model, the Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016) applies the concept of limited attentional capacity to the demands of listening in a variety of contexts. The current work examines novel combinations of the methods commonly used in the field of Cognitive Hearing Science to address questions about the nature of attention allocation when listening in noise and under the constraints of divided attention. I first combined listening in noise with a secondary continuous working memory task and measured pupillary response as an index of cognitive work and listening effort. Here, I found that listening task demands affect performance on the working memory task. The shared demands of listening and working memory were not, however, evident in the pupil dilation patterns. As a result, I followed these findings by employing a different divided attention method. With the use of a temporally discrete secondary task that either closely overlapped with the listening task or did not closely overlap, I found the same carryover effects of listening demands on secondary task. Most importantly, I found that these demands interacted and were clearly present in the pupil dilation patterns, demonstrating the importance of the timing of the task demands. Together, the studies in this thesis provide evidence that these two secondary tasks access the same attentional resources as those accessed in the primary listening task and that this overlapping demand for resources can be seen in the pupillary response.

ACKNOWLDGEMENTS

Thank you to Bruce Milliken for tirelessly working with me through the past couple of years. You helped me create something I'm proud of and I really couldn't have done it without you. I truly appreciate your enthusiasm for my research and your continuous encouragement. Thank you to my committee members, Judy Shedden and Jenny Campos. Your feedback and guidance were invaluable, but your endless support and kindness meant more to me than anything else. Thank you to Karin Humphreys and David Shore for your continued support when things got tough; it was good to know you were always in my corner.

Thank you to my parents, Lorie and Frank, and the rest of my family, for being cool with me being in school for forever and supporting me through all these years. Thank you to my CogSci family, Esther Manoian, Kaian Unwalla, Julie Bannon, Melissa Ptok, Julie Conder, Lila Danis, Marija Radenovic, and Hannah Song. You all came on this journey with me in different ways and I owe so much of the researcher I've become to your brilliance. Thank you to Blake Anderson, Alyssa Veltri, Stephanie Hatzifilalithis, Sarah Creighton, and Sitara Sharma for all your incredible support. Thank you to all my friends and colleagues for the good company over the years.

And finally, thank you to all the neighbourhood cats for putting up with me on my procrastination walks. I'm sorry if I scared you. I just wanted to pet you.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

LIST OF TABLES

CHAPTER 2

CHAPTER 3

DECLARATION OF ACADEMIC ACHIEVEMENT

All the experiments in this thesis were conceptualized and designed by Stefania Cerisano, Dr. Scott Watter, Dr. Karin Humphreys, and Dr. Jeff Crukley. The data were collected by members of the Cognitive Science Laboratory at McMaster University, including Stefania Cerisano, Esther Manoian, Marija Radenovic, and Hannah Song.

The General Introduction Chapter and General Discussion Chapter of this thesis were conceptualized and written by Stefania Cerisano with main edits from Dr. Bruce Milliken and additional edits provided by the supervisory committee, Dr. Judith Shedden and Dr. Jennifer Campos. Chapter 2 and Chapter 3 of this thesis were written by Stefania Cerisano with main edits from Dr. Bruce Milliken and additional edits provided by the supervisory committee, Dr. Judith Shedden and Dr. Jennifer Campos.

CHAPTER 1: GENERAL INTRODUCTION

The concept of attention is a ubiquitous thing, and everyone has some level of understanding of what it is. Its ubiquity makes it easy to talk about in a general sense but very hard to pin down from a theoretical, science-driven perspective. We may commonly hear parents and teachers telling kids to 'pay attention', but most who use this phrase do not consider the complex balance of costs and benefits associated with 'paying attention'. We 'attend' to tasks all the time, but do not necessarily think about the mechanisms that allow us to do so; we do not consider what it *means* to attend to something. Although we *know* that attention plays a pivotal role in our everyday lives, there are also countless areas of research that are dedicated to understanding *how* attention contributes to our lives. Some of this research focuses on applications of attention to everyday tasks, other research focuses on clinical studies of when attention goes wrong, and still more research digs into the theoretical and mechanistic ways in which attention works. The ubiquity of attention means it can be studied in a variety of ways and this allows us to understand how it works from different perspectives.

A benefit of attention being so familiar a construct is that intuition points to many ways in which it can be studied scientifically. The diversity of uses of attention has led to a diverse array of research areas that focus on the construct of attention. From a clinical perspective, attention can be studied to understand ways in which it differs between people and how it might not function correctly. For example, some clinical research has focused on Attention Deficit/Hyperactivity Disorder and autism, providing insight to how attention is different in neuroatypical individuals and how understanding this difference can point to interventions and medications that will improve their daily lives (Hagen & Hale, 1973; Klimkeit et al., 2005; Naber et al., 2008). From a more cultural and sociological clinical perspective, attention is also relevant to music therapy, which can be used in a variety of therapeutic settings and in conjunction with other treatments (Morton et al., 1990). Another therapeutic application aimed at attention is the practice of mindfulness. Mindfulness is the act of focusing on a single object of attention, be it a part of the body or breathing itself, and letting go of any other thoughts. This practice is focused on how attention can be controlled and directed to reduce stress and practice inward reflection (Garland et al., 2015; Jha et al., 2007; Semple, 2010; Tang et al., 2015; Valentine, & Sweet, 1999).

Another application of attention research aims to understand how it contributes to common ways in which we interact with the world around us. For example, a lot of attention research is dedicated to the task of driving. Remarkably, we seem able to attend to a variety of demands, including holding conversations, while operating a complex and potentially dangerous moving object (Alletto et al., 2016). Another currently relevant applied research topic is the impact of ever-increasing screen time on cognitive function, as well as how attention works in virtual situations (Schmidt & Vandewater, 2008). And, of course, a natural way to study attention is to look at the sensory organs that control attention, such as the eyes. We can learn much about attention by studying eye movements and how and where the gaze is directed (Baldi & Itti, 2010).

Beyond the clinical and applied contexts of attention in research, there are studies of how attention can be understood and utilized to improve communication in workplace function and student learning. In the workplace, there is an entire domain of research called organization science, which looks at how companies are organized and communicate. This area of study uses theories of attention to improve how the employees in a workplace can function cohesively (Ocasio, 2011). In addition to understanding how screen time affects attention, there is also research on how video communications are attended to and used in both learning and work environments (Kuzminykh, & Rintel, 2020). The recent Covid-19 pandemic has made understanding how attention works in various learning environments particularly important; the shift to online learning and online communications has made it imperative to optimize attention in contexts that require continuous use of screens and other virtual platforms.

Research in this vast array of topics helps us to understand attention in individuals, how attention affects everyday interactions with the world around us, and how we can optimize and utilize attention in everyday tasks. However, attention has also been the subject of more targeted basic theoretical research. This basic research focuses on identification of fundamental mechanisms of attention, how those mechanisms of attention work, and what are the best research methods to study those mechanisms. Whereas studying applied topics in attention focuses on behaviours that we practice routinely in real life, studying fundamental mechanisms of attention aims at a deeper understanding of how attention functions. One such understanding of attention is that it is limited and can be thought of as a pool of resources that we dip into when we attend to a variety of activities. As we are taking from this pool of resources, we are limiting what is now available until we have completed the activities we are attending to. Most relevant to the current thesis is basic research that aims to understand how attention mechanisms prevent these limited capacity mental processes from being overwhelmed by vast amounts of available information (Dayan et al., 2000; Driver, 2001; Johnston & Dark, 1986; Moray, 1967).

Foundational Ideas of Attention

One historically important line of research into attention conceptualizes it as a filter. The concept of attention working as a filter for incoming information builds off the idea of attention as a selective process (for reviews, see Driver, 2001; Hatfield, 1998). We are unable to process and assign meaning to *all* the sensory information in our environment so there must be a way to select what information is attended to and what is not. It is again important to note that selective attention implies there is a limit to our attentional abilities; we must select some information to attend to because we cannot attend to all of it.

Broadbent (1958) proposed a filter model of attention, where incoming information was selected for based on relatively simple perceptual characteristics processed "early" in information processing streams. The model consisted of a short-term storage, a filter, and then a "channel" into which information passes. Information from the environment enters the short-term storage, and then is selected to pass through the filter and into the "channel" in a serial fashion. In other words, new information cannot be processed in the "channel" until the current information is dealt with. In this sense, there is a limit associated with selective processing of information that passes through the filter. This limit imposed by early selective filters is what gives the sense that we attend to some information in our environment while seemingly ignoring other information, and that we can often attend to just one stimulus or piece of information at a time.

Broadbent's (1958) filter theory served as a foundation for additional theoretical work by his student, Anne Treisman. Treisman's (1964) attenuation model of attention allowed for the idea that whether information is filtered at any early stage of processing may depend on its importance. Whereas some information may be selected early based on physical attributes, such as a tone of voice or the loudness of a stimulus, other information may be selected later based on more semantic aspects of the signal and the importance of the information itself (Treisman, 1960; Treisman, 1964; Treisman & Geffen, 1967; Treisman & Gelade, 1980; Treisman &

Gormican, 1988; Sullivan, 1976). Together, the filter theories of Broadbent and Treisman formed a strong foundation for later theorizing on selective attention. Indeed, Kahneman (1973) integrated these foundational ideas into his work by emphasizing that central processing mechanisms in all filter models are limited in capacity. From this standpoint, incoming information may occur in parallel before being filtered, but once the information has passed a filter, processing must occur serially. Though Kahneman's model of limited capacity attention will be described in more detail in the next section of this Introduction, it is worth noting here that the 'selective' part of selective attention implies allocating limited processing capacity by prioritizing some sources of information over other sources of information.

Another influential view of selective attention considers attention to function like a spotlight. Spotlight models, in contrast to filter models, conceptualize the selection process to be akin to a moving spotlight that seeks out and chooses information for further processing. This concept of selective attention explains how information within the focus of attention benefits from more efficient and more complete processing than information outside of the focus or 'spotlight' (Cave and Bichot, 1999; Hurlbert & Poggio, 1985; Posner et al., 1980). Although the spotlight metaphor differs mechanistically from the filter metaphor, it retains the core idea that attention is selective and that we cannot attend to everything all at once but must focus instead on only some of the information in our environment at a time.

These foundational ways of thinking of attention both highlight the concept of attention as limited in some way and emphasize that we must be selective in what we attend to. In line with the idea that there is a limit to attention and how it functions, another foundational way to think of attention is as a pool of limited resources. The following sections will describe Kahneman's (1973) model of limited attentional capacity and set the context for thinking about attention as a limited cognitive resource that is tapped in accordance with the demands of present tasks and stimuli.

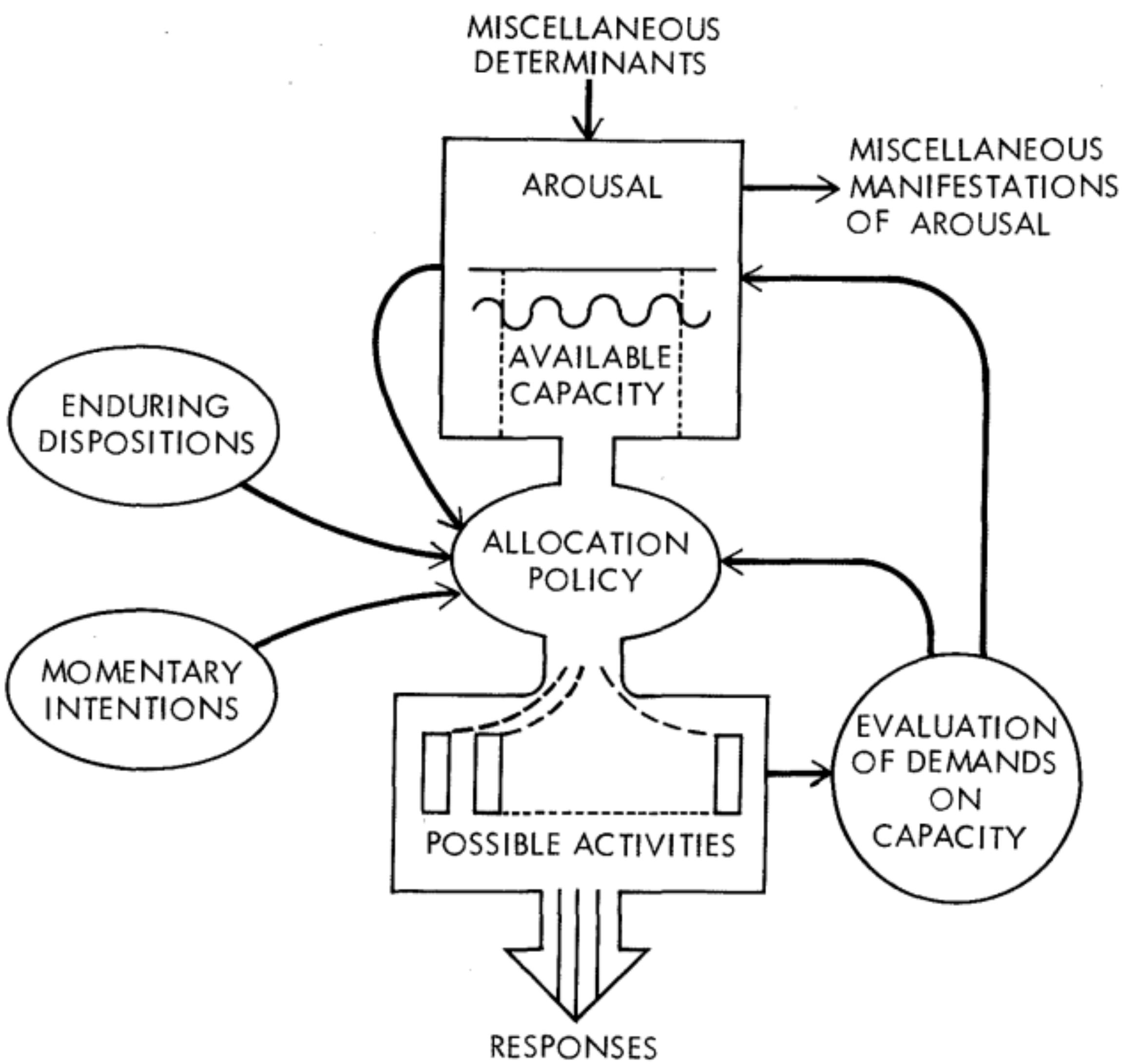
Attention as a Resource

An alternative to the idea that attention acts as a filter is that attention consists of a limited pool of general cognitive resources. Like the filter models of attention, the limited resource model of attention implies that capacity limitations require the prioritization of competing sources of information. However, resource theories incorporate a plethora of factors that influence the attentional resources available at any given point in time, as well as how these resources are parceled out and allocated to competing sources of information. The following section describes a seminal resource model of limited capacity attention (Kahneman, 1973).

Kahneman's model of attention

Kahneman's model of attention builds on foundational ideas about selective attention but adds the idea that attention has limited availability that varies with use (Kahneman, 1973). This limited capacity notion of attentional resources implies that the attention available at any point in time mediates what information is processed and how much effort is applied to processing that information (Kahneman, 1973). Figure 1 shows Kahneman's capacity model of attention. The model addresses three related questions about limited central capacity attention: (1) how do the demands of an activity relate to the allocation of attention?; (2) what influences the amount of attention that is available at any given point in time?; and (3) what factors determine when and how many resources are allocated? Together, answers to these three questions offered by Kahneman's model explain how attention and effort can be viewed as a limited capacity resource essential to effective interaction with the environment (Kahneman, 1973).

Figure 1: Kahneman's Model of Limited Attentional Capacity (reproduced with permission; Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.; https://researchonline.ljmu.ac.uk/id/eprint/3572/1/Pichora-Fuller_2016_EarHearing.pdf)



The model is centered on the idea that there is a wide set of "possible activities" that require attention and effort at any given point in time (Kahneman, 1973). Essentially, anything that requires attention to be processed is considered a "possible activity" and the likelihood of successfully completing these activities depends on sharing the resources and effort that are available for allocation. The allocation process is subject to supply and demand, where limited capacity attentional resources command an amount of attention in accordance with their importance. Available capacity is closely related to arousal. Changing states of arousal are associated with changes in available attentional resources. Much in the way that a fatigued person may have fewer resources to allocate attention to demanding tasks, a person who is more energized would have more resources available for such tasks. The state of arousal of an individual can itself be affected by task demands, but also by other factors such as physical energy levels or emotional states (Kahneman, 1973). The arousal of the individual can be measured in a variety of ways, including heart rate and pupil dilation. Levels of arousal and limited capacity then feed into the "allocation policy" that determines how much effort and to what "possible activities" attention will be allocated. The "allocation policy" is simultaneously informed by an evaluation of demands on the available capacity by the "possible activities". This evaluation allows for the prioritization of activities and for the provision of needed attentional resources. The "allocation policy" is additionally affected by external factors such as activities that demand involuntary attention, the intention of the observer in the moment, and how arousal states affect available attentional capacity. All together, these factors determine what activities are attended to and how much effort is devoted to each of these activities.

This model of attention has a lot of moving pieces—multiple internal and external factors affect capacity availability and capacity allocation. Though this complexity may be cast as a weakness of the model, attention is a complex concept that requires multiple factors to be taken into account. Careful consideration of each of these components can help determine which questions about attention need to be asked and how to ask them. Kahneman's limited capacity model allows us to speculate on how internal factors affect arousal and cognitive capacity, which in turn affects how capacity is allocated. The model allows us to hypothesize how motivation to prioritize one task over another might affect how capacity is allocated to these tasks. The model invites us to conceive experiments that measure how changes in arousal and task demands affect how attentional resources are allocated. The most important property of Kahneman's model of limited attentional capacity is the interconnectedness of its components. This interconnectedness makes the model more complex than other structural models of attention and information processing, but it also allows it to capture a rich array of behaviours observed in real-world attention.

Attention as a limited resource and working memory capacity

Beyond Kahneman's (1973) foundational model, a closely related area of research that considers attention to be a resource is that of working memory. Working memory is viewed as a multicomponent system for temporary storage, updating, and manipulation of information (Baddeley & Hitch, 1974). By this view, working memory consists of subsystems for temporarily storing information, and a central executive that controls updating and manipulation of the content of these subsystems. Importantly, the central executive is widely viewed as limited by attentional capacity (Baddeley, 2002; Heitz et al., 2005; Kane et al., 2001; Engle, 2018; Logie, 2011; Posner et al., 1980; Shipstead, Lindsey et al., 2014). The limited capacity central executive is crucial for information processing that is widely viewed as attentional in nature (e.g., selective attention, divided attention, task set control). From this perspective, attention can be seen as critical for the control of working memory (Cowan et al., 2005). The role of attention is an important feature of working memory theories. Indeed, a recent review of a wide range of working memory theories proposed that the role of attention is a fundamental feature that can be used to distinguish among various conceptualizations of working memory (Adams, Nguyen & Cowan, 2018). The limited capacity of attention shapes working memory; both what and how much is in working memory depends on how attention is allocated. Attention is an integral part of working memory and understanding attentional limitations, including the factors and systems in Kahenman's model, can further inform our understanding of working memory and the role it plays in individual cognitive abilities and the applications of these abilities to attentional tasks.[1]

Relating these ideas of working memory to Kahneman's model of limited resource capacity, we can see that all the components of Kahneman's model that feed into the "allocation policy" and influence attention are also likely to influence working memory. If the central executive in working memory is governed by attention, and attention itself is affected by all the factors Kahneman describes, then working memory and the ability to maintain and update information in the mind is also affected by these factors. By this logic, any task that requires working memory is also going to be affected by attention and all the ways attention allocation can be manipulated.

Measuring resource capacities

How does one measure a construct such as limited capacity attentional resources? One approach is the attention span tasks widely used in clinical and applied settings (Engle, 2002; Hill et al., 2010). These span tasks (digit span, reading span, operation span) measure an individual's memory span by loading working memory with sets of information that must be

[1]The limitations of attentional capacity and resources can also be conceived of in a more structural manner. Franconeri, Alvarez, and Cavanagh (2013) considered the concept of capacity-limited cognitive resources in relation to cortical mappings of visual attention and memory. These authors relate a limit of mental resources to competition for limited access to neurological space. A physiological basis of limited cognitive capacity lends support to the theoretical understandings of limited attentional resources.

remembered while also performing another task that demands attentional resources. These tests offer relatively simple measures of the limits of attention that map easily onto everyday situations and to clinical settings. Another similar and common way to measure limited capacity attention is with working memory tasks. For example, requiring participants to complete a task while also retaining a concurrent working memory load allows researchers to make inferences about attentional limits to performance imposed by the concurrent working memory load (de Fokert et al., 2001; Logan, 1979; Park, Kim & Chun, 2007). Loading working memory is assumed to constrain available attentional resources, and therefore performance on a simultaneous task is thought to reflect what remaining resources are available. This method highlights the idea that the central executive in working memory is reliant on attention, and that taxing attentional resources will compromise performance on any task with a working memory component.

As noted by Kahneman (1973), measures of physiological arousal can be used to understand attentional capacity use. In his model, Kahneman notes that internal states of arousal and available capacity are closely related, and in turn physiological measures of arousal can be used to understand how capacity changes with changing resource availability and use. More specifically, his model conceptualizes arousal in two ways: the first is how a state of arousal affects the available attentional capacity; and the second is how physiological arousal can be used as a way to measure the use of attentional capacity when applied to task demands. Measures such as eye response, cardiac changes, and skin conductance, therefore, can be used to infer changes in attention use and the demands on the available capacity. Saccadic eye movements, while not necessarily a measure of physiological arousal, can be used to understand where attention resources are being allocated (Groner & Groner, 1989). Another eye-related measure of physiological arousal is the pupillary response. The pupil dilates in response to a whole host of factors, the most pertinent of which is cognitive work and effort (Hoeks & Levelt, 1993; van der Wel & van Steenbergen, 2018). When the pupillary response is time-locked to a stimulus or event, it is called a task-evoked response (Kang et al., 2014; Konishi et al., 2017; Smallwood et al., 2011). The pupil response is used to study attention demands in a wide variety of tasks and is thought to provide a physiological measure of the use of limited attentional resources. Again, Kahneman's limited capacity model states that arousal is directly related to the available attentional capacity; physiological measures of arousal, in turn, can be used to make inferences about available capacity and capacity use.

Divided Attention

In the same book in which he describes the capacity model of attention, Kahneman (1973) describes the interference that occurs when engaging simultaneously in more than one activity that requires attention. Ninio and Kahneman (1974) point to a pattern of increased errors and reaction times in such divided attention tasks relative to a single task control. These divided attention costs follow from having to divide the limited available attentional resources between two competing "possible activities". Divided attention methods, therefore, provide a tool to study limited attentional resources—by taxing resources with one task and measuring the costs incurred in task performance on the other task. Dividing attention, or time sharing resources between tasks, comes with costs to task performance (Wickens, 1976).

Applications of divided attention

Knowing that divided attention has measurable outcomes that allow us to infer the allocation of attentional resources gives us many opportunities to apply these concepts to everyday life. We often feel like we are performing more than one task at a time or thinking more than one thought at a time. Colloquially, this is referred to as “multitasking”. Though multitasking may often feel like we are doing two things at once, research suggests that critical components of the two tasks may actually be done in succession (see Dzubak, 2008 for a review of multitasking research; Himi et al., 2019). Generally, people’s metacognition of their ability to multitask does not predict how well they actually perform on divided attention tasks (Finley et al., 2014). We may think we can multitask well, but in reality there are easily measured costs of dividing attention between more than one activity. As noted, the term “multitasking” itself is misleading, as we do not actually do multiple tasks simultaneously at full capacity, but instead share cognitive resources between tasks over a short period of time. This capacity sharing between tasks has implications related to tasks such as driving, and particularly how driving ability may be affected by aging.

While driving, many of us engage in divided attention by simultaneously holding a conversation. This division of attention is often framed in the context of resource models of cognition. Sharing resources between driving and conversing can negatively affect both tasks, but of course the effects on driving are the most concerning. In fact, the more complex a conversation is, the greater the resources it demands, and the greater detriment there is to driving performance (Iqbal et al., 2010). There is also evidence that performance in divided attention tasks decreases with age (Brouwer et al., 1991). Again, this relationship is often looked at in the context of driving, though it is unclear whether it applies generally to all divided attention contexts (McDowd & Craik, 1988; Ponds et al., 1988; Somberg & Salthouse, 1982; Svetina, 2016). In summary, the construct of limited capacity resources plays a helpful role in understanding how divided attention might affect an everyday task like driving, and how an individual difference variable such as aging and divided attention together may have important consequences for behaviour.

In addition to applied research, much basic research has been devoted to the study of divided attention in relation to working memory. When attentional resources are divided between two tasks, this challenge to limited attentional resources is mediated by the central executive component of working memory, which maintains and updates information in working memory. People with different working memory capacities are affected differently by divided attention demands. Though the most intuitive prediction would be that those with higher working memory capacity should perform better than people with lower working memory capacity, a range of outcomes is possible. The precise outcome can vary with task context, implicating the role of a highly flexible working memory central executive in divided attention contexts (Colflesh & Conway, 2007; Kane & Engle, 2000). For example, those with greater working memory capacity are sometimes affected particularly strongly by having to divide their attention, presumably because they have the attentional capacity available to process information from more than one task. Findings have been reported in both positive and negative directions, such that people with higher working memory capacity are better at divided attention tasks, but also that they may be more susceptible to interference from secondary tasks, further highlighting the role of attention in working memory. Attentional processing required when dividing attention between tasks determines how people with different working memory capacities are able to deal with competing processing demands.

Shared capacity or a structural bottleneck?

Although there is wide agreement that attentional capacity is limited, there has been debate over the nature of this limit for decades. Two models of how exactly attention limits affect dual-task performance have been studied extensively: a shared capacity model and a structural bottleneck model. Both of these models find support in the literature. The work presented in this thesis does not aim to make any definitive statement about the relative merits of these two views, and actually borrows from the research traditions of both views. A brief overview of these two views follows.

The bottleneck theory of limited cognitive capacity assumes that there is an information processing mechanism that is capable of attending to only one task at a time. This mechanism must process task-relevant information from a first task before it can move on to process task-relevant information from a second task. This structural bottleneck can be thought of as a stopping point, where the first task must be completed before the next task can proceed. Pashler (1989, 1994a, 1994b) describes how this structural bottleneck produces a Psychological Refractory Period (PRP). The PRP occurs when two tasks are being performed within a short timeframe; when the second task is presented in quick succession after the first, the second task takes extra time to complete. In contrast, when the second task is presented long after the first task, the second task can be completed without delay. This finding has been replicated in many studies and implicates some form of cognitive capacity limitation that slows performance for a second task when two tasks are presented in quick succession. According to Pashler, the slowdown is caused by the main processing mechanism—the stage in stimulus processing that does the actual cognitive work—being able to attend to only one task at a time. This bottleneck leads to divided attention costs; secondary tasks take longer to complete and suffer in performance accuracy when they demand resources at the same time—or close to the same time—as primary tasks. This bottleneck in processing has been studied in a variety of contexts by Pashler and his colleagues (Ruthruff & Pashler, 2001; Ruthruff, Pashler & Klaassen, 2001).

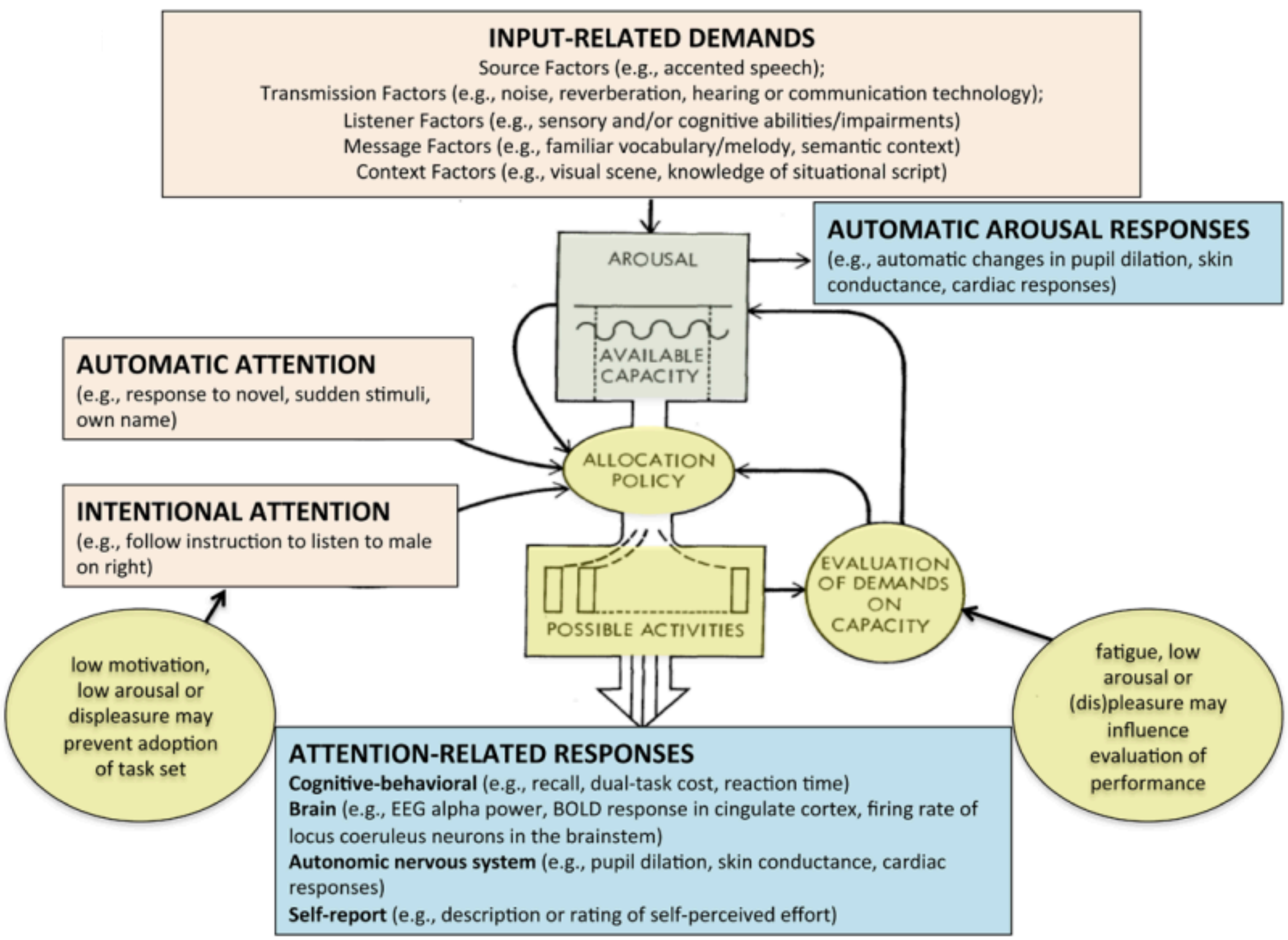
An alternative to the structural bottleneck theory instead proposes that the costs seen in divided attention are due to a sharing of limited available capacity. Instead of a central mechanism having to process two tasks in sequence, this theory states that available capacity is shared between tasks or items that need to be processed. Tombu and Jolicœur (2002, 2003) argue that a central capacity sharing model, in which the processing of two tasks occurs in parallel, can explain the PRP effect as well as results seen in other dual-task experiments. In summary, in contrast to a bottleneck, this model proposes that when two tasks occur temporally close to one another, the limited available cognitive resources are divided between the tasks, with the proportions of resources allocated to each being determined by the demands of the tasks themselves.

In proposing his limited capacity model of attention, Kahneman (1973) also addresses divided attention and the perspective that doing more than one thing at a time causes interference in the demand for available resources. Kahneman's model aligns with the central capacity sharing model of divided attention, where the interference between two tasks is dependent on the demands each "possible activity" places on the available attention capacity. By this view, divided attention costs are due to a sharing of the available resources that is determined by the allocation policy's evaluation of task demands and possible responses.

A Relevant Application of Attention as a Resource

Most relevant to the current research is a framework based on Kahneman's model of attention allocation, called the Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016). In this framework, the allocation of attention and effort is put into the context of listening to speech and other auditory signals. The goal of the FUEL is to provide a basis for understanding how effort and attention are allocated with regards to Cognitive Hearing Science (CHS). CHS is a field of research that combines the study of cognition with that of clinical approaches to hearing, hearing loss, and age-related changes in hearing. By expanding upon all the factors that can affect the different components of Kahneman's limited capacity model of attention, the FUEL addresses questions within the area of CHS. In this sense, the basis of Kahneman's model is updated with more contemporary understandings of attention and its correlates and considers attention from the perspective of listening and the factors associated with listening. Figure 2 shows the FUEL, which is based on Kahneman's (1973) model.

Figure 2: Pichora-Fuller et al.'s (2016) Framework for Understanding Effortful Listening (FUEL; reproduced with permission; Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.; https://researchonline.ljmu.ac.uk/id/eprint/3572/1/Pichora-Fuller_2016_EarHearing.pdf).



As with Kahneman's model, the FUEL is comprised of the same components of arousal, available capacity, the "allocation policy", disposition and intentions, "possible activities", and the evaluation of activity demands on the available capacity. This framework expands on each of these components in some ways that are specific to listening and hearing science and in other ways that are applicable to any area of attention research. If we start with the "possible activities" of Kahneman's model, the FUEL proposes that there are attention-related responses that can be broken down into a number of categories: 1) cognitive behavioural, which includes classic cognition measures such as memory and dual-task performance costs; 2) brain, which means using EEG or other methods to measure changes in brain activity; 3) autonomic nervous system, which includes the body's automatic physiological responses to attention allocation, like pupil dilation and cardiac changes; and 4) self-report, which measures a person's self-perception on how they use their attention and effort.

These "possible activities", and the responses that correspond with them, are evaluated to determine how much attentional capacity they demand, as also seen in Kahneman's model. The FUEL also identifies transient factors that may affect the evaluation of the demands placed on the allocation policy by the possible responses. The FUEL acknowledges that there are additional factors, such as fatigue and other predispositions that can affect how demands get evaluated, which in turn will affect how resources are allocated. For example, a person who has just completed a long shift at work may have fewer resources available than a person who is rested and alert. As a result, when evaluating the demands that possible activities and responses put on their available capacity, the tired individual might evaluate certain activities with a particularly high priority. For example, they might put more effort into listening to the traffic report on the radio so they can get home to rest than to listening to a conversation with their passenger.

In addition to the evaluation of demands feeding into the "allocation policy", Kahneman's model also has enduring dispositions and momentary intentions that affect the attention "allocation policy." In the FUEL, these factors are instead referred to as automatic attention and intentional attention, which reflect involuntary (e.g., hearing your name being called) and voluntary (e.g., intending to follow certain instructions) aspects of attention. These factors reflect task-related effects on attention allocation. Again, these components are expanded versions of those that Kahneman included in his original model; the goal of these changes is to explain how aspects of attention we can observe in both everyday life and the laboratory factor into how limited attention is used. To continue the example from above, if our tired person is not motivated to listen to their driving companion, that affects how they intend to use their attention and therefore affects the allocation of available resources.

The relationship between arousal and available capacity, and how those feed into the "allocation policy", also remains from Kahneman's model. Regarding these components, the FUEL also expands on what inputs affect the component of arousal and available capacity, and how this arousal can be measured. The input-related demands are very specific to listening and include information about the physical aspects of the incoming auditory signal (e.g., signal to noise ratio, hearing impairments, the use of hearing aids, whether the speech is accented or has reverberation) and also cognitive factors associated with the signal (e.g., semantic context, familiarity with the vocabulary). These factors are the inputs to the model and affect arousal and available capacity, which in turn feed into the "allocation policy" that determines how that limited available capacity will be used. Lastly, as in the Kahneman model, the FUEL identifies physiological measures that can be used to infer a person's state of arousal. These measures include pupil dilation and cardiac responses, both of which are indirect measures of attention responses.

To summarize, the FUEL is an expansion on Kahneman's (1973) limited capacity model of attention. While the FUEL was created for the purpose of listening and hearing experiments typical of CHS research, most of it does still apply to the study of attention and cognitive effort in general. The input demands, while specific to factors involved in listening, can be thought of in the context of any demand on attentional resources. The FUEL itself is focused on using Kahneman's notion of attention allocation to understand how this allocation works in listening and, more importantly, how listening effort can be understood in terms of attention and cognitive resource use. This framework, as with Kahneman's model, provides many different avenues for possible research that all center on the notion of attention as a limited resource. All in all, the general concept of limited attentional resources and their allocation, as conceptualized by Kahneman, can be applied to more specific areas of research that are inextricably related to attention—listening is one of those areas.

The utility of the FUEL in systematically breaking down the many different components that affect listening effort and attention provides the basis for the current research. The experiments reported in this thesis are aimed at attention limits and use in the context of listening to speech in noise and dual-task demands, and the FUEL offers a useful basis to identify research questions in this domain. With the FUEL as a theoretical foundation—along with all the attention research that came before it—I have conducted research that examines how attention allocation and cognitive effort while listening in noise are affected by a variety of input-related demands, all of which can be measured through a variety of attention-related responses.

Current Research

The goal of the current research is to combine Kahneman's (1973) and the FUEL's (Pichora-Fuller et al., 2016) approach to understanding attention and listening effort with a theoretically driven approach to divided attention. The current studies combine a listening in noise task with various dual-task methods and measures pupil dilation as an index of attention and effort. While using a dual-task paradigm to measure listening costs is a common method in the field of CHS (see Gagné et al., 2017 for a review of dual-task methods combined with listening tasks), the specific methodologies used in the current thesis have not been combined before. The current studies combine listening in noise with a secondary task in two dual-task methods, a continuous working memory task and a PRP-style task, and further include measuring pupillary responses to task demands. This novel combination allows for a greater understanding of how listening in noise, working memory, and divided attention can all be understood from the perspective of limited attentional capacity. It is important to combine these methods to answer theoretically-focused questions about attention in the context of listening in noise. The current research adopts methods that are commonly used in CHS and blends them with two areas of divided attention that have been thoroughly studied in the field of experimental psychology—working memory and the PRP. The aim is to better understand the complex nature of attention allocation through the lens and tools of basic experimental psychology.

The three approaches that are combined in the current research include experimental manipulations of listening in noise and dual-task requirements; and the outcome measure of pupil dilation as an index of cognitive effort. These three methodologies have a common focus on the allocation of limited resources to complete cognitive tasks. The FUEL puts listening, and input-related demands that can affect listening, in the context of an attention allocation framework that builds on Kahneman's (1973) work. Listening to speech in noise and other related tasks requires the allocation of limited capacity attentional resources, and this allocation of resources is thought to be akin to effort. By framing listening in this way, we can address questions of attentional resource allocation in listening tasks while participants complete other tasks known to tap attentional resources.

Dual-task paradigms have been viewed historically from the perspective of limited attentional resources. The term "divided attention" itself implies that doing more than one task simultaneously requires that available attention be divided. Kahneman (1973) proposed that interference occurs when two tasks are completed simultaneously because each of the tasks demands some of the limited available resources. In the studies reported here, I use divided attention paradigms to examine how resource allocation responds to different kinds of divided attention demands while listening to speech in noise. Kahneman (1973) also described how physiological measures, such as pupil dilation, can be measured to understand arousal and, by extension, the available attentional capacity and its allocation. The current research uses pupil dilation to understand cognitive effort and attention allocation in a way that is complimentary to the manipulations to cognitive load through divided attention and listening in noise. The current research, therefore, uses Kahneman's (1973) and Pichora-Fuller et al.'s (2016) conceptualizations of limited attention capacity as a foundation to study three interrelated issues: listening to speech in noise, divided attention, and pupil dilation as a measure of cognitive resource use.

There is broad research interest in the topics of listening in noise, divided attention, and pupil dilation as a measure of attention and effort, both on their own and in combination. For example, recent prior studies have focused on listening under divided attention (Gagné et al., 2017), divided attention and the pupillary response (van der Wel & Steenbergen, 2018), and listening and the pupillary response (Koelewijn et al., 2014). However, to my knowledge no prior study has tackled the challenge of combining all of these elements in the way that I have in this thesis; my studies focus on listening in noise under dual-task conditions, with a focus on pupil dilation as a measure of attentional resource allocation and effort. The current research is also unique in that I use two different dual-task methods: a method that invokes a continuous working memory secondary task, and a method used to study the PRP effect. Neither of these dual-task methods have been studied previously in combination with listening in noise and pupil dilation. In Chapter Two, I use a continuous working memory task to examine the connection between cognitive loads and a variety of listening input demands, with a focus on pupil dilation as a measure of arousal related to cognitive work and attentional resource use. In Chapter Three, I use a PRP-style method to tax attention at the time of listening to examine how increased attentional demands at specific times interact with listening input demands, again with a focus on pupil dilation as a measure of arousal related to cognitive work and attentional resource use. Together, pupil dilation measures in these two studies of listening input demands under divided attention offer a novel lens to the understanding of allocation of attention as a limited capacity resource.

References

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, speech, and hearing services in schools*, *49*(3), 340-355.

Alletto, S., Palazzi, A., Solera, F., Calderara, S., & Cucchiara, R. (2016). Dr (eye) ve: a dataset for attention-based tasks with applications to autonomous and assisted driving. In *Proceedings of the ieee conference on computer vision and pattern recognition workshops* (pp. 54-60).

Baddeley, A. D. (2002). Is working memory still working?. *European psychologist*, *7*(2), 85.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*(5), 649-666.

Broadbent, D. E. (1958). Perception and Communication. New York: Pergamon Press.

Brouwer, W. H., Waterink, W., Van Wolffelaar, P. C., & Rothengatter, T. (1991). Divided attention in experienced young and older drivers: lane tracking and visual analysis in a dynamic driving simulator. *Human factors*, *33*(5), 573-582.

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic bulletin & review*, *6*(2), 204-223.

Colflesh, G. J., & Conway, A. R. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic bulletin & review*, *14*(4), 699-703.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive psychology*, *51*(1), 42-100.

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature neuroscience*, *3*(11), 1218-1223.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*(1), 53-78.

Dzubak, C. M. (2008). Multitasking: The good, the bad, and the unknown. *The Journal of the Association for the Tutoring Profession*, *1*(2), 1-12.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*(2), 190-193.

Finley, J. R., Benjamin, A. S., & McCarley, J. S. (2014). Metacognition of multitasking: How well do we predict the costs of divided attention?. *Journal of experimental psychology: applied*, *20*(2), 158.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*(5509), 1803-1806.

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in cognitive sciences*, *17*(3), 134-141.

Gagné, J. P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in hearing*, *21*, 2331216516687287.

Garland, E. L., Farb, N. A., Goldin, P. R., & Fredrickson, B. L. (2015). The mindfulness-to-meaning theory: extensions, applications, and challenges at the attention–appraisal–emotion interface. *Psychological Inquiry*, *26*(4), 377-387.

Groner, R., & Groner, M. T. (1989). Attention and eye movement control: an overview. *European archives of psychiatry and neurological sciences*, *239*(1), 9-16.

Hagen, J. W., & Hale, G. A. (1973). The development of attention in children.

Hatfield, G. (1998). Attention in early scientific psychology. *Visual attention*, *1*, 3-25.

Heitz, R. P., Unsworth, N., & Engle, R. W. (2005). Working memory capacity, attention control, and fluid intelligence. *Handbook of understanding and measuring intelligence*, 61-77.

Hill, L., Williams, J. H., Aucott, L., Milne, J., Thomson, J., Greig, J., ... & Mon-Williams, M. A. R. K. (2010). Exercising attention within the classroom. *Developmental Medicine & Child Neurology*, *52*(10), 929-934.

Himi, S. A., Bühner, M., Schwaighofer, M., Klapetek, A., & Hilbert, S. (2019). Multitasking behavior and its related constructs: Executive functions, working memory capacity, relational integration, and divided attention. *Cognition*, *189*, 275-298.

Hoeks, B., & Levelt, W. J. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, *25*(1), 16-26.

Hurlbert, A., & Poggio, T. (1985). Spotlight on attention. *Trends in Neurosciences*, *8*, 309-311.

Iqbal, S. T., Ju, Y. C., & Horvitz, E. (2010, April). Cars, calls, and cognition: Investigating driving and divided attention. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1281-1290).

Jha, A. P., Krompinger, J., & Baime, M. J. (2007). Mindfulness training modifies subsystems of attention. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(2), 109-119.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual review of psychology*, *37*(1), 43-75.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of experimental psychology: General*, *130*(2), 169.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336.

Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil dilation dynamics track attention to high-level information. *PloS one*, *9*(8), e102463.

Klimkeit, E. I., Mattingley, J. B., Sheppard, D. M., Lee, P., & Bradshaw, J. L. (2005). Motor preparation, motor execution, attention, and executive functions in attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *11*(2), 153-173.

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hearing research*, *312*, 114-120.

Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention wanders: Pupillometric signatures of fluctuations in external attention. *Cognition*, *168*, 16-26.

Kuzminykh, A., & Rintel, S. (2020, April). Classification of functional attention in video meetings. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(2), 189.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current directions in Psychological science*, *20*(4), 240-245.

McDowd, J. M., & Craik, F. I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of experimental psychology: human perception and performance*, *14*(2), 267.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta psychologica*, *27*, 84-92.

Morton, L. L., Kershner, J. R., & Siegel, L. S. (1990). The potential for therapeutic applications of music on problems related to memory and attention. *Journal of Music Therapy*, *27*(4), 195-208.

Naber, F. B., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Dietz, C., Van Daalen, E., Swinkels, S. H., ... & Van Engeland, H. (2008). Joint attention development in toddlers with autism. *European Child & Adolescent Psychiatry*, *17*(3), 143-152.

Ninio, A., & Kahneman, D. (1974). Reaction time in focused and in divided attention. *Journal of Experimental Psychology*, *103*(3), 394.

Ocasio, W. (2011). Attention to attention. *Organization science*, *22*(5), 1286-1296.

Park, S., Kim, M. S., & Chun, M. M. (2007). Concurrent working memory load can facilitate selective attention: evidence for specialized load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1062.

Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, *21*(4), 469-514.

Pashler, H. (1994a). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Pashler, H. (1994b). Graded capacity-sharing in dual-task interference?. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 330.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.

Ponds, R. W., Brouwer, W. H., & Van Wolffelaar, P. C. (1988). Age differences in divided attention in a simulated driving task. *Journal of Gerontology*, *43*(6), P151-P156.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Ruthruff, E., & Pashler, H. (2001). Perceptual and central interference in dual-task performance. *The limits of attention: Temporal constraints in human information processing*, 100-123.

Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement?. *Psychonomic bulletin & review*, *8*(1), 73-80.

Schmidt, M. E., & Vandewater, E. A. (2008). Media and attention, cognition, and school achievement. *The Future of children*, 63-85.

Semple, R. J. (2010). Does mindfulness meditation enhance attention? A randomized controlled trial. *Mindfulness*, *1*(2), 121-130.

Shipstead, Z., Lindsey, D. R., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116-141.

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., ... & Schooler, J. W. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PloS one*, *6*(3), e18298.

Somberg, B. L., & Salthouse, T. A. (1982). Divided attention abilities in young and old adults. *Journal of Experimental Psychology: human perception and performance*, *8*(5), 651.

Sullivan, L. (1976). Selective attention and secondary message analysis: A reconsideration of Broadbent's filter model of selective attention. *Quarterly Journal of Experimental Psychology*, *28*(2), 167-178.

Svetina, M. (2016). The reaction times of drivers aged 20 to 80 during a divided attention driving. *Traffic injury prevention*, *17*(8), 810-814.

Tang, Y. Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews Neuroscience*, *16*(4), 213-225.

Tombu, M., & Jolicœur, P. (2002). All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. *Psychological research*, *66*(4), 274-286.

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 3.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248.

Treisman, A. M. (1964). Selective attention in man. *British medical bulletin*, *20*(1), 12-16.

Treisman, A., & Geffen, G. (1967). Selective attention: Perception or response?. *Quarterly Journal of Experimental Psychology*, *19*(1), 1-17.

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological review*, *95*(1), 15.

Valentine, E. R., & Sweet, P. L. (1999). Meditation and attention: A comparison of the effects of concentrative and mindfulness meditation on sustained attention. *Mental health, religion & culture*, *2*(1), 59-70.

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review*, *25*(6), 2005-2015.

Wickens, C. D. (1976). The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 1.

CHAPTER 2: LISTENING IN NOISE AND WORKING MEMORY: PUPIL DILATION AS AN INDEX OF ATTENTION ALLOCATION

Introduction

The research area of Cognitive Hearing Science (CHS) is concerned with the complexities of understanding language and especially understanding language in difficult listening conditions (Arlinger et al., 2009). Researchers in this area often take a multi-disciplinary approach, studying technologies to aid hearing and the changes in hearing that accompany development and aging, while also focusing on fundamental issues underlying the interaction of cognition and hearing. This multidisciplinary approach has the broad goal of understanding the complexities of everyday listening to language as well as differences across individuals and situations that affect how we hear.

Interestingly, higher order cognitive capacity, often measured in terms of working memory resources, has been shown to affect hearing ability. These effects on hearing ability have been measured as a function of both individual and situational differences in cognitive capacity—speech perception is affected both by an individual's cognitive ability and by cognitive loads applied from the environment to the individual (Akeroyd, 2008; Heinrich et al., 2020). The theory that hearing ability is tied closely to broader cognitive ability is central to the Ease of Language Understanding (ELU) model, proposed by Rönnberg et al. (2013). This model proposes that when an auditory speech signal is encountered, it is compared to expected phonology. If the speech signal matches the expected phonology, then it is easily processed. In contrast, when the speech signal does not match the expected phonology, perhaps due to factors such as background noise and distracted attention, the speech signal must be processed further using available cognitive resources to interpret the signal correctly.

Similar to the ELU model, the Framework for Understanding Effortful Listening (FUEL) is centered on the idea that limited cognitive resources are available to support listening in a variety of situations (Pichora-Fuller et al., 2016). The FUEL expands on Kahneman's model of attention, where available attentional resources are allocated to incoming information in accord with a variety of demands and priorities (Kahneman, 1973). The FUEL describes a host of input demands (which are any aspects of the auditory signal that can affect how demanding the signal is on attentional resources, such as background noise and speaker characteristics) and other factors (which can range from personal factors, such as fatigue, to external factors, such as external motivators to complete a task) that tax available cognitive resources, and that constrain resource allocation in ways that produce measurable outcomes. Many of these measurable outcomes are used as indicators of listening effort; the more attentional resources are required to deal with the input demands, the more effortful the listening. This effort can then be measured through a variety of outputs including subjective ratings of listening effort, physiological indices of listening effort, and behavioural costs associated with limited attentional resources.

The current study uses the FUEL as a framework to investigate listening effort and cognitive resource use. To our knowledge, no prior study has combined listening in noise with an n-back working memory task and measured pupillary response as an index of cognitive effort and attention allocation. To combine these methods, I had participants listen to speech in noise while simultaneously engaged in a working memory task. I examined behavioural measures, subjective effort, and pupil dilation as a physiological index of effort to investigate how cognitive resources are allocated in this dual-task listening in noise context. The specific rationale for the study rests on foundational work in the areas of working memory, the use of pupil dilation to study cognitive effort, and the effect of input demands on listening in noise.

Brief summaries of these three areas are provided below, followed by a description of the rationale for the present study.

Working Memory as a Measure of Cognitive Ability

Working memory is often thought of as a set of processes that maintain and update information in active memory (Baddeley, 1992; Baddeley & Hitch, 1974). The Baddeley and Hitch (1974) model of working memory includes a central executive that controls a phonological loop and a visuospatial sketchpad. These two subcomponents of working memory allow for maintenance and updating of auditory and visual information over timescales typically ascribed to short-term memory. The central executive connects this temporarily held information to long-term memory through the retrieval of relevant memories and the encoding of new memories (Baddeley & Logie, 1999). More recently, Baddeley (2000) updated the working memory model to include an episodic buffer, an additional temporary storage component that combines the different modalities of incoming information and integrates them with episodic long-term memory. This conception of working memory therefore consists of components that allow new auditory and visual information to be encoded in temporary memory, and then also updated, manipulated, and applied to further downstream processing.

Given the important role of maintenance, updating, and manipulation of information in this model of working memory, tasks that focus on these more central processing components of the model have been used to measure and understand capacity limits of working memory. In fact, it is now well-established that there are reliable individual differences in working memory capacity. Conceptualized as a limited pool of cognitive resources, working memory capacity defines a limit to what and how much can be maintained and updated in active memory (Cowan, 2010; Engle, 2002; Wilhelm et al., 2013). There are many ways to measure working memory capacity, including the use of complex span tasks (Conway, 1996; Conway et al., 2005; Daneman & Carpenter, 1980); and there are many tasks—such as verbal or visual memory tasks or motor tracking—that are assumed to tax the working memory system in a variety of modalities (Cocchini et al., 2002; Engle, 2002). Engle (2002) described working memory capacity to be dependent on one's ability to use attentional resources. Working memory capacity is therefore thought to depend on the allocation of attentional resources and the control of this allocation process, much like the "allocation policy" in Kahneman's (1973) model of attention and effort, and later Pichora-Fuller et al.'s (2016) framework of listening effort.

In addition to working memory tasks being used to measure cognitive ability, they can be used to add a cognitive load to other tasks. The logic here is that tasks that have memory updating and maintenance demands will tap a limited capacity resource available for such activities, and thereby constrain the working memory resources available for other tasks. One way to load working memory in this way is using the n-back task. The n-back task requires participants to compare a stimulus on the current trial to one n trials back, effectively forcing participants to maintain n stimuli in mind while also updating what they are holding in memory every trial. For example, in a numerical 2-back task, participants observe the stimuli presented on trials one and two (e.g., the numbers 4 and 8) and then compare the stimuli presented on trials three, four, and so on with the stimuli from two trials back. So, for this example, if the stimulus on trial three were the number 4, it would match the number from two trials back; if the following stimulus on trial four were the number 7, then it would not match the number from two trials back. The n-back is a common working memory task used in studies with a wide range of methods, including behavioural tasks and fMRI and EEG approaches (Engle, 2010; Meegan et al., 2004; Owen et al., 2005). The n-back, therefore, is commonly used to add a cognitive load to working memory (Jaeggi et al., 2010; Kane et al., 2007; Redick & Lindsey, 2013).

Pupil Dilation Indicates Cognitive Effort

A common measure of cognitive effort associated with the load produced by a working memory task is pupil dilation. Pupil dilation is widely recognized to be an autonomic physiological correlate of cognitive work and effort (Beatty, 1982; Beatty & Lucero-Wagone, 2000; Kahneman & Beatty, 1966). Although the pupil also dilates in response to sensory changes, such as changes in luminance, the psychosensory pupil response (PPR) specifically reflects the link between psychological effort and the body's physiological response to that effort (Mathôt, 2018; Unwsorth and Robison, 2018). By holding constant the sensory components of a task and only varying the cognitive demands, pupil dilation differences across conditions can be attributed to cognitive work rather than sensory differences. Greater pupil dilation is associated with greater cognitive effort and work.

Following early studies on the PPR, the pupillary response has since been used extensively with working memory tasks, to measure the load on working memory and the amount of cognitive effort being used to maintain and update information in working memory. Pupil dilation studies have examined individual differences in working memory capacity (Miller et al., 2019; Unsworth & Robison, 2015), as well as the contributions of working memory and cognitive effort to visual search, learning, and long-term memory encoding (Attar et al., 2016; Heitz et al., 2007; Miller et al., 2019; Sibley et al., 2011). In accord with early work in this area, there is much converging evidence that the pupil dilates in response to greater working memory load (Attar et al., 2016; Hjortkjær et al., 2020; Van Gerven et al., 2004; Zekveld et al., 2019).

Of particular relevance to the current study, many prior studies have examined the pupillary response to cognitive load using n-back tasks, in some cases looking at pupillometry in concurrence with EEG or other physiological measures (Hjortkjær et al., 2020; Scharinger et al., 2015). One useful property of the n-back is that it comes in many shapes and sizes—the value of n and the modality (e.g., visual or auditory) in which the n-back stimuli are presented can vary. Past studies have indeed found that the pupil dilates in response to greater memory loads during the n-back, meaning a 3-back task would elicit a greater pupillary response than a 2-back task, and so on, though there is a limit to the memory load beyond which the pupil will no longer dilate further (Ewing & Fairclough, 2010). This finding reflects the engagement and motivation of participants, where a task that is too high in cognitive load, and therefore too difficult, leads to a withdrawal of cognitive and attentional resources instead of additional allocation of resources. The n-back is most commonly a visual task, but studies have also been done with auditory versions of the n-back, which similarly find greater pupil dilation as a result of greater memory load (Hjortkjær et al., 2020). Overall, there is strong support in the literature for the idea that the pupil dilates in response to memory load in the n-back task and that it therefore provides a measure of cognitive work in such tasks.

Listening to Speech in Noise, Working Memory, and Pupil Dilation

Studies in the field of CHS have often explored the resource demands of listening to speech in noise by adding a concurrent secondary task to the primary listening in noise task. Such secondary tasks can provide a measure of the resource demands of listening in noise—specifically, secondary task performance may worsen when the level of input demands of the listening task are high compared to when the demands are low. Alternatively, the addition of a secondary task could result in costs to the performance of the listening in noise task itself, particularly when input demands are high (for a review of secondary tasks used with a primary listening task, see Gagné et al., 2017). These dual-task studies have used a variety of secondary tasks such as visual probes, tactile pattern recognition, and recall of the presented auditory stimuli (Anderson-Gosselin & Gagné, 2011; Hornsby, 2013; Ng et al., 2015; Picou & Ricketts, 2014). A common feature of many secondary tasks is that they occur simultaneously and continuously with the primary listening task. For example, past studies have used listening as the primary task with temporally simultaneous secondary tasks such as visual-motor tracking, where the participant continuously tracks a moving visual target with a motor-controller, and driving tasks, where staying on-target means staying in the designated lane and not deviating (Desjardins & Doherty, 2013; Desjardins & Doherty, 2014; Wu et al., 2014; Xia et al., 2015).

Much like these continuous tracking tasks, working memory tasks can be used as secondary tasks to maintain a cognitive load throughout performance of a primary task. Application of this research strategy to the primary task of listening to speech in noise is a particularly interesting possibility because of what is known about pupil dilation both in studies of working memory and in studies of listening to speech in noise. In the field of CHS, the pupillary response has been measured while participants listen to speech in noise, and the results suggest that increased pupil dilation in response to more difficult listening conditions captures the amount of effort one expends to hear correctly (Miles et al., 2017; Pichora-Fuller et al., 2016; Winn et al., 2018). In the FUEL, physiological measures such as pupil dilation fall under the category of output responses, which measure the use of cognitive and attentional resources when listening under a variety of demands. The fact that there is also strong evidence that pupil dilation responds to memory load in the n-back task points to the possibility that working memory tasks such as the n-back and listening to speech in noise may be constrained by similar resource limits to those measured by pupil dilation.

The Present Research

Pupillometry has been used in the field of CHS to study input demand effects on listening effort, and it has also been used to study attention allocation and effort in the context of working memory tasks. Moreover, listening in noise studies have also utilized dual-task methods to capture input demand effects on limited capacity attention. The effect of input demands on the allocation of attention and effort are revealed by input demand effects on secondary task performance. Despite the strong thematic overlap associated with these three methodologies, to my knowledge they have not previously been examined together in the same study. Therefore, the goal of this study is to observe the effects of listening input demands and divided attention constraints on attention allocation and, as they are assumed to tap into the same resources, to see how these effects are reflected in pupil dilation patterns. To that end, the current study draws on ideas from CHS and the FUEL to examine how listening in noise affects the simultaneous completion of a secondary n-back working memory task (as measured by behavioural responses in the n-back task), and in particular how this effect is reflected in the pupillary response. Given that listening in noise is well-understood to be attentionally demanding, and that the n-back task is a working memory task similarly understood to draw on attentional resources, I examined whether these two tasks draw on the same pool of limited resources. Moreover, given that pupil dilation reflects cognitive effort and attention allocation, I examined whether the combined draw on resources of these two tasks would be reflected in the pupillary response. Put most simply and broadly, my aim was to examine whether the attentional demands of a continuous working memory load and those of listening in noise are related.

A useful framework to understand potential outcomes of this combination of methods is the FUEL (Pichora-Fuller et al., 2016). This framework considers how a variety of input demands affect resource allocation and listening effort, and how these processes map onto a variety of output measures. The current study uses the n-back task, a task commonly used in neurophysiological studies, as a secondary task to tax working memory while participants complete a listening in noise task. A key component of the study is the combination of behavioural and physiological methods; the pupillary response is used here as an indicator of cognitive work and listening effort. This particular combination of tasks and measures will allow me to address the following aims. First, I aim to replicate previous findings showing that the effect of noise on speech identification can be measured in the pupillary response, to further support the use of pupil dilation as a measure of cognitive effort when listening in noise. Second, I aim to replicate previous findings showing that the effect of a working memory load, implemented here through use a secondary n-back task, can also be measured in the pupillary response; again, this is to ensure that I am measuring what I intended to measure. Most important, by combining these methods, I aim to study the interaction between input demand effects in the listening in noise task and working memory load effects in the n-back task. As both of these effects are known to affect the pupillary response, and the pupillary response is deemed to reflect the allocation of limited capacity cognitive resources and effort, it seems likely that these two effects would interact. The FUEL would predict that the input demands of differing noise conditions in the primary task should tax attentional resources and be reflected in the pupillary response. Additionally, the cognitive load of the n-back task should simultaneously tax attentional resources and also be reflected in the pupillary response. Pupils are expected to dilate more when performing the listening task and the n-back compared to performing just the listening task alone. Each of these task demands is expected to have its own effect on attention allocation, as reflected in the pupillary response, but I am also interested to know if the combination of the two tasks will interact and if that will be seen in the pupillary response. This experiment will highlight how the pupil responds to competing demands on attentional resources from both cognitive load and listening effort.

Methods

Participants

A total of 53 undergraduate students participated in this experiment for partial course credit. Data from two participants were excluded from analysis because of technical difficulties, and data from another 23 participants were excluded due either to performance that was more than one standard deviation below mean performance on the listening task (15 participants) or performance that was at chance (or worse) on the secondary task (8 participants).The exclusion criteria of one standard deviation below the mean was used for the listening task because there was no chance level available as the task was to repeat the presented words out loud. The exclusion criteria of chance performance was used for the secondary task because participants only had two possible responses for the n-back task. Data from the remaining 28 participants (19 females and 9 males; mean age = 18.3) were used for the analyses. Table 1 displays a summary of participants (both included and excluded) and their corresponding individual signal-to-noise-ratio (SNR) values. The Speech Reception Threshold (SRT) procedure used to determine these SNRs is described below. All participants had self-reported normal or corrected-to-normal vision and self-reported normal hearing; all participants learned English before the age of 7.

Apparatus and Stimuli

The visual stimuli used in the secondary, n-back task were presented on on a Dell UltraSharp 2001FP LCD 17-inch monitor running from an HP Pro 3130 MT computer, using Presentation® software. The stimuli for the secondary task were single digit numbers, presented in 24-point Arial font, in isoluminant green on a grey background. An isoluminant fixation cross was also presented on each trial until the secondary task stimulus was presented. Participants were seated with their chin in a chinrest that was positioned 20 cm above the desk and 40 cm from the screen to maintain eye position throughout the experiment. Participants responded using a standard mouse, making either a right or left mouse click to each of the secondary task stimuli (when a response was necessary).

The listening task auditory stimuli were presented through two Genlec 8020D Bi-amplified Active Studio Monitor Speakers running from the same HP Pro 3130 MR computer. One speaker was located directly above the monitor that presented the n-back task stimuli visually, while the other was located approximately 65 cm behind the participant. The volume was set such that the target stimuli were presented at 70 dB and the background noise was set in accord with each participant's SRT (procedure described below). The primary task stimuli were the 200 high predictability sentences of the Revised Speech Perception in Noise Test (R-SPiN), where predictability refers to how predictable the final word is based on the rest of the sentence. The background noise was either shaped noise, created in MATLab using code from Hristo Zhivomirov (2013), or a four-talker babble track produced by Auditec (1971). Participants' verbal responses to the sentences presented in shaped noise, babble noise, or silence were manually recorded by the experimenter into an Excel spreadsheet in real-time. During the SRT measurement, participants were faced away from the monitor, directly facing the back speaker. Individual SRT values were determined for each participant using a speech-in-noise algorithm in MATLAB (Rao & Starkey, n.d.).

Working memory span was measured separately from the main experiment, using the Operation Span and Reading Span, provided with permission by the Attention and Working Memory Lab at Georgia Institute of Technology (Foster et al., 2015; Oswald et al., 2015). These tasks were presented in E-Prime from the same computer and monitor as the secondary task stimuli, and were fully automated with all instructions on-screen. These tasks provided a measure of each participant's working memory span.

Pupil dilation was recorded by a Model ET1000 EyeTribe eye tracker that was placed right beneath the monitor, turned up toward the participant. The EyeTribe software was used to calibrate and record the pupil data and was synced with the experiment data using Presentation®. Trigger markers from Presentation® were compiled with time-stamped sampling data from the EyeTribe in LabStreamingLayer, and output as .xdf files that were used for later analysis. The data were collected at a rate of 30Hz.

Procedure

Operation Span

The Operation Span task (Foster et al., 2015; Oswald et al., 2015) presented participants with math equations (e.g., 8-3 = 5) and the participants indicated, using a right or left mouse click, if the equation was true or false. Beneath the equation, a letter was presented and participants were instructed to remember the letter. After two to seven trials of this task, a matrix of letters appeared on the screen and the participant had to click the letters in the order in which they appeared under the equations. There were 15 blocks of these trials.

Reading Span

The Reading Span task (Foster et al., 2015; Oswald et al., 2015) was the same as the Operation Span task, except instead of equations, participants were shown sentences (e.g., The sky was full of spaghetti) and asked to indicate whether the sentence made logical sense semantically or not. Once again, each trial presented a letter that the participant had to remember and recall at the end of the block. There were six blocks of these trials.

SRT Measurement

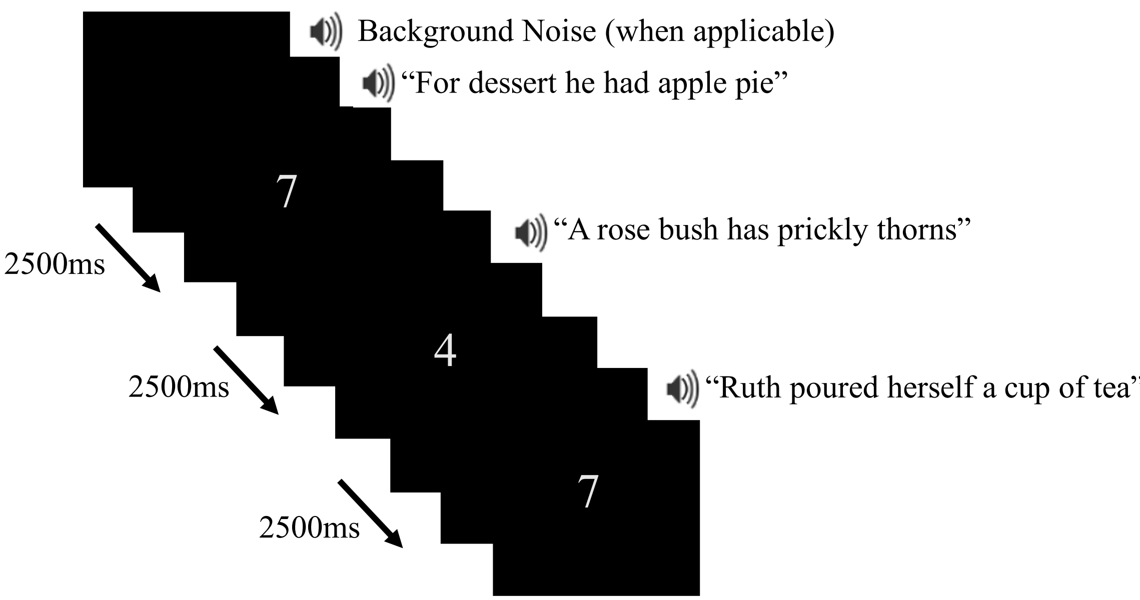
Participants sat between the two speakers and were instructed to keep their head facing the back speaker. The SRT software used a staircase adjustment to determine the individual participants' SRT at 60 percent correct critical word identification when repeating the sentences aloud. The critical words, of which there were several in each sentence, were the ones whose identification contributed to participants' overall accuracy on that sentence. Sentences were presented in babble noise, with the sentences presented from the speaker in front of the participant and babble noise from the speaker behind, beginning at an SNR of zero. Participants were instructed to listen to the whole sentence and then repeat the entire sentence aloud. The experimenter input the number of critical words within the sentence that the participant correctly repeated; the software then adjusted the SNR accordingly over 20 sentences. The measured SNR was then recorded and used in the listening task.

Listening Task

Figure 1 depicts an example of a sequence of trials in the listening task. Before completing the listening task of the experiment, the eye tracker was calibrated to each participant's eyes using the EyeTribe software. A total of 192 trials was presented—twelve blocks of 16 trials each. In the blocks where there was background noise present – shaped noise and babble noise – the noise was presented from the back speaker, and all sentence stimuli were presented from the front speaker. The background noise played continuously throughout the block, and thus the sentence always overlapped temporally with the background noise. The sentences ended with a target word, and participants were instructed to repeat that target word aloud. The experimenter then recorded the participant's response. A single digit appeared on the screen 2500 ms after the onset of the target word. In the single task blocks, participants spoke aloud the target word and ignored the visually presented digit that followed. In the dual-task blocks, participants first spoke aloud the target word and then performed a 2-back task with the digit that appeared 2500 ms after the auditory target word; participants compared the digit in the current trial to the digit that was presented two trials back (starting on the third trial). Participants indicated using a left or right mouse click whether the current digit was the same as two trials previous (left mouse click) or different from two trials previous (right mouse click). Accuracy and response time on this task were recorded. After each block of 16 trials, participants gave their rating of listening effort.

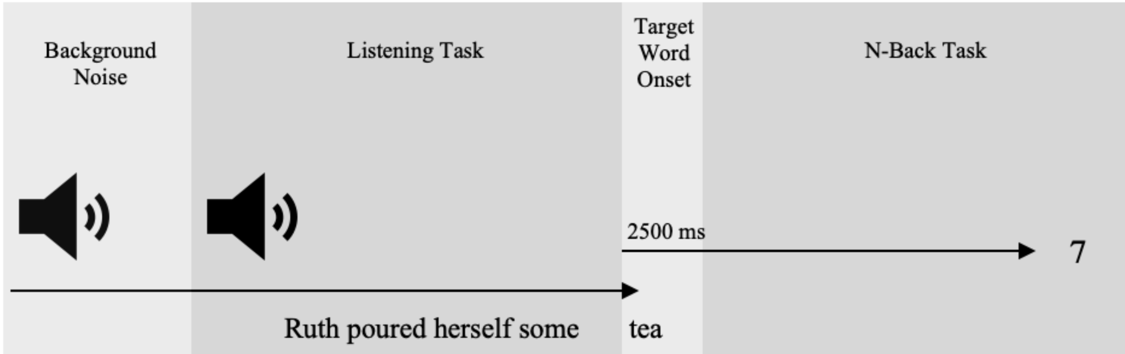
Figure 1

Figure 1A: Listening Task Trial Sequence



A sequence of trials in the experiment. The background noise turns on at the start of each block and a number is presented 2500 ms after the onset of the final word in each auditory sentence. In this example, a participant would be correct if they responded that the number 7 on the third trial is the same number as the one from two before.

Figure 1B: Sequence of Events in a Trial



The sequence of events throughout the trial, with the sentence stimulus presented in the background noise and then the number for the n-back task presented visually 2500 ms after the onset of the target word at the end of the sentence.

Design

All participants completed the experiment in the following order: Operation Span, Reading Span, and listening task. The listening task consisted of 12 blocks of 16 auditory trials, for a total of 192 trials. For half the blocks, participants only responded to the listening task (the single-task condition), in which they were instructed to ignore the secondary task stimulus (a number presented visually) that appeared 2500 ms after the onset of the target auditory word. For the other half of the blocks, participants completed the dual-task condition, in which they were instructed to pay attention to the number that appeared 2500 ms after the onset of the target auditory word, and to perform an n-back task (comparing the number to one from two trials previous). Sets of three blocks alternated between single- and dual-task, always starting with the single-task condition. In this manner, participants completed three blocks of single-task, three blocks of dual-task, three blocks of single-task, and then three blocks of dual-task for a total of twelve blocks.

Each participant started with the ‘single-task’ condition for the first three background noise condition blocks, and then alternated with the ‘dual-task’ condition for the remaining 9 blocks. Within those sets of three blocks the conditions cycled between silence, shaped noise, and babble noise, the latter two of which were counterbalanced across participants. Every set of three blocks started with the silence condition first (e.g., single-task silence, single-task shaped noise, and single-task babble noise; dual-task silence, dual-task babble noise, and dual-task shaped noise; single-task silence, single-task babble noise, and-single task shaped noise; and dual-task silence, dual-task shaped noise, and dual-task babble noise). Whether shaped noise or babble noise followed silence in the first block was also counterbalanced across participants, where the opposite pattern of shaped noise and babble noise order was used for alternating participants. Additionally, there were four lists of sentences, which were also counterbalanced across background conditions and participants, such that the sentence lists were evenly distributed across the two counterbalanced orders of background conditions.

Subjective listening effort was rated after each block, using Johnson et al.'s (2015) listening effort scale. When prompting the participant to rate their listening effort, the experimenter specifically asked for listening effort, and informed the participant that this was not a rating of overall effort but just of how effortful it was to hear and identify the words.

Pupil Data Processing

Pupil data were processed in Matlab and the post-processing data were exported to .csv files to be used in further analysis. Trial markers collected from Presentation® experiment output were timestamp matched to pupil data samples and used to integrate behavioural data (performance on the listening task) with the pupil data. Only trials on which participants correctly identified the target word were included in the pupil analyses. The pupil data were processed as separate left and right eye channels until they were combined and averaged at the end of processing. Participants with one or more conditions that had more than one third of the data samples missing were excluded from further processing, resulting in exclusion of data from 6 participants and leaving data from 22 participants for analyses of pupillary responses. For the included participants, missing data points were filled in using linear interpolation between existing data points. A three-sample moving average filter was then used to smooth the data. No gaze correction algorithm was applied to the data because participants used a chin rest and maintained their gaze at the fixation point. Baseline corrected pupil size was calculated as the average measured pupil size from -2500 ms to 0 ms before the onset of the sentence, during which time the background noise was playing (where applicable). This baseline value was subtracted from each pupil size sampled and then used to determine the proportion change from baseline pupil size. The trial length was set to 7500 ms, with the onset of the start of the sentence marking the start of a trial (the 0 ms mark).

Data Analysis

In the analysis of behavioural data, the dependent variable of interest in the primary listening task was the proportion of target words correctly identified, whereas the dependent variables of interest in the secondary n-back task were mean response times (RT) from n-back stimulus onset, accuracy rates, and Inverse Efficiency Scores (IES; mean RT divided by proportion correct). Subjective listening effort measured after each block was also analyzed. For the primary listening task and listening effort measure, the dependent variables were analyzed using a repeated measures ANOVA that treated Background Condition (silence/shaped noise/babble noise) and Task (single/dual) as factors. For the secondary n-back task, the dependent variables were submitted to one-way repeated measures ANOVAs that treated Background Condition as a factor.

Reading Span and Operation Span results were scored using the partial-credit unit score method (Conway et al., 2005). The correlation between these scores and secondary task IES was computed. The rationale underlying this analysis is that secondary task performance may reflect residual resources after completion of the primary task. Those with fewer cognitive resources, and therefore lower scores on the Reading Span and Operation Span tasks, may show greater costs in secondary task performance, meaning worse performance on the n-back task, leading to a positive correlation between working memory capacity and secondary task performance. The correlation between Reading Span and Operation Span and two pupillary response measures was also computed. The two measures were absolute pupil size in the baseline period and mean peak dilation amplitude. The rationale underlying these analyses is similar to that of IES; the pupillary response is a measure of cognitive work and therefore those with fewer cognitive resources are expected to require more effort to complete the tasks than those with greater available resources, as reflected by greater absolute pupil size in response to background noise and greater peak pupil dilation in response to the listening task.

In the analysis of pupil dilation, the data were separated into three measures: 1) mean absolute pupil size during the baseline period (-2500 ms to 0 ms); 2) latency of peak pupil dilation calculated using baseline corrected data (using the baseline period of -2500 ms to 0 ms)—this measure identified the time point within each trial at which the pupil reached peak dilation relative to baseline, averaged within conditions for each participant; and 3) peak dilation amplitude calculated using the baseline corrected data—this measure identified the mean amplitude of pupil dilation relative to baseline in a 500 ms time window centered on the peak latency for each condition (plus and minus 250 ms, rounded to the nearest 33.33 ms due to the sampling rate of 30 Hz) and then averaged across participants. The time windows used for peak dilation amplitude are detailed in Table 1A. All pupil data were also submitted to repeated measures ANOVAs that treated Background Condition (silence/shaped noise/babble noise) and Task (single/dual) as factors.

For the behavioural data, Figures 2 to 6 include error bars with the Morey correction applied (O'Brien & Cousineau, 2014), which accounts for within-subjects designs when calculating the standard error of the mean.

Table 1

Individual participant inclusion and exclusion with criteria for failure to include and the corresponding SNR values

Participant	**Individual SNR (dB)**	**Inclusion/Exclusion**	**Criteria (reason for exclusion)**
P1	-13.25	Excluded	Listening task performance
P2	-11.94	Included	
P3	-6.13	Excluded	Technical difficulties
P4	-9.12	Included	
P5	-7.24	Included	
P6	-11.75	Excluded	Listening task performance
P7	-5.44	Excluded	N-back task performance
P8	-3.38	Excluded	N-back task performance
P9	-8.53	Excluded	N-back task performance
P10	-9.93	Excluded	Listening task performance
P11	-7.35	Included	
P12	-9.47	Excluded	Listening task performance
P13	-8.12	Included	
P14	-7.5	Included	
P15	-6.8	Included	
P16	-11	Excluded	Listening task performance
P17	-10.87	Excluded	Listening task performance
P18	-8.65	Excluded	N-back task performance
P19	-5.59	Included	
P20	-10.18	Excluded	Listening task performance

P21	-17.25	Excluded	Technical difficulties
P22	-12.88	Excluded	Listening task performance
P23	-5.5	Included	
P24	-9.82	Included	
P25	-11.29	Included	
P26	-4.88	Included	
P27	-9.35	Excluded	Listening task performance
P28	-10.5	Excluded	N-back task performance
P29	-6.75	Included	
P30	-6.5	Included	
P31	-14.06	Excluded	Listening task performance
P32	-10.75	Excluded	Listening task performance
P33	-16.37	Excluded	Listening task performance
P34	-9.35	Included	
P35	-12.29	Excluded	Listening task performance
P36	-6.5	Included	
P37	-8.88	Excluded	N-back task performance
P38	-8.19	Included	
P39	-5.56	Included	
P40	-7.94	Included	
P41	-7.35	Included	
P42	-5.25	Included	
P43	-6.35	Included	
P44	-9.94	Excluded	Listening task performance
P45	-14.29	Excluded	Listening task performance
P46	-2.41	Included	
P47	-7.13	Excluded	N-back task performance
P48	-10.13	Included	
P49	-9.57	Excluded	N-back task performance
P50	-9	Included	
P51	-5	Included	
P52	-5.19	Included	
P53	-6.24	Included	

The average SNR of participants excluded due to performance lower than one standard deviation below the mean on the listening task was -11.76 dB; the average SNR of the included participants was -7.25 dB.

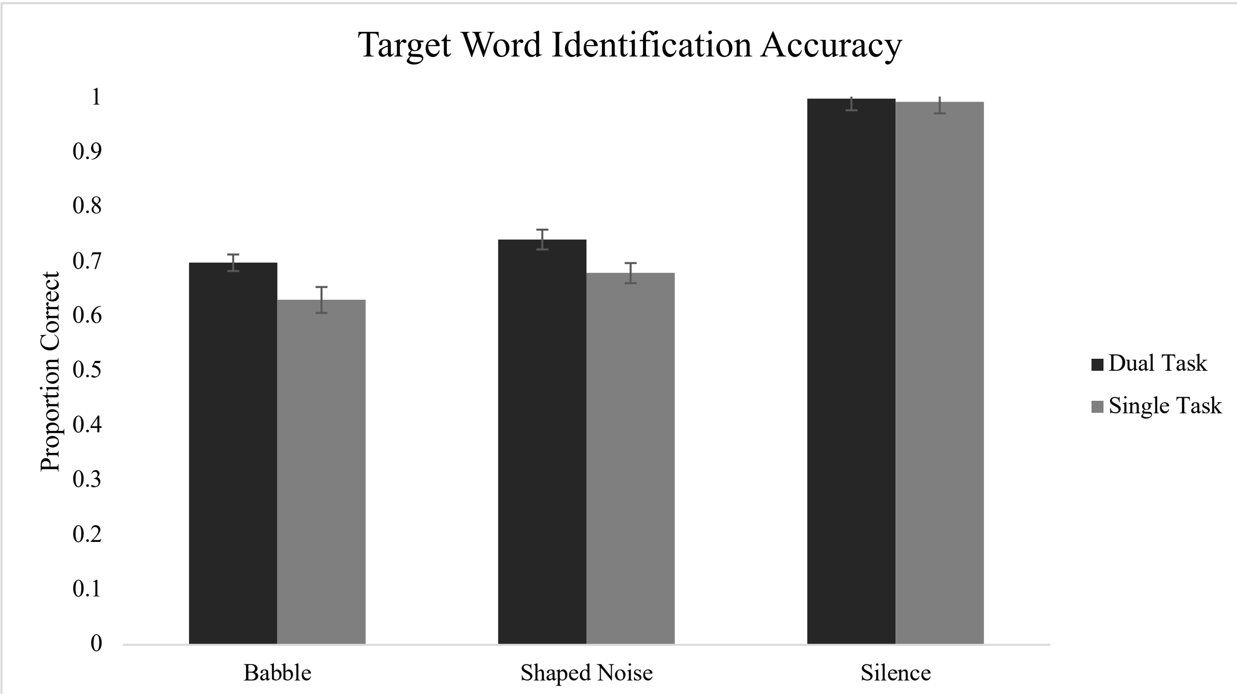
Results

Listening Task: Target Word Identification Accuracy

Figure 2 shows the proportion of correctly identified target words by condition in the listening task. There was a significant main effect of Background Condition, $F(2, 54) = 88.89$, $p < .001$, $\eta p^2 = .77$. Post-hoc Tukey tests indicated that target word identification accuracy was greater for silence (.996) than for both shaped noise (.71), $p < .001$, and babble noise (.67), $p < .001$, but that there was no significant difference between babble noise and shaped noise. There was also a main effect of Task, $F(1, 27) = 15.93$, $p < .001$, $\eta p^2 = 0.37$, with greater accuracy in the dual-task condition (.81) than the single-task condition (.77). The interaction between Background Condition and Task approached significance, $F(2,54) = 3.47$, $p = .05$, $\eta p^2 = .11$. Separate analyses of the effect of Task for each Background Condition revealed that accuracy was greater for dual- than single-task in both the babble (.70 vs .63, $p < .001$) and shaped (.74 vs .68, $p = .02$) noise conditions but not in the silence condition ($p > .10$).

Figure 2

Target Word Identification Accuracy



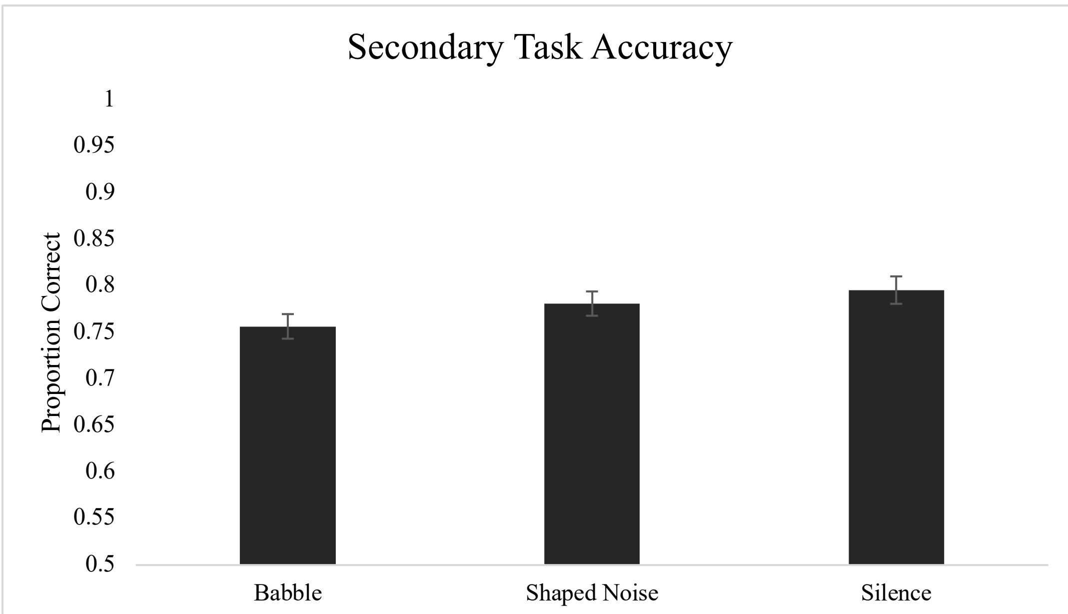
Error bars are standard error of the mean, with Morey correction applied.

N-Back Task: Accuracy, Reaction Time, and Inverse Efficiency Score

Figures 3, 4, and 5 display proportion correct, mean RT, and inverse efficiency scores respectively, for the secondary n-back task as a function of Background Condition. In the analyses of both proportion correct and mean RT, the effect of Background Condition was not significant, $F(2,54) = 2.08$, $p = .14$ and $F(2.54) = 2.01$, $p = .14$, respectively. However, in the analysis of IES, there was a significant effect of Background Condition, $F(2,54) = 4.09$, $p = .02$, $\eta p^2 = .13$. Post-hoc Tukey tests revealed that IES was higher for the babble noise condition (1529 ms) than for the silence condition (1377 ms), $p = .03$. The difference in IES between babble noise (1529 ms) and shaped noise (1390 ms) approached significance, p = .057. The difference between silence and shaped noise was not significant.

Figure 3

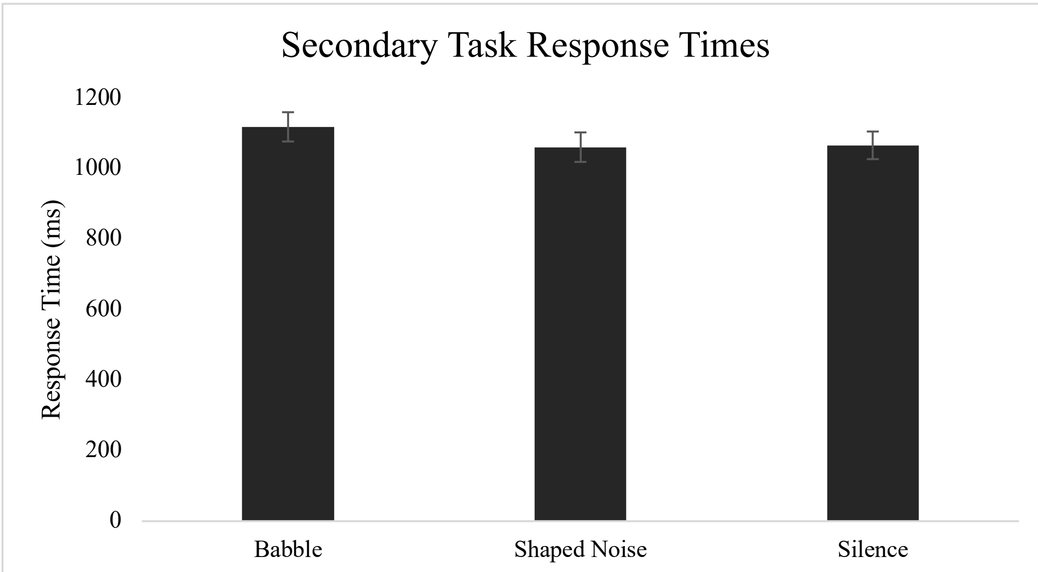
Secondary Task Accuracy



Error bars are standard error of the mean, with Morey correction applied.

Figure 4

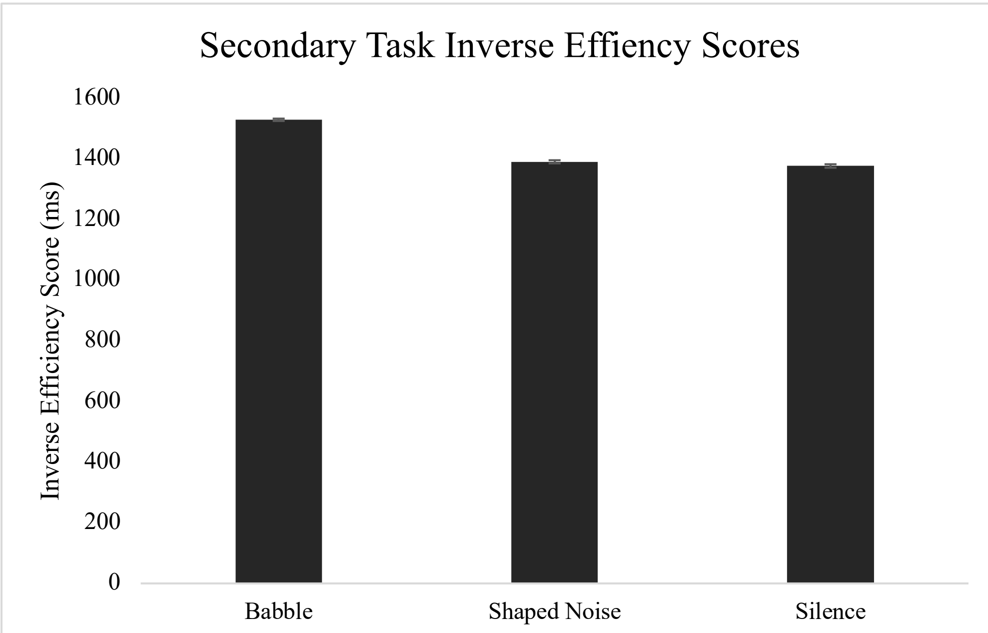
Secondary Task Response Times



Error bars are standard error of the mean, with Morey correction applied.

Figure 5

Inverse Efficiency Scores (IES)



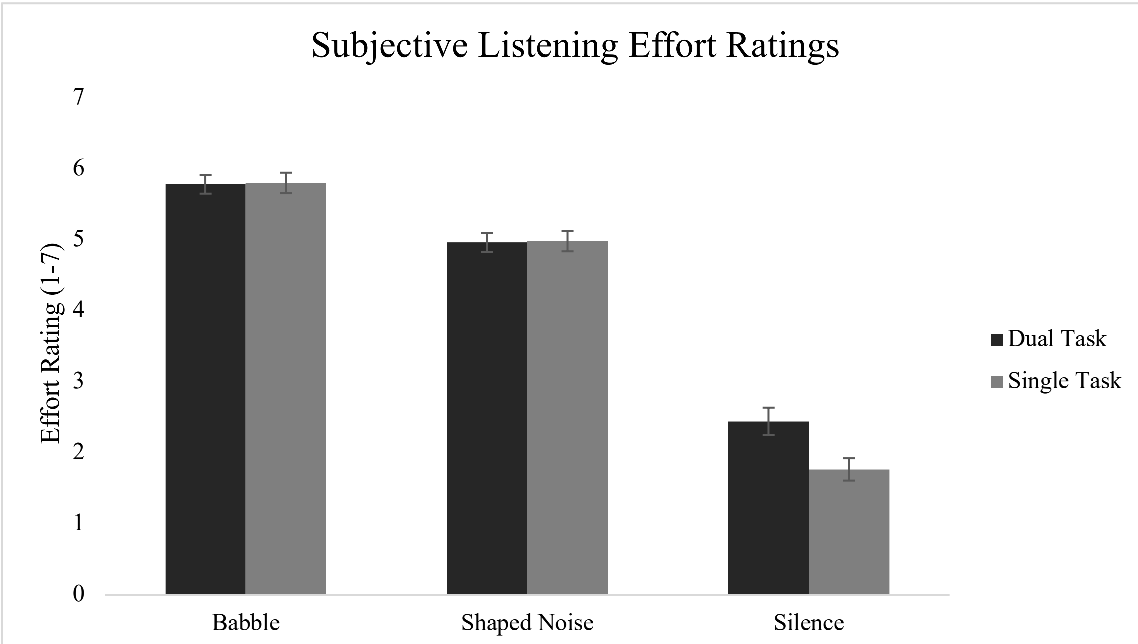
Error bars are standard error of the mean, with Morey correction applied.

Listening Effort

Figure 6 shows the subjective ratings of listening effort by condition. There was a significant main effect of Background Condition, $F(2, 54) = 203.3$ $p < .001$, $\eta p^2 = .88$. Post-hoc Tukey tests showed that subjective listening effort differed between all pairwise combinations of conditions, $p < .001$; it was highest for babble noise (5.79), intermediate for shaped noise (4.97), and lowest for silence (2.11). There was no main effect of Task, but there was a significant interaction between Background Condition and Task, $F(2.54) = 8.35$, $p < .001$, $\eta p^2 = .24$. For the silence condition, subjective ratings of listening effort were higher in the dual-task condition (2.45) than in the single-task condition (1.77), $p < .01$. In both the babble noise and shaped noise conditions, the difference in ratings between dual- and single-task conditions was not significant.

Figure 6

Listening Effort



Error bars are standard error of the mean, with Morey correction applied.

Working Memory

Reading Span and Operation Span scores were significantly positively correlated with each other $r = 0.64$, $p < .001$. However, neither of these measures of working memory correlated significantly with IES, absolute pupil size, or mean peak dilation amplitude in any condition. Appendix A shows all the correlations with working memory tasks.

Pupillary Response

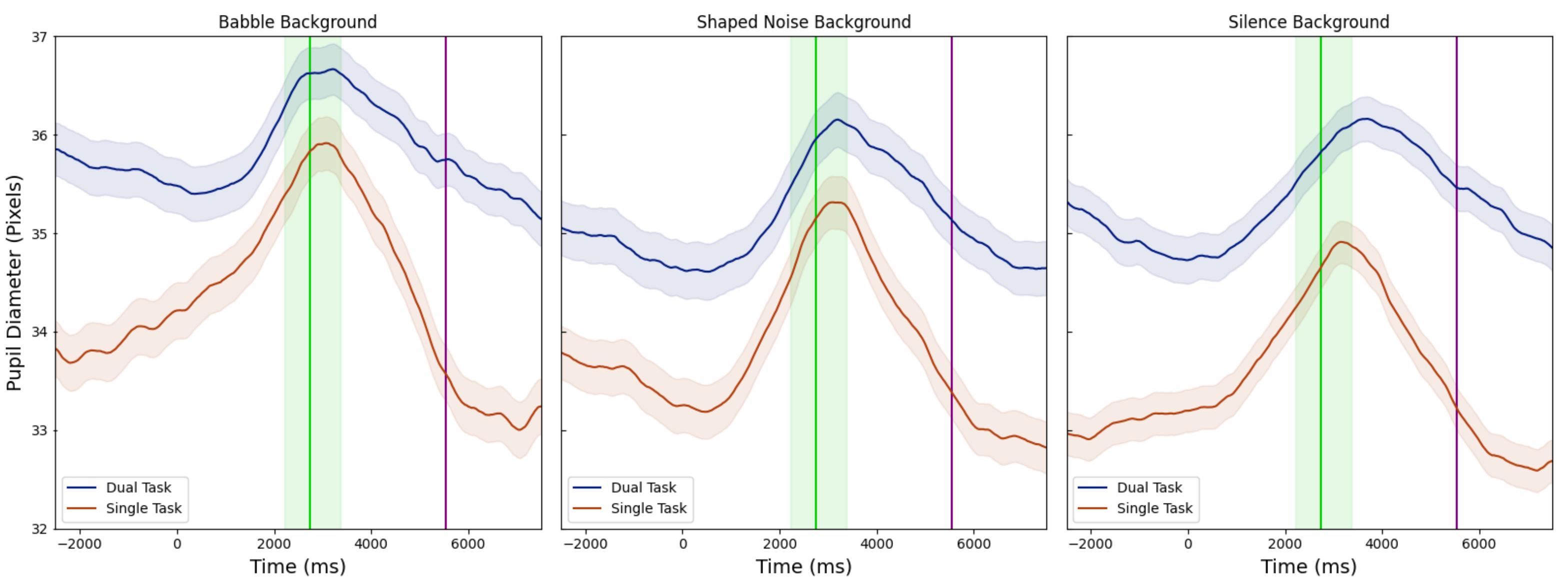
Absolute pupil size

Figure 7 shows the absolute pupil size across the trial time course for each combination defined by the Background Condition and Task variables. There was a significant main effect of

Background Condition, $F(2,42) = 10.53$, $p < .001$, $\eta p^2 = .33$. Post-hoc comparisons revealed that absolute pupil size in the baseline period (-2500 ms to 0 ms) was greater in the babble noise condition (34.7 pixels) than in both the silence condition (33.8 pixels), $p < .001$, and shaped noise condition (34.1 pixels), $p = .011$, whereas the difference between shaped noise and silence was not significant. There was also a significant main effect of Task, $F(1,21) = 14.76$, $p < .001$, $\eta p^2 = .41$. Absolute pupil size in the baseline period was greater in the dual-task condition (34.9 pixels) than in the single-task condition (33.5 pixels). The interaction between Background Condition and Task was not significant, $F(2,42) = 1.43$, $p = .25$.

Figure 7

Absolute Pupil Dilation



Pupil diameter, measured in pixels, for each Background Noise condition. Red lines represent single-task conditions, in which participants were only completing the listening task; and blue lines represent dual-task conditions, in which participants were completing the n-back task in addition to the listening task. Zero milliseconds is the onset of the sentence stimulus with the vertical green line representing the average onset time of the target word (2742 ms), the shaded green region representing the time window in which the target word onset occurs (2220 ms to 3367 ms), and the vertical purple line representing the average onset of the number for the n-back task (5542 ms).

Baseline corrected pupil dilation

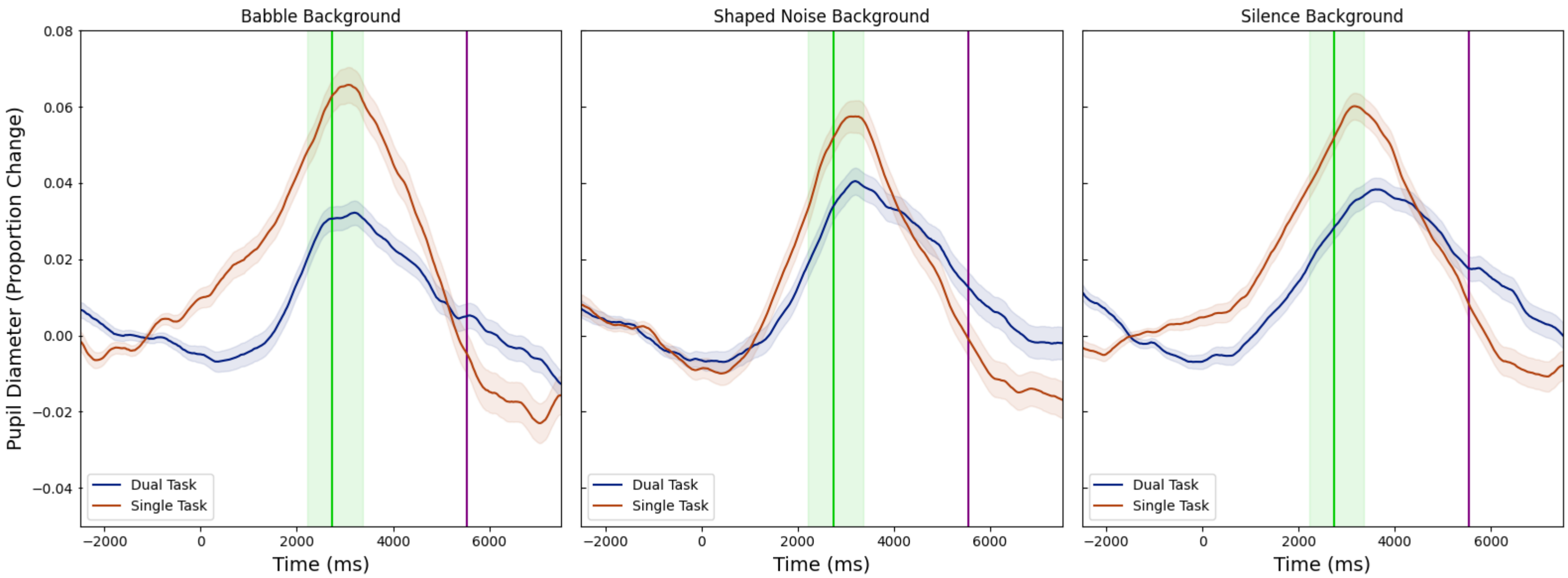
Figure 8 shows the baseline corrected pupil dilation across the trial time course, separated by Background Condition and Task. Tables 2A and 2B display the mean peak dilation amplitude and peak dilation latency for each condition.

In the analysis of peak dilation latency, there was a significant effect of Background Condition, $F(2,42) = 4.94$, $p = .012$, $\eta p^2 = .19$. Peak dilation latency was later for the silence condition (2928 ms) than for the shaped noise condition (2532 ms), $p = .013$. Neither of the two differences involving the babble noise condition (2618 ms) were significant. The effect of Task was non-significant, $F(2,21) = .03$, $p = .87$, as was the interaction between Background Condition and Task, $F(2,42) = .63$, $p = .54$.

In the analysis of peak dilation amplitude, there was a significant effect of Task, $F(1,21) = 20.09$, $p < .001$, $\eta p^2 = .49$. Peak dilation amplitude was greater in the single-task condition (.054 proportion change from baseline) than in the dual-task condition (.029 proportion change from baseline), $p < .001$. The effect of Background Condition was non-significant, $F(2,42) = 1.17$, $p = .32$, and the interaction between Background Condition and Task was also non-significant, $F(2,42) = 1.46$, $p = .25$.

Figure 8

Change in Pupil Dilation from Baseline



Proportion change in pupil diameter from baseline period of -2500 ms to 0 ms for each Background Noise condition. Red lines represent single-task conditions, in which participants were only completing the listening task; and blue lines represent dual-task conditions, in which participants were completing the n-back task in addition to the listening task. Zero milliseconds is the onset of the sentence stimulus with the vertical green line representing the average onset time of the target word (2742 ms), the shaded green region representing the time window in which the target word onset occurs (2220 ms to 3367 ms), and the vertical purple line representing the average onset of the number for the n-back task (5542 ms).

Table 2A

Average pupil dilation at peak dilation latency (in proportion change dilation from baseline)

	Babble Noise	Shaped Noise	Silence
Single-Task	.061 (2444 – 2944 ms)	.046 (2297 – 2797 ms)	.056 (2618 – 3118 ms)
Dual-Task	.026 (2292 – 2792 ms)	.028 (2267 – 2767 ms)	.033 (2738 - 3238 ms)

Values in brackets are the time windows (in milliseconds after sentence onset at 0 ms) over which pupil dilation was averaged. This time window was determined as the 500 ms surrounding the peak dilation latency (plus and minus 250 ms); peak dilation latencies are detailed in Table 2B.

Table 2B

Average peak dilation latency (in milliseconds) in single- and dual-task conditions

	Babble Noise	Shaped Noise	Silence
Single-Task	2694	2547	2868
Dual-Task	2542	2517	2988

Discussion

The goal of this study was to examine the joint consequence of resource demands associated with listening in noise, and resource demands associated with a concurrent working memory task. A key feature of the study was use of the pupillary response as an indication of resource use. To our knowledge, the present study was the first to examine these two resource demand issues in the same experimental context. Studying these demand issues in this context provides insight into how attention allocation and cognitive resource use—which is assumed to be integral to listening in noise, working memory, and the pupillary response—changes in response to certain task demands and manipulations.

Prior listening in noise studies suggest that increased input demands are met with increases in the allocation of cognitive resources (Pichora-Fuller et al., 2016; Rönnberg et al., 2013). According to the ELU model (Rönnberg et al., 2013), when a speech signal is not easy to parse from noise, additional cognitive resources are allocated to help interpret the speech signal, and listening is deemed subjectively to be more effortful. Similarly, in the FUEL (Pichora-Fuller et al., 2016), the input demands of the incoming speech signal are subject to an attentional resource allocation process, with more resources allocated to process the speech signal when input demands are high, together with higher subjective estimates of listening effort. The FUEL also posits that the outputs of attentional resource allocation can be measured in a variety of ways, including physiological measures of effort.

Similarly, prior working memory studies suggest that increased working memory load is met with an increase in the allocation of cognitive resources. As more information is encoded and maintained in working memory, more resources appear to be drawn from a limited capacity pool that supports working memory (Cowan, 2010; Engle, 2002; Wilhelm et al., 2013). As in studies of listening in noise, physiological measures have been used to index the allocation of effort and resources in working memory tasks. The pupillary response is one such measure, with the pupil dilating in response to greater cognitive work and effort in working memory tasks (Mathôt, 2018; Unwsorth & Robison, 2018). Of particular relevance here, the pupil is known to dilate in accord with working memory load in the n-back task (Ewing & Fairclough, 2010).

In the present study, I examined cognitive resource allocation associated with listening in noise concurrently with cognitive resource allocation associated with working memory demand. Prior research suggested that each of these two tasks should be sensitive to resource demands as reflected in the pupillary response (Mathôt, 2018; Unwsorth & Robison, 2018). The primary goal here was to study the joint effects of resource demands from both tasks completed concurrently.

In the behavioural results, I found many of the anticipated effects of input demands on performance. There was an effect of background noise on listening task accuracy, where listening in silence led to better performance than listening in the two 'noise' conditions. This effect of background noise also carried over to the n-back task, where IES for the n-back task was highest in the babble noise condition. Interestingly, the dual-task manipulation (single-task vs n-back task) also affected listening task performance; there was greater listening accuracy in the dual-task condition than in the single-task condition (speculation on the source of this counterintuitive result is provided below). And finally, subjective listening effort varied with input demands largely as expected, with highest subjective effort in the babble noise condition and lowest subjective effort in the silence condition. Overall, these behavioural findings show that input demands do affect both listening task performance and working memory performance, and that these demands are reflected in subjective perceptions of effort.

In the physiological results, I again observed effects that were aligned with prior studies. Pupil size in the baseline period was highest for the babble noise condition and lowest for the silence condition (Koelewijn et al, 2012; Wendt et al., 2018; Zekveld et al., 2014). Pupil dilation in the baseline period was also higher for the dual-task condition (the n-back task) than for the single-task condition. Interestingly, the task-evoked pupil response (the baseline corrected measure of peak amplitude) showed greater dilation for the single-task condition than for the dual-task condition. This result points to the possibility of overlapping resource demands for the n-back task and listening task. The implications of these findings are discussed in further detail below.

Listening task input demands: Effects on listening task performance, n-back performance, and subjective listening effort

Background noise was manipulated during the listening task, where two 'noise' conditions—shaped noise and babble noise—and a silence condition were studied. As expected, the presence or absence of background noise did affect listening task performance, with better performance for the silence condition than for the two 'noise' conditions. Surprisingly, the two 'noise' conditions did not differ. This result was unexpected because previous studies have found that the combination of signal and informational masking associated with babble noise leads to worse word identification than signal masking alone associated with shaped noise (Bennet et al., 2012; Desjardins & Doherty, 2013; Hall et al., 2002; Schneider et al., 2007). While participant performance was titrated to 60% accuracy using the babble noise background, the SNRs appear to have dampened word identification performance in both 'noise' conditions to the same extent.

It is unclear why the two 'noise' conditions did not differ, but perhaps this occurred because the most salient contrast was between either type of 'noise' and silence.

As the input demands in a listening task increase, performance on a secondary task can also be affected (see Gagné et al., 2017 for a review of listening in noise while performing an additional task). Indeed, in the current study listening task input demands affected IES in the n-back task. Specifically, inverse efficiency scores were higher (i.e., efficiency was lower) in the babble noise condition than in the shaped noise and silence conditions. This effect of listening task input demands on secondary task performance supports the proposal that cognitive resources recruited to manage listening in noise are related to those used to maintain and update the information in working memory in the n-back task.

Input demands in the listening task also affected subjective listening effort. Subjective listening effort was greatest for the babble noise condition, intermediate for the shaped noise condition, and lowest for the silence condition. These ratings reflect the anticipated patterns of subjective effort between the listening conditions (Klink et al., 2012; Krueger et al., 2017). Interestingly, there was also an interaction between Background Noise condition and Task condition; listening effort was only greater for the n-back condition than for the single-task condition when performed in silence. This result may reflect a ceiling effect of sorts, with the effect of dual-task demands only evident when the listening task was subjectively easy.

Dual-task (n-back) effects on listening task performance

Interestingly, listening task performance was better in the dual-task (n-back) than single-task condition. This result occurred for both babble noise and shaped noise conditions whereas one might have anticipated the opposite result—that listening task performance would suffer under dual-task demands. However, this type of result is not without precedent. A dual-task facilitatory effect has been shown in studies in which the addition of a secondary walking or posture task improves performance on a cognitive task, such as auditory word recognition (Huxhold et al., 2006; Nieborowska et al., 2019; Tomporowski & Audiffren, 2014). This faciliatory effect may owe to the inverted-U arousal function proposed by Yerkes and Dodson (Arent & Landers, 2003). According to this hypothesis, the single-task condition may be associated with lower than optimal arousal, which can negatively affect performance on the listening task. In contrast, the dual-task n-back condition may be associated with an arousal level that is closer to optimal for listening task performance. Indeed, the greater pupil dilation in the dual-task condition than in the single-task condition (see following section) aligns with the idea that the dual-task condition is associated with greater physiological arousal than the single-task condition, an arousal level that could conceivably contribute positively to listening task performance.

Listening task input demands, dual-task (n-back) demands, and the pupillary response

Physiological indices of listening effort, such as the pupillary response, can be used to measure the effect of input demands in listening tasks (Pichora-Fuller et al., 2016). Additionally, physiological indices like the pupillary response have frequently been used to index the cognitive work done in working memory tasks, and more specifically to index the cognitive work done in n-back tasks (Attar et al., 2016; Hjortkjær et al., 2020). Our findings also showed that both listening input demands and n-back task demands affected the pupillary response. In the baseline period, before the onset of the auditory sentence, two critical results were observed. First, pupil size was larger for the babble noise condition than for the shaped noise and silence conditions.

Second, pupil size was larger in the dual-task condition than in the single-task condition. These results are consistent with prior studies in demonstrating that the pupil responds to both tasks' demands.

In the baseline corrected data, the key finding was that peak dilation amplitude was greater for the single-task condition than for the dual-task condition. This finding has at least two interpretations. One possible interpretation is that the dual-task demands associated with the secondary n-back task leave fewer resources available to allocate to target identification in the listening task. As a result, task-evoked pupil dilation is lower for the dual-task condition. This is a particularly noteworthy possibility in light of our interest in the relation between resources tapped by the n-back and listening tasks. However, the dual-task effect on the task-evoked pupil response did not vary across listening task input demands. Thus, these results together paint a somewhat complex picture. Whereas it appears that listening task and n-back task tap similar resources, variation in listening task input demands does not modulate this effect. A second possible interpretation is that greater dilation in the dual-task condition during the baseline period prevents further dilation in response to the listening task stimulus because of physical limits to the amount that the pupil can dilate. If this is the case, then the present set of results cannot be used to evaluate the relation between resources used to meet listening task input demands and resources used to complete the n-back task. By this view, the different baseline pupil dilations for the single-task and dual-task (n-back) condition constitute an obstacle in the present study to determining whether listening task input demands tap the same resources as the n-back task.

Working memory correlates

There was a strong correlation between individual Reading Span and Operation Span scores. However, neither of these measures correlated with n-back IES performance or pupillary response in any condition. N-back task performance or pupil dilation is where I might have expected to see individual differences in cognitive ability play a role; the carryover effects of the listening task were expected to affect n-back performance as a function of participants' available cognitive capacity and to similarly be reflected in pupillary response. A participant with a high working memory span might be expected to have substantial spare capacity, and therefore perform well on the n-back task and have less pupil dilation in response, despite devoting significant resources to listening in noise, processing the sentence stimulus, and identifying the target word.

Conclusions

The results of this experiment provide preliminary support for the theory that the cognitive work associated with listening in noise and the working memory load in an n-back task are similarly taxed when the two tasks are done simultaneously. The strongest evidence from this study that these two tasks tap into similar cognitive resources came from the behavioural measures: primary task input demands affected n-back task performance, and dual-task demands affected listening task performance. Together, these findings provide support for the idea that the working memory resources used in the n-back task are related to the resources tapped by input demands in the listening task.

Though the pupil dilation data did not provide converging support for this idea, a challenge related to interpreting task-evoked pupil dilation in this study must be noted.

Specifically, differences in baseline pupil dilation between the single- and dual-task condition made it impossible to use task-evoked pupil dilation in the listening task to make inferences about the relation between listening task resource use and dual-task resource use. A method in which baseline pupil dilation is equated prior to onset of listening task targets is needed to properly interpret dual-task effects on the task-evoked pupillary response to the listening task, and ultimately to make strong inferences about overlap between resource use in dual-task and listening task contexts.

Of note, the timing of the two tasks—specifically, requiring the response to the n-back task two and a half seconds after the listening task target word—may have played a role in there being very little interaction between the two tasks, especially in the pupillary response. Despite the n-back task requiring continuous updating and maintenance, thus constantly engaging working memory resources, the response selection and actual comparison of the current number to the one from two trials earlier did not occur temporally close to the listening task. Therefore, the work being done to maintain the numbers in working memory, while overlapping with the listening task, may not have been cognitively taxing enough to interact with target word identification, outside of the faciliatory effect produced by assumed increased arousal and attention allocation. This pattern likely holds for the pupillary response as well, where each task appears to individually affect pupil dilation, but the listening task and the n-back task do not interact. An interesting subject to investigate further would be to see how tasks interact when the response selection and cognitive strain of both tasks temporally overlap and how these interactions are captured by the pupillary response.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International journal of audiology*, *47*(sup2), S53-S71.

Anderson-Gosselin, P. A., & Gagné , J. P. (2011). Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. International Journal of Audiology, 50(11), 786–792.

Arent, S. M., & Landers, D. M. (2003). Arousal, anxiety, and performance: A reexamination of the inverted-U hypothesis. *Research quarterly for exercise and sport*, *74*(4), 436-444.

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scandinavian journal of psychology*, *50*(5), 371-384.

Attar, N., Schneps, M. H., & Pomplun, M. (2016). Working memory load predicts visual search efficiency: Evidence from a novel pupillary response paradigm. *Memory & cognition*, *44*(7), 1038-1049.

Auditec of St. Louis. (1971). Four-talker babble. 2515 S. Big Bend Boulevard, St. Louis, MO, 63143–2105.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, *91*(2), 276.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of psychophysiology*, *2*(142-162).

Bennett, K. O. C., Billings, C. J., Molis, M. R., & Leek, M. R. (2012). Neural encoding and perception of speech signals in informational masking. *Ear and Hearing*, *32*(2), 231.

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, *30*(7), 1086-1095.

Conway, A. R. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, *4*(6), 577-590.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why?. *Current directions in psychological science*, *19*(1), 51-57.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Desjardins, J. L., & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. Ear and Hearing, 34(3), 261–272.

Desjardins, J. L., & Doherty, K. A. (2014). The effect of hear- ing aid noise reduction on listening effort in hearing- impaired adults. Ear and Hearing, 35(6), 600–610.

Engle, R. W. (2002). Working memory capacity as executive attention. Current directions in psychological science, 11(1), 19-23.

Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Current anthropology*, *51*(S1), S17-S26.

Ewing, K., & Fairclough, S. (2010). The impact of working memory load on psychophysiological measures of mental effort and motivational disposition. *Human Factors: A system view of human, technology and organisation. Maastricht: Shaker Publishing*.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. Memory & cognition, 43(2), 226–236.

Gagné, J. P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in hearing*, *21*, 2331216516687287.

Hall III, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, *23*(2), 159-165.

Heinrich, A., Ferguson, M. A., & Mattys, S. L. (2020). Effects of cognitive load on pure-tone audiometry thresholds in younger and older adults. *Ear and hearing*, *41*(4), 907.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, *45*(1), 119-129.

Hjortkjær, J., Märcher-Rørsted, J., Fuglsang, S. A., & Dau, T. (2020). Cortical oscillations and entrainment in speech processing during working memory load. *European Journal of Neuroscience*, *51*(5), 1279-1289.

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. Ear and Hearing, 34(5), 523–534.

Huxhold, O., Li, S. C., Schmiedek, F., & Lindenberger, U. (2006). Dual-tasking postural control: aging and the effects of cognitive demand in conjunction with focus of attention. *Brain research bulletin*, *69*(3), 294-305.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394-412.

Johnson, J., Xu, J., Cox, R., & Pendergraft, P. (2015). A comparison of two methods for measuring listening effort as part of an audiologic test battery. *American journal of audiology*, *24*(3), 419-431.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585.

Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615.

Klink, K. B., Schulte, M., & Meis, M. (2012). Measuring listening effort in the field of audiology—A literature review of methods (part 2). *Zeitschrift für Audiol*, *51*(2), 60-67.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and hearing*, *33*(2), 291-300.

Mathôt, S. (2018). Pupillometry: psychology, physiology, and function. *Journal of Cognition*, *1*(1).

Meegan, D. V., Purc-Stephenson, R., Honsberger, M. J., & Topan, M. (2004). Task analysis complements neuroimaging: an example from working memory research. *Neuroimage*, *21*(3), 1026-1036.

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., De Lissa, P., Graham, P., & Lyxell, B. (2017). Objective assessment of listening effort: Coregistration of pupillometry and EEG. *Trends in hearing*, *21*, 2331216517706396.

Miller, A. L., Gross, M. P., & Unsworth, N. (2019). Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation. *Journal of Memory and Language*, *104*, 25-42.

Ng, E. H., Rudner, M., Lunner, T., & Rönnberg, J. (2015). Noise reduction improves memory for target language speech in competing native but not foreign language speech. Ear and Hearing, 36(1), 82–91.

Nieborowska, V., Lau, S. T., Campos, J., Pichora-Fuller, M. K., Novak, A., & Li, K. Z. (2019). Effects of age on dual-task walking while listening. *Journal of Motor Behavior*, *51*(4), 416-427.

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *The quantitative methods for psychology*, *10*(1), 56-67.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, *25*(1), 46-59.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior research methods*, *47*(4), 1343-1355.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.

Picou, E. M., & Ricketts, T. A. (2014). The effect of changing the secondary task in dual-task paradigms for measuring listening effort. Ear and Hearing, 35(6), 611–622.

Rao & Starkey (n.d.) Speech Reception Threshold Program [computer software]. Starkey Hearing Industries

Redick, T. S., & Lindsey, D. R. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic bulletin & review*, *20*(6), 1102-1113.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*, *7*, 31.

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology*, *52*(10), 1293-1304.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*, *18*(07), 559-572.

Sibley, C., Coyne, J., & Baldwin, C. (2011, September). Pupil dilation as an index of learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 237-241). Sage CA: Los Angeles, CA: SAGE Publications.

Tomporowski, P. D., & Audiffren, M. (2014). Dual-task performance in young and older adults: speed-accuracy tradeoffs in choice responding while treadmill walking. *Journal of aging And physical activity*, *22*(4), 557-563.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757-765.

Van Gerven, P. W., Paas, F., Van Merriënboer, J. J., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167-174.

Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., & Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hearing research*, *369*, 67-78.

Wilhelm, O., Hildebrandt, A. H., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it?. *Frontiers in psychology*, *4*, 433.

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in hearing*, *22*, 2331216518800869.

Wu, Y. H., Stangl, E., Zhang, X., Perkins, J., & Eilers, E. (2016). Psychometric functions of dual-task paradigms for measuring listening effort. Ear and Hearing, 37(6), 660–670.

Xia, J., Nooraei, N., Kalluri, S., & Edwards, B. (2015). Spatial release of cognitive load measured in a dual-task paradigm in normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America, 137(4), 1888–1898.

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*, *101*, 76-86.

Zekveld, A. A., Kramer, S. E., Rönnberg, J., & Rudner, M. (2019). In a concurrent memory and auditory perception task, the pupil dilation response is more sensitive to memory load than to auditory stimulus characteristics. *Ear and hearing*, *40*(2), 272.

Zhivomirov, H. (2018). A method for colored noise generation. *Romanian journal of acoustics and vibration*, *15*(1), 14-19.

APPENDIX A
Working Memory Correlations

Table 1
Working Memory Task Correlations with Secondary Task Inverse Efficiency Scores

	Babble	**Shaped Noise**	**Silence**
Operation Span	.017	.133	.054
Reading Span	-.070	-.011	-.146

Table 2
Working Memory Task Correlations with Peak Pupillary Dilation Amplitude

	Babble		**Shaped Noise**		**Silence**	
	Dual-Task	**Single-Task**	**Dual-Task**	**Single-Task**	**Dual-Task**	**Single-Task**
Operation Span	-.159	.067	.040	-.047	-.168	-.208
Reading Span	-.178	.006	.062	-.258	.070	-.310

Table 3
Working Memory Task Correlations with Absolute Pupillary Dilation Size in the Baseline Period

	Babble		**Shaped Noise**		**Silence**	
	Dual-Task	**Single-Task**	**Dual-Task**	**Single-Task**	**Dual-Task**	**Single-Task**
Operation Span	.004	-.178	.019	-.091	.033	-.012
Reading Span	.125	-.048	.110	.080	.097	.105

CHAPTER 3: LISTENING IN NOISE AND THE PSYCHOLOGICAL REFRACTORY PERIOD: PUPIL DILATION AS A MEASURE OF ATTENTION ALLOCATION

Introduction

Perceiving the world around us often feels effortless. Imagine sitting alone in a subway car while texting a friend you are meeting up with about where to grab some lunch; it is quiet on the subway and therefore easy to hear the station announcements so you do not miss your stop, despite holding a conversation over text at the same time. However, as perceptual demands increase, so too does the need for attention, and the feeling that effort is required to perceive accurately. Now imagine the subway car starts to fill with people as you get closer to your destination; there is a lot of people-watching to do and conversations happening around you, and you are still texting your friend and making food-related decisions, making it harder to attend to the station announcements for your stop. The idea that subjective effort accompanies the allocation of limited capacity attentional resources has a long history in experimental psychology (Kahneman, 1973). There is a lot going on in that subway car and you are trying to hold a conversation over text, so you now have to put in the effort to attend to the station announcements, as you have limited ability to pay attention to everything in your environment; you must prioritize some information over other information so you do not miss your stop.

The allocation of limited capacity attentional resources is a central construct in many frameworks aimed at the study of attention. Kahneman (1973) proposed that we have limited resources available to us that can be applied to the demands in our environment and that this limitation dictates how we allocate attention. Two particular areas of research are the focus of the present study: (1) Listening to speech in noise benefits from the allocation of attentional resources (Mattys et al., 2012); and (2) Performance on two tasks suffers because limited attention cannot be allocated to both tasks simultaneously (Pashler, 1994). Here I combine common methods of these two research areas to examine the concept of limited attentional resources. Brief summaries of these two research domains are provided below, followed by an overview of the research strategy that motivated my empirical work.

Attentional Resource Allocation: Listening to Speech in Noise

Struggling to properly hear a conversation in a loud environment is a common experience. Listening to speech in noise requires effort. The effort required to listen in noise relative to listening in quiet can be measured in a variety of ways. One such method is to examine costs on performance in an unrelated secondary task that occurs with the listening task (Gagné et al., 2017), such as a tactile or visual discrimination task, or a continuous visuomotor tracking task. A second method is to use a physiological index, such as pupil dilation (Mathôt, 2018). Both of these measures can be used to determine how the recruitment of additional cognitive resources contributes to listening in effortful situations. The field of Cognitive Hearing Science (CHS) seeks to understand how the recruitment of cognitive resources interacts with the physical sound components of speech perception, and to measure this interaction objectively in a variety of ways (Arlinger et al., 2009). This chapter examines this interaction by measuring pupillary responses to speech in noise under dual-task conditions.

The idea that central cognitive resources affect hearing ability and speech perception is well established (Davis & Johnsrude, 2007; Francis & Nusbaum, 2009; Mitterer & Mattys, 2017; Murphy et al., 2000; Piquado et al., 2010; Wingfield et al., 1994; Wingfield & Tun, 2007). Such effects are commonly experienced in our everyday listening environments – we listen while our attention is divided with other tasks and find that divided attention makes listening difficult.

Clearly, hearing the world around us is not simply a sensory task. Rather, higher order cognitive abilities play an important role in navigating our auditory world. Akeroyd (2008) reviewed twenty studies that examined speech perception of individuals with normal and impaired hearing and found that speech perception is indeed related to an individual's cognitive ability: those with greater cognitive ability were better able to recognize speech in noise. Moreover, there is evidence that this relation between hearing and cognition may be bi-directional; mice with induced hearing loss early in life performed worse on cognitive tasks than those that maintained normal hearing (Park et al., 2016). Together, these studies illustrate that cognition affects hearing and hearing affects cognition.

Rönnberg et al. (2013) proposed the Ease of Language Understanding (ELU) model to explain findings in this area. This model describes how cognitive resources are utilized in difficult listening conditions. The model includes a process by which an incoming auditory signal is compared to an expected phonology and cognitive resources associated with working memory (WM) are enlisted to help when that comparison results in a mismatch. These cognitive resources are critical in interpreting the signal, as well as for further processing of the signal. The general idea that listening effort is related to cognitive ability ties together studies of listening in noise with those focusing on individual differences in cognitive ability (Daneman & Carpenter, 1980; Engle et al., 1999). These individual differences have been studied for decades and hold potential utility to understand not only higher order cognitive performance but also the seemingly more perceptual task of hearing under difficult listening conditions. As memory load, a concurrent task, or signal degradation are introduced to a hearing task, an individual's ability to hear correctly decreases. A way to understand such effects is to assume that cognitive resources are used to hold onto items in memory, to attend to a concurrent task, and to pick out a signal presented in noise. Consequently, assuming humans have a finite capacity of cognitive resources, the use of some of that finite capacity on other cognitive activities limits what is available for simultaneous hearing under difficult listening conditions. The ELU model, therefore, explains listening ease in terms of an incoming signal and the cognitive resources an individual has available to allocate to the processing of that signal (Wingfield et al., 2015). As a consequence of the interaction between the signal and an individual's cognitive ability, higher working memory capacity is associated with better performance and less subjective effort in listening tasks. A recent revisit of the model has found it to support a variety of hypotheses regarding the link between working memory capacity and listening effort (Rönnberg et al., 2019). Cognitive resources and their use play a role in listening and affect an individual's level of listening effort.

The Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016) aims to understand listening effort from the perspective of incoming factors (e.g., background noise or the amount of semantic information in an auditory stimulus) and outgoing measures (e.g., physiological response). It looks at the various types of difficult listening conditions that can affect listening effort as well as a variety of ways to measure that effort. The framework takes into account inputs such as the listening condition, context, and individual cognitive abilities, and examines how these inputs interact with levels of arousal, attention, and other cognitive factors. These interactions are then behaviourally and physiologically measured by means of, for example, pupil dilation, EEG alpha power, and task performance measures such as reaction times and dual-task costs. To summarize, the FUEL assumes that numerous factors affect hearing ability and listening effort, and that these various factors interact with the allocation of cognitive resources in listening tasks. Relative to the ELU model, the FUEL is a broader, more all-encompassing framework that considers cognitive ability but also the effect of external factors and other internal factors on listening effort. It also outlines a variety of methods that can be used to measure the amount of effort being expended when listening in noise or in other contexts of effortful listening.

Attentional Resource Allocation: The Pupillary Response

The pupil dilates in response to greater cognitive engagement (Beatty, 1982; Hess & Polt, 1964; Kahneman & Beatty, 1966; Neagu et al., 2019). Mathôt (2018) refers to this measure as the psychosensory pupil response (PPR). Though the pupil does dilate spontaneously in response to external stimuli, pupil size increases systematically in response to increases in cognitive task load (Mathôt, 2018; Unsworth & Robison, 2018). These changes in pupil size have been documented by psychologists since the 1960s, and appear to reflect the body's response to effortful cognitive work. The pupillary response is related to activity in the locus coeruleus, where increased activity is correlated with increased dilation (Laeng et al., 2012). It has also been found that when participants are overloaded with information or given too high a cognitive load, the rate of pupil dilation decreases and the pupils can even begin to constrict (Peavler, 1974). In line with the notion of cognitive overload, an insensitivity of the pupil dilation response may reflect a ceiling level of cognitive engagement, while pupil constriction may reflect disengagement from the task and task-relevant information. The PPR has also been used as an indication of cognitive effort in a variety of divided attention methods, including dichotic listening (Koelewijn et al., 2014), task switching (Katidioti et al., 2014; McCloy et al., 2017), and those used to measure working memory span (Padilla et al., 2020). Overall, the pupillary response has proven to be a versatile tool used to measure cognitive costs and the effort expended in numerous types of tasks.

Attentional Resource Allocation: Doing Two Tasks at Once

A well-studied and fundamental question in the field of attention concerns whether there is a performance cost to doing more than one task at the same time (Welford, 1952; Logan, 1979). Our everyday experience suggests that we can sometimes perform two tasks simultaneously with relatively little cost, perhaps when one of the tasks is well-practiced and can be performed 'automatically'. However, there are many other situations in which simultaneous performance of two tasks clearly produces worse performance on one or both tasks than when either of those tasks are performed on their own. This reduction in performance efficiency when two tasks are performed at the same time is known as a dual-task cost.

Dual-task costs have been studied with many methods, but the method of primary interest here examines carefully the temporal overlap between tasks to identify which processing components of two tasks interfere with each other. Research on this issue dates back more than half a century (Welford, 1952), but has been a particular focus in the field of human cognition since the seminal work of Pashler (1984; for a review, see Pashler, 1994). The method used in these studies presents participants with a primary task followed at varying temporal intervals by a secondary task. These tasks need not be complex—the primary task can be as simple as discriminating whether a tone is low or high in pitch, and the secondary task might require something as simple as word identification. Nonetheless, despite the simple nature of the primary and secondary tasks, many studies have demonstrated that performance on the secondary task is delayed substantially when it is presented temporally close in time to the primary task (e.g., within 250 ms) relative to when it is presented after a longer temporal interval following the primary task. This delay in performance of the secondary task that occurs for short temporal intervals between primary and secondary tasks suggests a form of processing bottleneck; that is, some aspect of processing must be dedicated to the primary task for a period of time, and only when primary task processing is complete can that aspect of processing occur for the secondary task.

The delay in secondary task response when primary and secondary tasks are separated by a short temporal interval is known widely as the psychological refractory period (PRP). Substantial research has been devoted to understanding the precise nature of the PRP. One view of the PRP is that it reflects the allocation of limited capacity attentional resources to processing of the primary task, leaving few residual resources for processing of the secondary task when it overlaps closely in time with the primary task. In contrast, when primary and secondary tasks are spaced apart in time, limited capacity attentional processing of the primary task is already complete upon onset of the secondary task, and consequently there is no shortage of resources to complete the secondary task with optimal efficiency. An alternative view of the PRP is that it reflects a structural limitation rather than a processing limitation. By this view, there is a stage of processing that is necessarily serial—for structural reasons, it can be done for only one task at any given point in time. As a result, when primary and secondary tasks are presented in rapid succession, this stage must be completed first for the primary task before it can be started for the secondary task. If the duration of this processing stage is longer than the temporal interval between tasks, then a bottleneck in processing will result. By comparing short and long intervals between primary and secondary task, one can then estimate the duration of this bottleneck processing stage. Proponents of this view point to the bottleneck as occurring after a stimulus is perceived and categorized, but before a response is executed, at the stage of response selection (McCann & Johnston, 1992; Pashler, 1994).

The Present Research Strategy

Research on listening to speech in noise and research on dual-task costs both focus on the fundamental issue of limited capacity attention allocation. Yet, to my knowledge, no research has combined the methods of these two domains to study them together. In the present study, I combined methods from these two domains to ask how the limited capacity attentional resources that help one decipher speech in noise are related to the capacity limit that produces a bottleneck when we try to do two tasks at once.

The experiments used a dual-task method as in studies of the PRP; that is, a primary task was followed at either a short or long temporal interval by a secondary task. I assumed that when the asynchrony between onset of primary and secondary tasks (stimulus onset asynchrony, or SOA) was short (i.e., 100 ms in my study), then key aspects of processing of primary and secondary tasks would overlap. In this sense, any slowing that occurred for this short SOA condition relative to the long SOA condition can be thought of as a dual-task cost. Most important for my purpose, the primary task was a listening to speech in noise task. The input demands of this task were manipulated through background noise and sentence predictability. In Experiment 1, two kinds of background noise were used: shaped noise and babble noise, where the shaped noise provides signal masking and the babble noise provides signal masking and informational masking, making the babble noise a more demanding condition (Bennett et al., 2012; Hall et el., 2002; Schneider, Li & Daneman, 2007). Additionally, in Experiment 1, there were high predictability and low predictability sentences, where the low predictability leads to greater demands on resources due to there being less semantic information available to the listener (Hunter, 2020; Hunter & Pisoni, 2018; Sheldon et al., 2008). In Experiment 2, the only primary task manipulation was background noise, with the introduction of a silence condition to compare with the shaped noise and babble noise backgrounds. In line with the literature on listening to speech in noise, manipulation of primary listening task difficulty was assumed to affect the cognitive resources devoted to interpreting the auditory signal. Therefore, any input demands on the listening task, such as background noise or manipulations to the semantic information in the stimulus, would affect the amount of resources needed to identify speech presented in noise. A key research question concerned how the allocation of resources to the auditory signal in the listening task would affect performance in the secondary task, meaning that the demand on resources in the primary task determines how resources can be allocated in the secondary task. The secondary task was visual, a simple discrimination task: color discrimination for a target square in Experiment 1, and odd/even parity for a target number in Experiment 2.

If the limited capacity attentional resources that aid listening in noise contribute to the capacity limit associated with doing two tasks at once, then a number of straightforward predictions can be made. First, and most simply, it should be possible to measure a PRP effect using identification of speech in noise as the primary task. That is, performance should be worse for the short SOA condition than for the long SOA condition of the secondary task. Second, and more importantly, this PRP effect—defined as worse performance when two stimuli are presented at a short SOA compared to when they are presented at a longer SOA—should depend on the input demands of the listening in noise task, being larger when the input demands of the listening task are high (e.g., a babble noise background) than when those input demands are low (e.g., silence). This prediction follows from the view that a more demanding listening task results in more resources being allocated to complete that task, leaving fewer residual resources for the secondary task when the two tasks overlap in time. Input demands of the listening task were manipulated in both Experiments 1 and 2 by varying the semantic predictability of the listening task target (Experiment 1), and the nature of noise that accompanied the entire speech signal (babble noise versus shaped noise in Experiment 1; babble noise versus shaped noise, both compared to silence in Experiment 2).

In addition to using secondary task performance as an indication of limited capacity resource allocation, I used the pupillary response as another objective measure of resource allocation and effort. Past research has often used mean peak dilation latency and mean peak dilation amplitude as measures of interest (Zekveld et al., 2010). These measures of when the pupil reaches its maximum dilation and how much the pupil dilates allow comparisons of the amount and timing of cognitive effort expended throughout the task across experimental conditions of interest. Of particular relevance here, pupil dilation has been employed in prior studies of effortful listening. Cognitive effort and allocation of limited capacity resources should increase when listening in noise and, in accord with this idea, pupil dilation has been shown to vary in response to listening difficulty (Miles et al., 2017; Pichora-Fuller et al., 2016; Winn et al., 2018).

Returning to the key predictions in the present study, to my knowledge, no prior study has examined pupil dilation in the context of a PRP dual-task method. Again, if limited capacity attentional resources that aid listening in noise contribute to the capacity limit associated with doing two tasks at once, and if this resource allocation is captured in measures of pupil dilation, then three key predictions follow. First, pupil dilation patterns should reflect the input demands of the primary listening in noise task. Specifically, pupil dilation should vary directly with the listening in noise task input demands, meaning greater demands lead to increased pupil dilation. Second, it should be possible to measure a pupil dilation effect that aligns with the behavioural PRP effect, with task-evoked pupil dilation being greater for the short SOA condition than for the long SOA condition. This prediction follows from the view that maximum resource allocation and effort is likely to occur when resources are demanded by both tasks simultaneously. Third, under conditions in which the behavioural PRP effect is larger for higher listening task input demands, pupil dilation also ought to be correspondingly large. In other words, pupil dilation should be particularly pronounced when the primary and secondary tasks overlap substantially (i.e., in the short SOA condition), and when the primary listening task input demands are highest.

An additional straightforward way to measure subjective listening effort was also adopted here. I simply asked the participant to estimate their listening effort. I used Johnson et al.'s (2015) 7-point listening effort scale to measure subjective listening effort. This subjective measure can be used as a simple check that my task conditions vary in perceived difficulty as I expected and can also provide insight into how well subjective and objective measures of effort correspond.

Altogether, both experiments aimed to combine a listening in noise task with a PRP-style secondary task to determine the nature of attention allocation when task demands overlap closely in time and when the task demands are more temporally spread out. Experiment 1 manipulates two factors of listening task input demands (background noise and sentence predictability), while Experiment 2 instead focuses on background noise as the manipulation of interest and introduces a silence condition with which to compare the effects of background noise. Both experiments use a visual secondary task and measure pupil dilation as an indicator of cognitive load.

Experiment 1

This experiment employed a PRP method that combined a listening to speech in noise primary task with a simple colour discrimination secondary task. The secondary task target appeared either 100 ms or 2500 ms after the primary task target. The broad objective was to evaluate the relation between limited capacity resources used in service of understanding speech in noise, and capacity limits that produce the PRP. Would high input demands in the listening task (i.e., babble noise or low sentence predictability) modulate the PRP effect in the secondary task such that greater listening demands affect secondary task performance more at the short SOA than at the long SOA? And would this interaction between primary task demands and secondary task timing be reflected in pupil dilation measures, such that the pupil would respond more when the primary and secondary tasks overlap more closely temporally?

Methods

Participants

A total of 34 undergraduate students participated in this experiment for partial course credit. Data from one participant were excluded from analysis because of technical difficulties, and data from another participant were excluded from analysis because they learned English later than my cut-off age of 7 years. Data from the remaining 32 undergraduate students (25 females and 7 males; mean age = 20.2 years) were subject to analysis. All participants had self-reported normal or corrected-to-normal vision and self-reported normal hearing.

Apparatus and Stimuli

The visual stimuli used in the secondary colour identification task were presented on a Dell UltraSharp 2001FP LCD 17-inch monitor running from an HP Pro 3130 MT computer, using Presentation® software. The stimuli for the secondary task were coloured squares presented in the centre of the screen, measuring 2cm x 2cm on the screen, in one of four colours: red (255, 0, 0); blue (0, 127, 255); green (0, 255, 0); or yellow (255, 255, 0). Participants were seated approximately 60 cm from the computer monitor and used the space bar, "z", and "/" keys on a standard keyboard to respond to the secondary task by indicating whether the presented coloured square was a warm colour (red or yellow) or a cool colour (blue or green).

The auditory stimuli were presented over a Mackie CR3 Creative Reference speaker running from the same HP Pro 3130 MT computer. The speaker was positioned on a shelf right above the monitor, facing the participant. The volume was set such that the target stimuli were presented at 70 dB and the background noise was set 3 dB louder at 73 dB. The stimuli were the Revised Speech Perception in Noise Test (R-SPiN) sentences, which include 200 high predictability sentences and 200 low predictability sentences, where predictability refers to how predictable the final word is based on the rest of the sentence. Pairs of low and high predictability sentences shared their final target word, so there was a total of 200 target words in the stimulus set. The background noise was either shaped noise, created in MATLab using code from Zhivomirov (2013), or a four-talker babble track produced by Auditec (1971). Participants' verbal responses to the sentences presented in noise were manually recorded by the experimenter into an Excel spreadsheet in real-time.

Participants also completed two working memory span tasks before completing the listening task: the Operation Span and Reading Span, provided with permission by the Attention and Working Memory Lab at Georgia Institute of Technology (Foster et al., 2015; Oswald et al., 2015). These two tasks were used to obtain a measure of participants' working memory spans. These tasks were presented in E-Prime from the same computer and monitor as the visual stimuli and experiment, and were fully automated with all instructions on-screen. The rationale for measuring working memory span is that it has been implicated as the cognitive mechanism at work during response selection (Ellenbogen & Meiran, 2008), identified as a potential processing locus of PRP effects. Those with better performance on working memory tasks also have better secondary task performance when identifying a stimulus in noise (Desjardins & Doherty, 2013) and therefore I analyzed the correlations of these two measures with secondary task performance and pupillary response.

Pupil dilation was recorded by a Model ET1000 EyeTribe eye tracker that was placed right beneath the monitor, turned up toward the participant. The EyeTribe software was used to calibrate and record the pupil data, and was synchronized with the experiment data from Presentation®. Trigger markers from Presentation® were compiled with time-stamped sampling data from the EyeTribe in LabStreamingLayer, and output as .xdf files that were used for later analysis. The data were collected at a rate of 30Hz.

Design

All participants completed the experiment in the following order: Operation Span, Reading Span, and the main task which involved listening to and identifying speech in noise and performing a secondary visual task. The listening task consisted of 12 blocks of 16 auditory trials, for a total of 192 trials. Background Condition (shaped noise/babble noise) varied between blocks, alternating between shaped noise and babble noise regularly from one block to the next in a counterbalanced manner across participants. Sentence Predictability (high/low) and

Secondary Task SOA (100 ms/2500 ms) varied randomly within the blocks of 16 trials, meaning that there was an approximately equal number of each of the four conditions across the experiment as a whole. There were four lists of sentences, which were counterbalanced across background conditions and participants. Subjective listening effort was rated after each block—thereby obtaining ratings for the two background noise conditions, which was the only blocked condition—using Johnson et al.'s (2015) listening effort scale, which is a Likert scale that ranges from 1 (no effort) to 7 (extreme effort). When prompting the participant to rate their listening effort, the experimenter specifically asked for listening effort, and informed the participant that this was not a rating of overall effort but just of how effortful it was to hear and identify the words.

Procedure

Operation Span

The Operation Span task (Foster et al., 2015; Oswald et al., 2015) presented participants with math equations (e.g., 8 -3 = 5) and the participants indicated, using a right or left mouse click, if the equation was true or false. Beneath the equation, a letter was presented and participants were instructed to remember the letter. After two to seven trials of this task, a matrix of letters appeared on the screen and the participant had to click the letters in the order in which they appeared under the equations. There were 15 blocks of these trials.

Reading Span

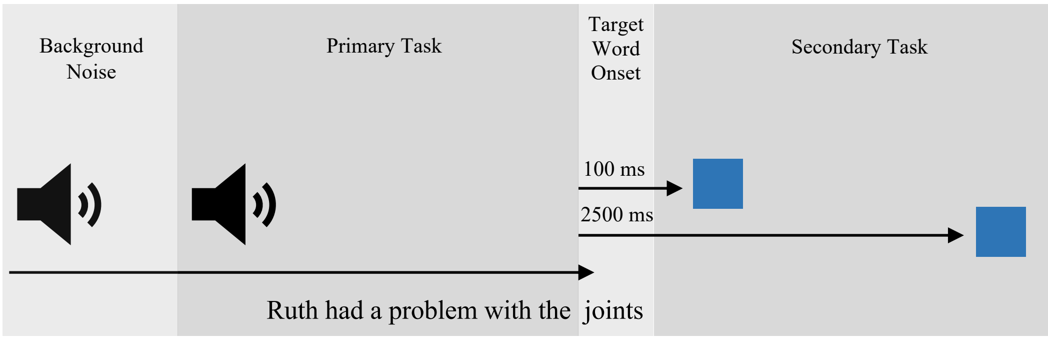
The Reading Span task (Foster et al., 2015; Oswald et al., 2015) was the same as the Operation Span task, except instead of equations, participants were shown sentences (e.g., The sky was full of spaghetti) and asked to indicate whether the sentence made sense semantically or not. Once again, each trial presented a letter that the participant had to remember and recall at the end of the block. There were six blocks of these trials.

Listening Task

Figure 1 depicts the trial sequence in the listening task. Before completing the listening task of the experiment, the eye tracker was calibrated to each participant's eyes using the EyeTribe software. Each of the 192 trials began with the background noise coming on one second before the start of the sentence, and thus the sentence always overlapped temporally with the noise. The predictability of the sentence was randomized across trials. Both high- and low-predictability sentences ending with a particular target word were presented, and participants were instructed to repeat aloud that final word before completing the secondary task. The experimenter then recorded the participant's spoken response. Following each sentence, the secondary task (colour discrimination task) stimulus appeared on the screen; one of four coloured squares either 100 ms or 2500 ms after onset of the target word from the preceding sentence. Participants were asked to make one of two responses to each of the four possible secondary task stimuli. The coloured squares were grouped as 'warm' and 'cool' colours; red and yellow were designated warm, and blue and green were designated cool. Participants responded with one key (either the '/' or 'z' key) for yellow or red, and the opposite key for blue or green; these response mappings were counterbalanced. After each block of 16 trials, participants gave their rating of listening effort.

Figure 1

Experiment 1 Listening Task Trial Sequence



This figure depicts the sequence of events throughout the trial, with the background noise followed by the sentence stimulus, and then the secondary colour discrimination task either 100 ms or 2500 ms after the onset of the target word at the end of the sentence.

Pupil Data Processing

Pupil data were processed in Matlab and the post-processing data were exported to .csv files to be used in further analysis. Trial markers collected from Presentation® experiment output were timestamp matched to pupil data samples and used to integrate behavioural data (performance on the listening task) with the pupil data. Only trials on which participants correctly identified the target word were included in the pupil analyses. The pupil data were processed as separate left and right eye channels until they were combined and averaged at the end of processing. Participants with one or more conditions that had more than two-thirds of the data samples missing were excluded from further processing, resulting in inclusion of 23 out of 32 participants' data in analyses of pupillary response. For the included participants, missing data points were filled in using linear interpolation between existing data points. A three-sample moving average filter was then used to smooth the data. No gaze correction algorithm was applied to the data because participants used a chin rest and maintained their gaze at the fixation point. Baseline pupil size was calculated as the average measured pupil size from -2500 ms to 0 ms before the onset of the sentence. This baseline value was subtracted from each pupil size sampled and then used to determine the proportion change in pupil size from baseline to the pupil's size over the course of the trial. The trial length was set to 7500 ms, with the onset of the start of the sentence marking the start of a trial (the 0 ms mark).

Data Analysis

The dependent variables analyzed included the proportion of primary task target words correctly identified and secondary task response times (RT) and accuracy. Inverse Efficiency Scores (IES; mean RT divided by proportion correct) were also calculated for secondary task performance. IES can be used as a wholistic measure of task performance by taking into account accuracies and response times together but should be examined in addition to accuracy and response time performance, rather than on its own (Bruyer & Brysbaert, 2011). Each of these dependent variables was analyzed using a repeated measures ANOVA that treated Background Condition (shaped noise/babble noise), Sentence Predictability (high/low), and Secondary Task SOA (100 ms/2500 ms) as factors. Subjective listening effort measured after each block was analyzed using a one-way ANOVA that treated Background Condition (shaped noise/babble noise) as its factor[1].

Reading Span and Operation Span results were scored using the partial-credit unit score method (Conway et al., 2005). The correlation between these scores and secondary task IES was computed. The rationale underlying this analysis is that secondary task performance may reflect residual resources after completion of the primary task. Therefore, those with fewer cognitive resources may show greater costs in secondary task performance, leading to a positive correlation between working memory capacity and secondary task performance. The correlation

[1] This ANOVA is mathematically equivalent to a matched sample t-test; it is labelled an ANOVA here for the sake of consistency with other analyses in this experiment and in Experiment 2.

between these scores and two measures of pupillary response was also computed. The two measures were absolute pupil size in the baseline period and mean peak dilation amplitude (as detailed below). The rationale for this analysis is similar to that of the IES correlations; the pupillary response is a measure of cognitive work and therefore those with fewer cognitive resources are expected to require more effort to complete the tasks than those with greater available resources, as reflected by greater absolute pupil size in response to background noise and greater peak pupil dilation in response to the listening task.

The raw pupil data were used to generate three different measures. The first measure was the mean absolute pupil size in the baseline period from 2500 ms prior to the onset of the sentence (i.e., -2500 ms) to the time of onset of the sentence (i.e., 0 ms). This absolute dilation in the baseline period is presumed to measure the demands on cognitive resources before the auditory stimuli were presented and encompasses the one second in which the background noise is presented prior to onset of the sentence. The other two measures were labelled 'baseline corrected' pupil measures; that is, they measured properties of pupil size after subtraction of mean pupil size from the baseline period. The two baseline corrected measures of pupil size were: (1) peak dilation latency and (2) peak dilation amplitude. Peak dilation latency was defined as the timepoint within each trial at which the pupil reached its maximum size. This value was then used to compute the mean peak dilation latency in each condition for each participant. Peak dilation latency is presumed to reflect the relative point in time at which the pupil is maximally responding to the presented stimuli. Peak dilation amplitude was measured as the mean pupil size (relative to baseline) in the 500 ms time window centered on the peak dilation latency (plus and minus 250 ms, rounded to the nearest 33.33 ms due to the sampling rate of 30 Hz) for each condition, averaged across participants. The time windows used for peak dilation amplitude are detailed in Table 1A. Peak dilation amplitude is presumed to reflect the amount of cognitive work being done at the point at which cognitive work is at its highest. All pupil measures were also submitted to repeated measures ANOVAs that treated Background Condition (shaped noise/babble noise), Sentence Predictability (high/low) and Secondary Task SOA (100 ms/2500 ms) as factors.

For the behavioural data, Figures 2 to 5 include error bars with the Morey correction applied (O'Brien & Cousineau, 2014), which accounts for within-subjects designs when calculating the standard error of the mean.

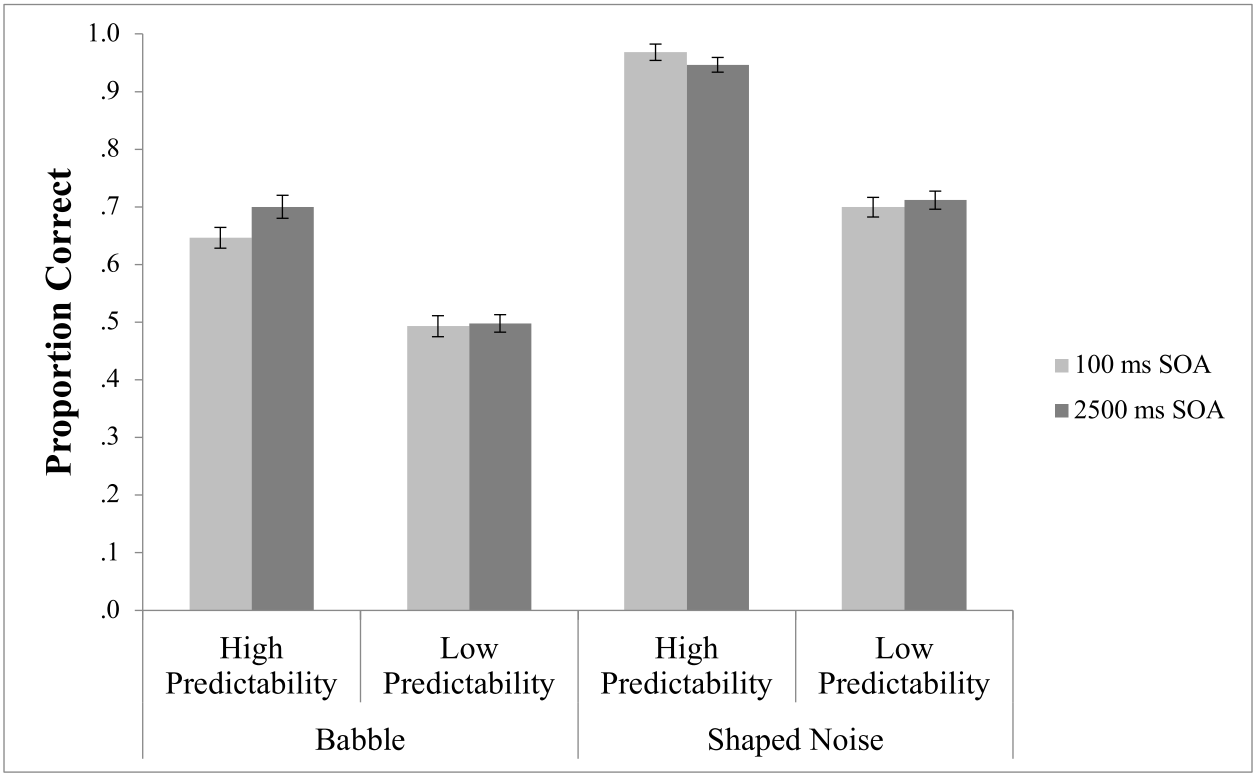
Results

Listening Task: Target Word Identification Accuracy

Figure 2 shows the proportion of correctly identified target words by condition. There was a main effect of Background Condition, $F(1, 31) = 252.2$, $p < .001$, $\eta p^2 = .89$. Target word accuracy was higher for shaped noise (.83) than for babble noise (.58). There was also a main effect of Sentence Predictability, $F(1,31) = 446.5$, $p < .001$, $\eta p^2 = .94$. Target word accuracy was higher for high (.82) than low (.60) sentence predictability. There was also a significant interaction between Background Noise and Sentence Predictability, $F(1,31) = 9.67$, $p < .01$, $\eta p^2 = .24$; this effect indicates that the Sentence Predictability effect was larger for shaped noise (.25) than for babble (.18). There was also a significant three-way interaction between Background Condition, Sentence Predictability, and Secondary Task SOA, $F(1,31) = 4.63$, $p = .04$, $\eta p^2 = .13$. Post-hoc t-tests found that there were Secondary Task SOA effects for high sentence predictability for both babble noise, $t(31) = 2.1$, $p = .022$ and shaped noise, $t(31) = 1.77$, $p = .043$. Importantly, these two effects were opposite in direction, with better performance for the 2500 ms SOA (.70) than the 100 ms SOA (.65) in the babble noise, and better performance for the 100 ms SOA (.97) than for the 2500 ms SOA (.95) in the shaped noise. These opposite effects of SOA for the two noise conditions appear to be the source of the significant three-way interaction. Neither difference between Secondary Task SOAs in the low sentence predictability conditions was significant, both p values > .30. These results confirm that there were large and reliable effects in behaviour of the two input demand manipulations in the listening to speech in noise task.

Figure 2

Word Identification Accuracy



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

Secondary Task: Colour Discrimination Accuracy, Reaction Time, and Inverse Efficiency Score

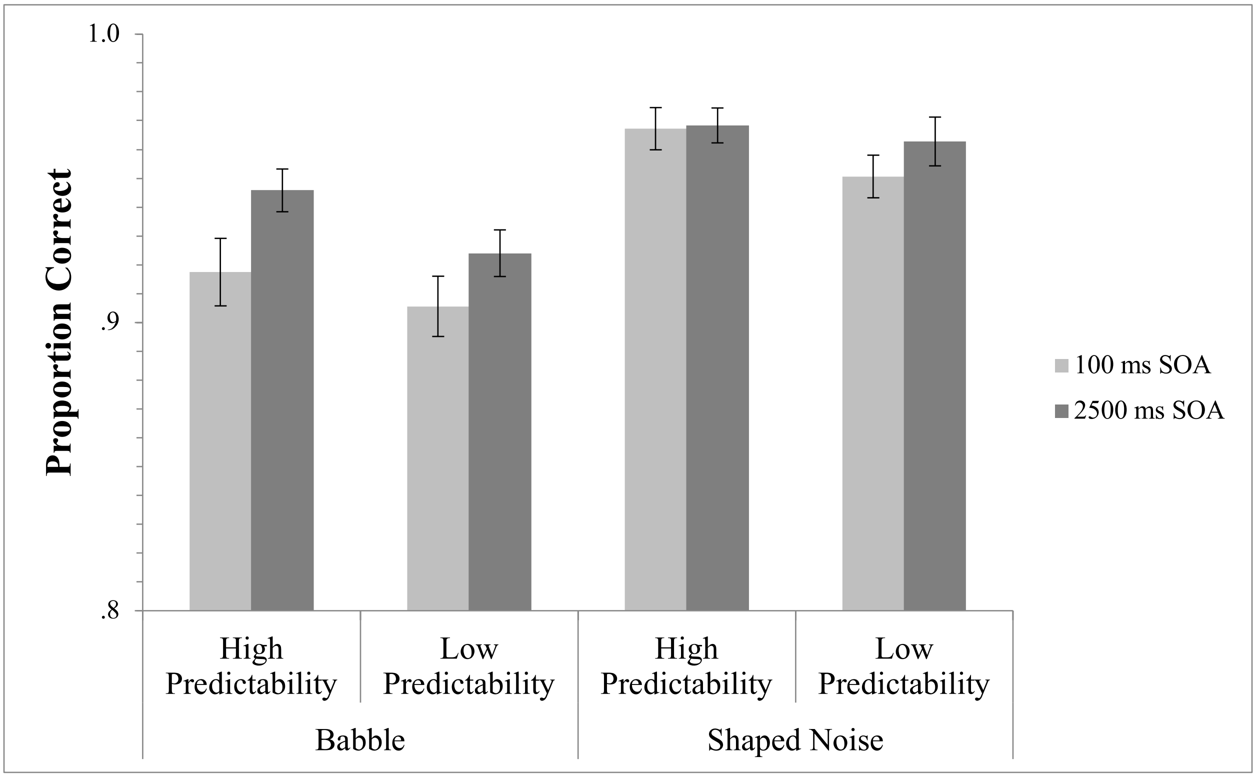
Figures 3 and 4 display proportion correct and mean RT, respectively, for the secondary task.

Accuracy

In the analysis of secondary task accuracy, there was a main effect of Background Condition, $F(1,31) = 19.8$, $p < .001$, $\eta p^2 = .39$; accuracy was higher for secondary task performance following shaped noise (.96) than for babble noise (.92). There was also a significant main effect of Sentence Predictability, $F(1,31) = 6.5$, $p < .05$, $\eta p^2 = .17$; accuracy was higher on the secondary task following the high predictability sentences (.95) than for the low predictability sentences (.93). And there was a main effect of Secondary Task SOA, $F(1,31) = 7.2$, $p < .05$, $\eta p^2 = .19$; accuracy was higher when the secondary task followed at the 2500 ms SOA (.95) than for the 100 ms SOA (.94). Neither Background Condition nor Sentence Predictability interacted with Secondary task SOA, both F values < 1.0.

Figure 3

Secondary Task Accuracy



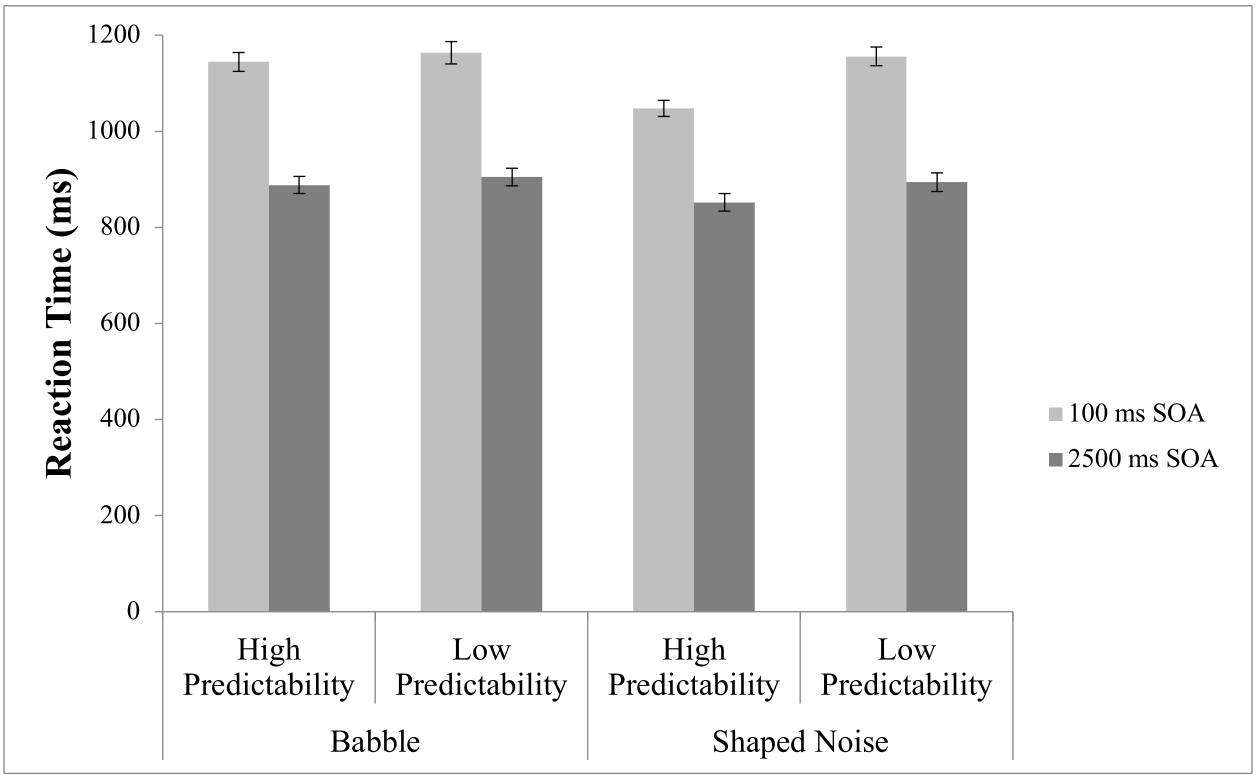
Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

Reaction Time

In the analysis of mean RT, there was a main effect of Sentence Predictability, $F(1,31) = 23.8$, $p < .001$, $\eta p^2 = .43$; RTs in secondary task performance following the high predictability sentences (983 ms) were faster than following the low predictability sentences (1029 ms). The main effect of SOA was also significant, $F(1,31) = 149.5$, $p < .001$, $\eta p^2 = .83$, with faster secondary task responses for the 2500 ms SOA (885 ms) than for the 100 ms SOA (1128 ms). There was also a significant interaction between Background Condition and Sentence Predictability, $F(1,31) = 10.12$, $p < .01$, $\eta p^2 = .25$. In the shaped noise condition, secondary task RTs were faster following high predictability (950 ms) than low predictability (1025 ms) sentences, $p < .001$. In the babble noise condition, secondary task RTs did not differ following the high predictability (1016 ms) and low predictability sentences (1034 ms), $p = .53$. Again, neither Background Condition nor Sentence Predictability interacted with Secondary Task SOA, both $F < 1.0$.

Figure 4

Secondary Task Reaction Time



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

Inverse Efficiency Scores

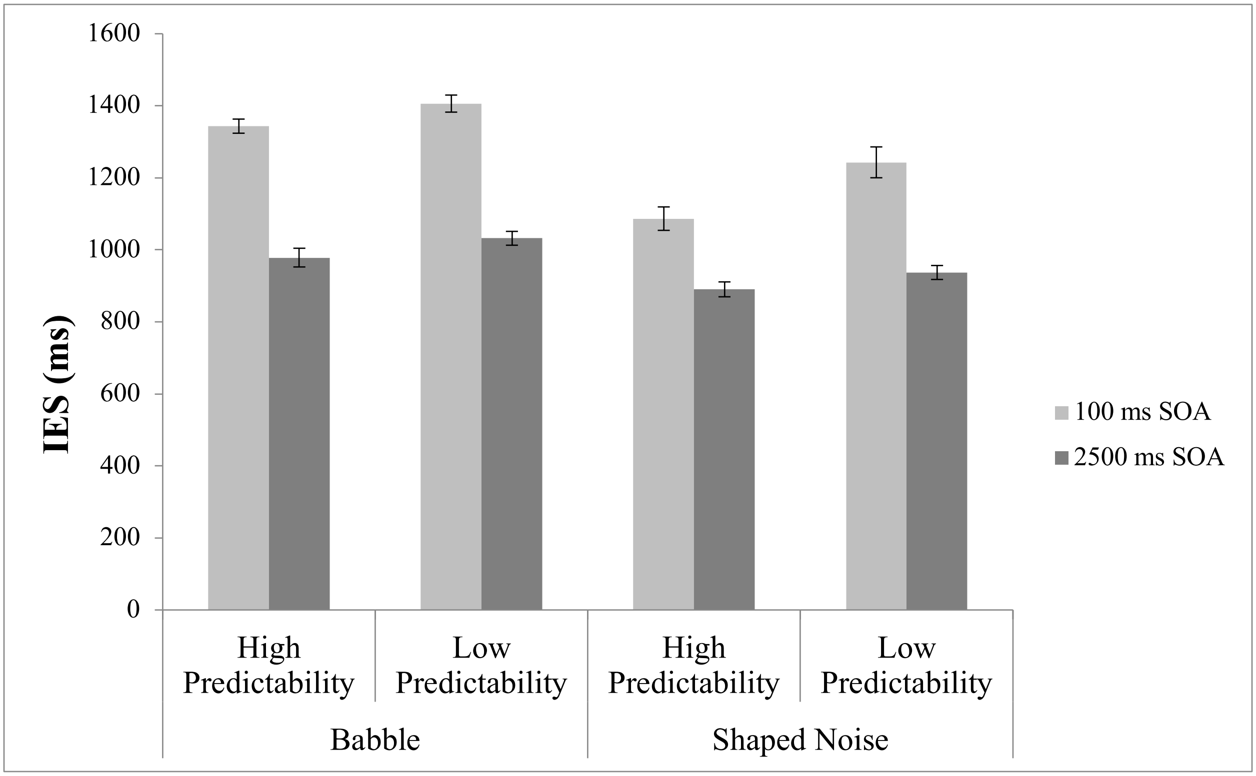
Figure 5 displays inverse efficiency scores (IES) for the secondary task. There was a main effect of Background Condition, $F(1,31) = 34.8$, $p < .001$, $\eta p^2 = .53$; secondary task IES was higher following babble noise (1190 ms) than following shaped noise (1039 ms). There was also a main effect of Sentence Predictability, $F(1,31) = 22.7$, $p <. 001$, $\eta p^2 = .42$; secondary task IES was higher following low predictability sentences (1155 ms) than following high predictability sentences (1075 ms). Finally, there was a main effect of Secondary Task SOA, $F(1,31) = 121.4$, $p < .001$, $\eta p^2 = .80$; secondary task IES was higher for the 100 ms SOA (1270 ms) than the 2500 ms SOA (960 ms).

Of particular importance, there were also two significant two-way interactions in the analysis of IES. First, the interaction between Background Condition and Secondary Task SOA was significant, $F(1,31) = 13.6$, $p < .001$, $\eta p^2 = .30$. Secondary task IES was higher for the 100 ms SOA than for the 2500 ms SOA following both babble noise (1375 vs 1005 ms; $p < .001$) and shaped noise (1165 vs 914 ms; $p < .001$), but the magnitude of this effect was greater for babble noise[2]. Second, the interaction between Sentence Predictability and Secondary Task SOA was significant, $F(1,31) = 4.5$, $p = .04$ $\eta p^2 = .13$. Secondary Task IES was higher for the 100 ms SOA than for the 2500 ms SOA following both high predictability sentences (1215 vs 934 ms; $p < .001$) and low predictability sentences (1324 vs 985 ms; $p < .001$), but the magnitude of this effect was greater for low predictability sentences.

[2] Stated differently, I see that the difference in IES between babble noise and shaped noise was greater at the 100 ms SOA than at the 2500 ms SOA. For the 100 ms SOA, babble noise versus shaped noise (1375 vs. 1165 ms/proportion correct; $p < .001$), and for the 2500 ms SOA, babble noise versus shaped noise (1005 vs 914 ms/proportion correct; $p < .05$).

Figure 5

Inverse Efficiency Scores (IES)



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

These analyses of secondary task performance reveal three key findings. First, a large and reliable PRP effect was observed in this experiment—mean RT was substantially faster for the long SOA than for the short SOA. Second, the input demands (Background Condition, Sentence Predictability) in the listening task carried over to affect performance in the secondary task—high input demands in the primary listening task resulted in slower RTs, lower accuracies, and higher IES in the secondary task. Most important, for the IES measure, this effect of input demands in the listening task was sensitive to secondary task SOA—higher input demands in the primary task resulted in a larger PRP effect in the secondary task.

Listening Effort

The only significant effect in the analysis of subjective listening effort was a main effect of Background Condition, $F(1,31) = 35.01$, $p < .001$, $\eta p^2 = .75$. Subjective listening effort measured on a 1 to 7 Likert scale was higher for babble (5.51) than for shaped noise (4.03).

Working Memory

Reading Span and Operation Span scores were significantly positively correlated with each other, $r = 0.39$, $p < .05$. However, neither of these measures of working memory correlated significantly with IES, absolute pupil size, or mean peak dilation amplitude in any condition. Appendix A shows all the correlations with working memory tasks.

Pupillary Response

Absolute Pupil Size

The main effects of Background Condition, Sentence Predictability, and Secondary Task SOA on mean absolute pupil size at baseline (-2500 ms – 0 ms) were not significant. Absolute pupil size at baseline is the measure of pupils before the start of a trial, and therefore is not anticipated to be affected by manipulations that occur during the trial (Sentence Predictability and Secondary Task SOA). Furthermore, the non-significant effect of Background Condition, $F(1,22) = .53$, $p = .47$, $\eta p^2 = .02$, was not surprising, given that the baseline time window encompasses one and a half seconds in which there was no background noise stimulus. Furthermore, the one second in which the background noise was presented before the sentence stimuli does not provide the pupil much time to respond and therefore it is not completely surprising that there was no difference between the background noise conditions in the baseline period. The only significant effect in this analysis was the interaction between Background Condition and Secondary Task SOA, $F(1,22) = 4.86$, $p = .038$, $\eta p^2 = .18$. This interaction seems likely to be spurious given that Secondary Task SOA was randomized from trial to trial, and therefore was undefined during the baseline period.

Baseline Corrected Pupil Dilation

Figure 6 displays baseline corrected pupil dilation across the trial time course, with babble in the left panel and shaped noise in the right panel. Tables 1A and 1B display the mean peak dilation amplitude and peak dilation latency for each condition. The baseline corrected data are those that have had the baseline average subtracted from their values and therefore are a measure of the change from baseline pupil size to the pupil size observed across the trial.

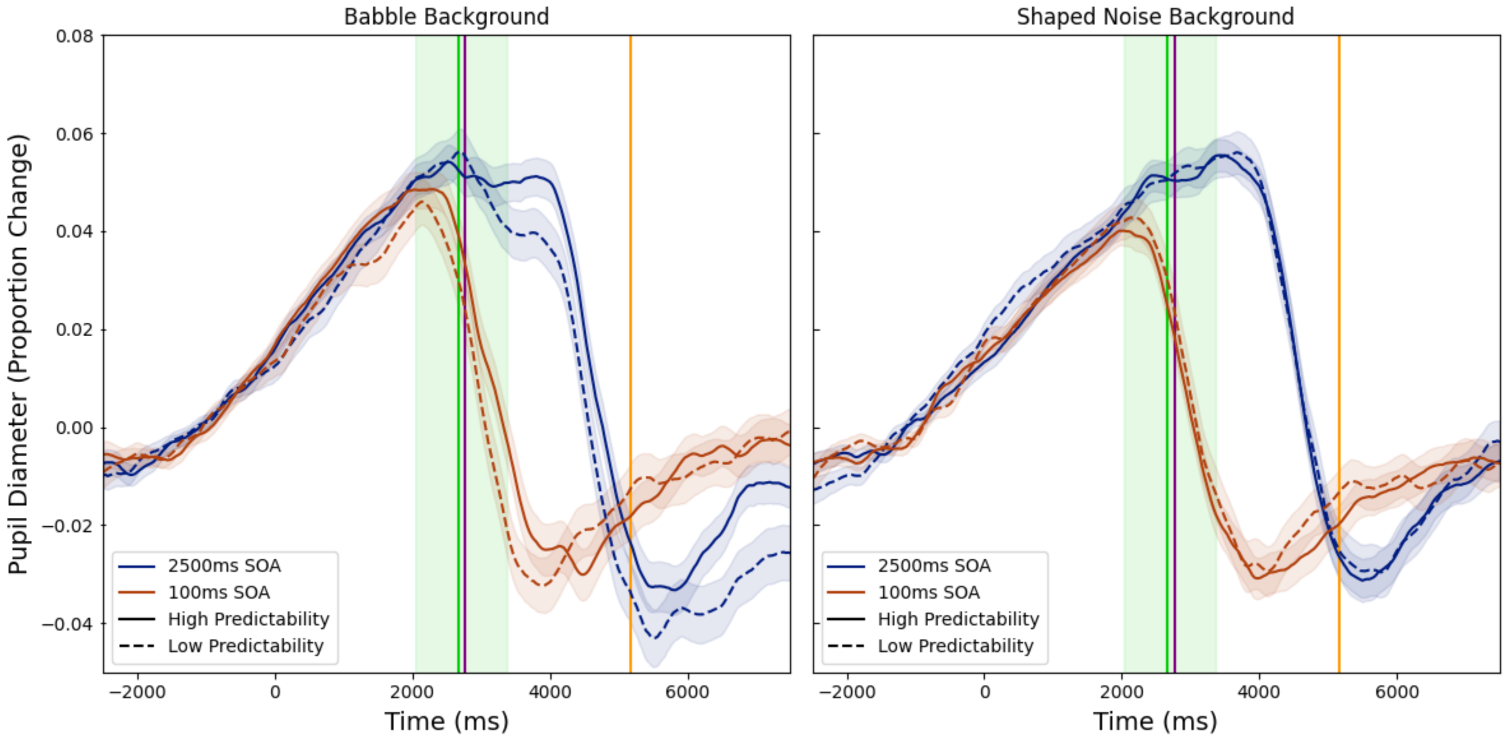
In the analysis of peak dilation latency, there was a significant main effect of Secondary Task SOA, $F(1,22) = 11.43$, $p < .01$, $\eta p^2 = .34$. Peak latency was longer for the 2500 ms SOA (2940 ms) than for the 100 ms SOA (2499 ms).

In the analysis of peak dilation amplitude, there was a significant main effect of Secondary Task SOA, $F(1,22) = 20.56$, $p < .001$, $\eta p^2 = .48$. Peak dilation amplitude was greater for the 2500 ms SOA (.05 proportion change from baseline) than for the 100 ms SOA (.04 proportion change from baseline). There was also a significant interaction between Background Condition and Secondary Task SOA, $F(1,22) = 9.16$, $p < .01$, $\eta p^2 = .29$. In the shaped noise condition, pupil dilation was greater for the 2500 ms SOA (.05 proportion change from baseline) than for the 100 ms SOA (.03 proportion change from baseline), $p < .001$. In contrast, in the babble noise condition, there was no difference between pupil dilation for the 2500 ms SOA (.05 proportion change from baseline) compared to the 100 ms SOA (.04 proportion change from baseline), $p = .22$.

In summary, there was no effect of Background Condition on the baseline corrected pupil dilation measures. However, peak dilation latency was longer for the 2500 ms SOA than for the 100 ms SOA, and peak dilation amplitude was greater for the 2500 ms SOA than for the 100 ms SOA, particularly so for the shaped noise condition.

Figure 6

Change in Pupil Dilation from Baseline



Proportion change in pupil dilation from baseline period of -2500 ms to 0 ms. Long secondary task SOA conditions are shown in blue, and short SOAs are shown in red. High predictability sentences are solid lines, and low predictability sentences are dashed lines. Shaded areas represent standard error of the mean. The 0 ms timepoint is the start of the auditory sentence. The vertical green line represents the average time of target word onset (2670 ms), with the shaded green section representing the time window in which the target word onset occurs (2035 ms to 3367 ms). The vertical purple line represents the average onset of the short SOA secondary task (2770 ms), 100 ms after the average onset of the target word; and the vertical orange line represents the average onset of the long SOA secondary task (5170 ms), 2500 ms after the average onset of the target word. The left panel shows long and short secondary task SOA and high and low sentence predictability conditions in the babble background condition; and the right panel shows long and short secondary task SOA and high and low sentence predictability conditions in the shaped noise background condition.

Table 1A

Average pupil dilation at peak dilation latency (in proportion change dilation from baseline)

	Babble		Shaped Noise	
	High Predictability	Low Predictability	High Predictability	Low Predictability
100 ms SOA	.043	.045	.027	.037
	(2348 – 2848 ms)	(2047 – 2547 ms)	(2356 – 2856 ms)	(2246 – 2746 ms)
2500 ms SOA	.053	.050	.051	.053
	(2682 – 3186 ms)	(2684 – 3184 ms)	(2725 – 3225 ms)	(2666 – 3166 ms)

Values in brackets are the time windows (in milliseconds after sentence onset at 0 ms) over which pupil dilation was averaged. This time window was determined as the 500 ms surrounding the peak dilation latency (plus and minus 250 ms); peak dilation latencies are detailed in Table 1B.

Table 1B

Average peak dilation latency (in milliseconds) at short and long SOAs

	Babble		Shaped Noise	
	High Predictability	Low Predictability	High Predictability	Low Predictability
100 ms SOA	2598	2297	2606	2496
2500 ms SOA	2936	2934	2975	2916

Discussion

This experiment combined the methods from two areas of research that center on allocation of limited capacity resources in support of human cognition. One of these areas focuses on the role of limited capacity resources in listening to speech in noise, while the other area focuses on doing two tasks at once. By combining the methods from these two domains, I aimed to address whether the attentional resources that help one understand speech in noise are related to the attentional resources that are required when we do two tasks simultaneously. I start with a comprehensive summary of the results from this experiment and then summarize how those results address the aims of my study.

Listening Input Demands and Listening Task Performance

Input demands (background noise and sentence predictability) affected listening task performance as expected. Word identification was better with shaped noise than with babble noise. This finding replicates prior studies showing that the informational masking associated with babble noise masks the speech signal more than the simpler signal masking associated with shaped noise (Bennett et al., 2012; Hall et el., 2002; Schneider, Li & Daneman, 2007). Word identification was also better for high predictability sentences than for low predictability sentences. This finding also replicates prior studies that have reported an effect of sentence predictability on target word identification (Hunter, 2020; Hunter & Pisoni, 2018; Sheldon et al., 2008). A highly predictable sentence context can assist word identification, possibly decreasing listening effort and the need for limited capacity cognitive resources.

Listening Input Demands and Absolute Pupil Size

When listening to speech in noise, interference produced by background noise and degraded speech conditions have been shown to be accompanied by increases in pupil size (Wendt et al., 2018; Zekveld et al., 2014; Zekveld & Kramer, 2014). Single-talker maskers are often more interfering than multi-talker maskers, which in turn are more interfering than non-talker noise (Koelewijn et al., 2012). In contrast, the effects of sentence predictability on the pupillary response are not as well-studied (but see Piquado et al., 2010). Given that the Background Condition was only introduced one second before the sentence stimulus, which marked the zero millisecond timepoint, it is reasonable to find that there was no difference between the two background conditions in absolute pupil size during baseline. Sentence Predictability was not manipulated between blocks and therefore should not affect baseline pupil dilation.

Listening Input Demands and Secondary Task Performance

Listening task input demands can also be revealed through their effect on performance in a secondary task (see Gagné et al., 2017 for a review of dual-task paradigms used with listening tasks)—secondary task performance is sensitive to carry-over effects of completing primary tasks of varying difficulty. In the present study, input demand in the primary listening task affected performance on the secondary task. Specifically, secondary task performance was better when following a shaped noise primary task stimulus than for a babble noise primary task stimulus (as measured by accuracy and IES), and better for high predictability than low predictability sentences (as measured by RT, accuracy, and IES). These effects of background noise and sentence predictability on secondary task performance are consistent with attentional resource allocation assumptions in the FUEL; input demands draw cognitive resources needed to complete the primary task, and the effect of this resource depletion is then measured as a dual-task cost in the secondary task.

More important, the present study allowed me to examine the effect of listening task input demands on the PRP effect in the secondary task. A typical PRP effect on secondary task performance was observed in RTs, accuracy, and IES. All three dependent measures revealed more efficient performance for the longer SOA, when the two tasks did not overlap in time, than for the shorter SOA, when the two tasks did overlap (Pashler, 1992; Pashler 1994). Of particular importance, the IES measure revealed that input demands interacted with this PRP effect. The PRP effect was larger for babble noise than for shaped noise, and larger for low predictability sentences than for high predictability sentences. These results demonstrate that input demands in the listening task tap resources that are related in some manner to the PRP effect.

Listening Input Demands, Secondary Task Performance, and the Pupillary Response

Although the effects of a PRP-style secondary task on pupil dilation have not been reported in past literature, the results from conceptually similar paradigms, like the attentional blink and sustained attention in mind-wandering, have been reported and demonstrate greater dilation in response to tasks that overlap in time (Unsworth & Robison, 2018; Zylberberg et al., 2012). Therefore, I predicted that pupil dilation would be greater for the 100 ms SOA than for the 2500 ms SOA. The results showed the opposite pattern: peak dilation amplitude was greater for the 2500 ms SOA than for the 100 ms SOA. Moreover, this effect interacted with Background Condition—it was significant for shaped noise but not for babble. In line with the behavioural results, this interaction demonstrates that input demands in the listening task affect the PRP effect in the secondary task, implying that the two tasks are not attentionally independent.

I also observed an effect of secondary task SOA on peak pupil dilation latency; peak latency occurred later for the long SOA than for the short SOA. This effect may be a result of the uncertainty inherent in the primary task, where the participant does not know when the end of the sentence, and therefore the target word, is going to occur. I suspect this uncertainty was lower in the short SOA condition than in the long SOA condition because the onset of a secondary task target in the short SOA condition provided an overt marker of the end of the sentence.

Listening Input Demands and Subjective Listening Effort

Prior studies have also measured perceived listening effort with different kinds of background noise (Klink et al., 2012; Krueger et al., 2017). These subjective measures do not always align with behavioural measures of listening effort, but nonetheless do reflect some facet of the perceived difficulty level of the listening task (Francis et al., 2016). In the current experiment, in accord with the behavioural findings, listening effort was rated as higher for babble noise than for shaped noise.

Working Memory Correlates

I also examined the correlation between working memory scores and inverse efficiency (IES) and two measures of pupil dilation for each condition of the secondary task. According to the FUEL, secondary task performance is where I would expect to see an effect of individual cognitive abilities: variability in resources leftover from the primary task would lead to variability in secondary task performance, which would correlate with individual WM scores.

However, no such correlation was observed in my study. Similarly, given that pupil dilation is an index of cognitive work, I anticipated the absolute pupil size and the mean peak pupil dilation to be greater for those with lower working memory scores, as they would have to allocate proportionally more of their resources to complete the two tasks. Again, I did not observe such a correlation. A number of factors may have contributed to these null results. Undergraduate students generally have high WM scores, which may have restricted the range of the WM measures. In line with this view, the correlation between Reading Span and Operation Span scores was just .39, far from the size of correlation one would expect if the two tests of WM reliably measured the same construct. Given this modest correlation between the two measures, the opportunity for either WM test to correlate with secondary task performance or pupillary response was limited.

Summary of Experiment 1

Varying input demands produced strong effects on listening task performance, and varying the stimulus onset asynchrony of a secondary task following the primary listening task produced a strong PRP effect. If the resources that support listening in noise are related to the resources that are used when having to do two tasks at once, then the input demands of the primary listening task should modulate the PRP effect. There was evidence for this effect in the IES for secondary task performance. Both input demand manipulations (Background Condition and Sentence Predictability) interacted significantly with secondary task SOA, producing a larger PRP for higher primary task input demands. This result implies that the processes underlying input demand effects in the listening task are related in some manner to processes underlying the PRP effect in secondary task performance.

Following prior research showing that pupil dilation varies directly with cognitive effort, and assuming that cognitive effort should be maximal when the tasks overlap temporally, I was also interested in patterns of pupil dilation. In accord with the PRP seen in behaviour, I speculated that pupil dilation should be greater for the short SOA condition than for the long SOA condition. The pupil dilation pattern did not support this prediction; in fact, the opposite effect occurred. A second pupil dilation prediction hinged on finding a behavioural effect in which input demands in the listening task interacted with the PRP effect. If high input demands produce a larger PRP effect, then perhaps this effect will also be manifest in pupil dilation. Indeed, peak pupil dilation amplitude produced an interaction between Background Condition and Secondary Task SOA, although again the nature of this interaction was not as predicted. The Secondary Task SOA effect (i.e., the pupil dilation version of the PRP effect) was larger for shaped noise than for babble noise.

All told, the predicted behavioural effects were observed in the primary listening in noise task (the effect of input demands), and in the secondary task (the PRP). Furthermore, the effect of input demands in the primary task interacted with the PRP effect in the secondary task as predicted. Although the pupil dilation results were not as predicted, they were clear and demonstrate an interesting complement to the behavioural findings: (1) peak latency was later for the long SOA than the short SOA; (2) peak amplitude was higher for the long SOA than the short SOA; and (3) this effect of secondary task SOA on peak amplitude was greater for shaped noise than for babble noise. Together, the behavioural and pupil dilation results both suggest that listening task input demands tap resources that are related in some manner to the PRP effect in secondary task performance.

Experiment 2

There were two issues in Experiment 1 that motivated changes of the method for Experiment 2. First, the design of Experiment 1 did not allow me to measure how input demands in the listening task affect the absolute pupil size measure during the baseline period; the background noise did not play throughout the entire baseline period in Experiment 1, but in this experiment, the background noise is continuous throughout the entire block allowing me to measure absolute pupil size in response to noise before the trial. In Experiment 2, I examined this issue again using a stronger manipulation of input demands and adjusting the trial design. Specifically, I included a silence condition to contrast with the babble noise and shaped noise conditions. If silence versus noise constitutes a stronger input demand manipulation than babble noise versus shaped noise, then the method in Experiment 2 may offer a better opportunity to measure and understand how input demands in the listening task translate to absolute pupil size in the baseline period. Additionally, the background noise was adjusted to play continuously throughout each noise condition block, providing ample time before stimulus onset to measure the pupillary response to noise. Second, the longer peak latency for the long SOA condition than for the short SOA condition was a strong result that was not anticipated. I surmised that this result may have been caused by the temporal uncertainty associated with onset of the listening task target word when presented in the context of variable carrier sentences. To address this issue, temporal uncertainty about onset of the listening task target was minimized by using a single, unchanging carrier sentence prior to onset of the target word. The aim was to provide a second opportunity, this time with minimal temporal uncertainty about secondary task target onset, to observe the effect of listening task input demands on both performance and pupil size in a PRP dual-task context.

Methods

Participants

Thirty-six undergraduate students (25 females and 11 males; mean age = 18.6 years) who did not participate in Experiment 1 participated in this experiment for partial course credit. All participants spoke English fluently, had learned to speak fluent English by age 7 years, and had self-reported normal or corrected-to-normal vision and self-reported normal hearing.

Apparatus and Stimuli

The apparatus and stimuli were similar to Experiment 1, with just a few changes as noted below.

The visual stimuli were presented from an HP Pro 3130 MT computer, using Presentation® software, on a Dell UltraSharp 2001FP LCD 17-inch monitor. The secondary task stimuli were numbers, which were presented in the centre of the screen in 24pt Arial font. These stimuli were isoluminant green, presented on a grey background, and replaced an isoluminant fixation cross. Participants were seated approximately 60 cm from the computer and used a standard keyboard to respond to the secondary task (determining if a number is odd or even) by pressing the space bar, or one of the "z" or "/" keys.

The auditory stimuli were the Northwestern University Auditory Test Number 6 (NU6) list, which consisted of a single speaker saying a word, preceded by the carrier phrase "say the word". These stimuli were presented from a Genelec 8020D Bi-amplified Active Studio Monitor speaker that was positioned above the monitor. The volume was set such that the target stimuli were presented at 70 dB and the background noise was set 3 dB louder at 73 dB. The background noise conditions were the same as those in Experiment 1 (shaped noise or babble noise) with the addition of a silence condition.

The working memory tasks, listening effort scale, and pupillometry recording were the same as in Experiment 1.

Design

The design of Experiment 2 was the same as that of Experiment 1, with the exception that there were three different listening conditions (silence, shaped noise, and babble noise; counterbalanced across participants and lists of words), and the secondary task required participants to categorize numbers as even or odd (instead of categorizing colours).

Procedure

Operation Span and Reading Span

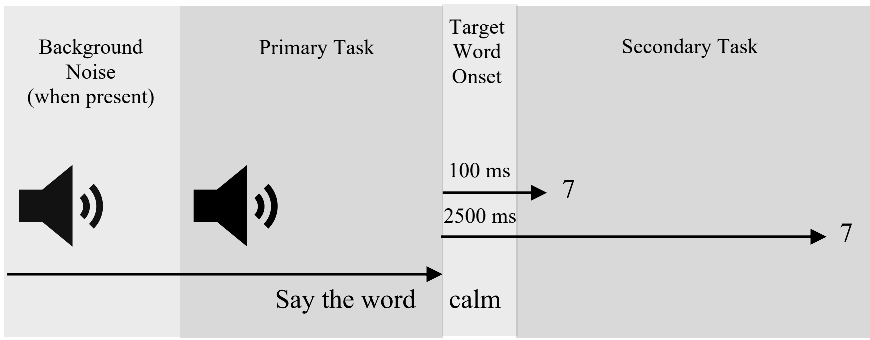
These tasks followed the same procedure as in Experiment 1.

Listening Task

Figure 7 depicts the trial sequence in the listening task. Once again, the eye tracker was calibrated to the participant before beginning the experiment. There were again 192 trials, divided into 12 blocks, but this time there were three different background noise conditions (silence, shaped noise, and babble noise). Unlike Experiment 1, each block began with the background noise turning on (when there was background noise), and this noise played continuously throughout the block. Each trial consisted of the auditory stimulus "say the word ____" where the blank was replaced with the target word for that trial. Participants were then required to say aloud the target word, which the experimenter recorded. Following each auditory stimulus, a number appeared on the screen either 100 ms or 2500 ms after onset of the target word. Participants responded using the keyboard to indicate whether the number was odd or even. Subjective listening effort was rated at the end of each block.

Figure 7

Experiment 2 Listening Task Trial Sequence



This figure depicts the sequence of events throughout the trial, with the background noise followed by the carrier phrase and target word stimulus, and then the secondary number categorization task either 100 ms or 2500 ms after the onset of the target word.

Pupil Data Processing

Pupil data were processed as in Experiment 1 with the exception that the start of the carrier phrase "Say the word…" was now defined as the trial onset (marked as 0 ms). Additionally, the same exclusion criterion was used as in Experiment 1; participants with one or more conditions that had more than three-quarters of the data samples missing were excluded from further processing. Application of this criterion resulted in 18 out of 36 participants' data being included in further analyses.

Data Analysis

The dependent variables analyzed were the same as in Experiment 1: the proportion of primary task target words correctly identified and secondary task response times (RT), accuracy, and Inverse Efficiency Scores (IES). Each of these dependent variables was analyzed using a repeated measures ANOVA that treated Background Condition (silence/shaped noise/babble noise) and Secondary Task SOA (100 ms/2500 ms) as factors. Subjective listening effort measured after each block was analyzed using a one-way ANOVA that treated Background Condition (silence/shaped noise/babble noise) as its factor. Where applicable, post-hoc testing was done using Tukey's HSD.

Reading Span and Operation Span were once again scored using the partial-credit scoring method (Conway et al., 2005), and correlated with secondary task IES, absolute pupil size in the baseline period, and mean peak dilation amplitude as in Experiment 1. As in Experiment 1, these tasks were used as measures of participants' working memory span and were run before the main task of the experiment.

Pupil data were analyzed in the same three ways as Experiment 1: (1) absolute pupil size in the baseline period (-2500 ms to 0 ms); (2) mean peak dilation latency in the baseline corrected data; (3) mean peak dilation amplitude in the baseline corrected data. Time windows used for peak dilation amplitude are detailed in Table 2A. These data were submitted to repeated measures ANOVAs that treated Background Condition (silence/shaped noise/babble) and Secondary Task SOA (100 ms/2500 ms) as factors. Again, where post-hoc tests were performed, they were done using Tukey's HSD.

For the behavioural data, Figures 8 to 11 include error bars with the Morey correction applied (O'Brien & Cousineau, 2014), which accounts for within-subjects designs when calculating the standard error of the mean.

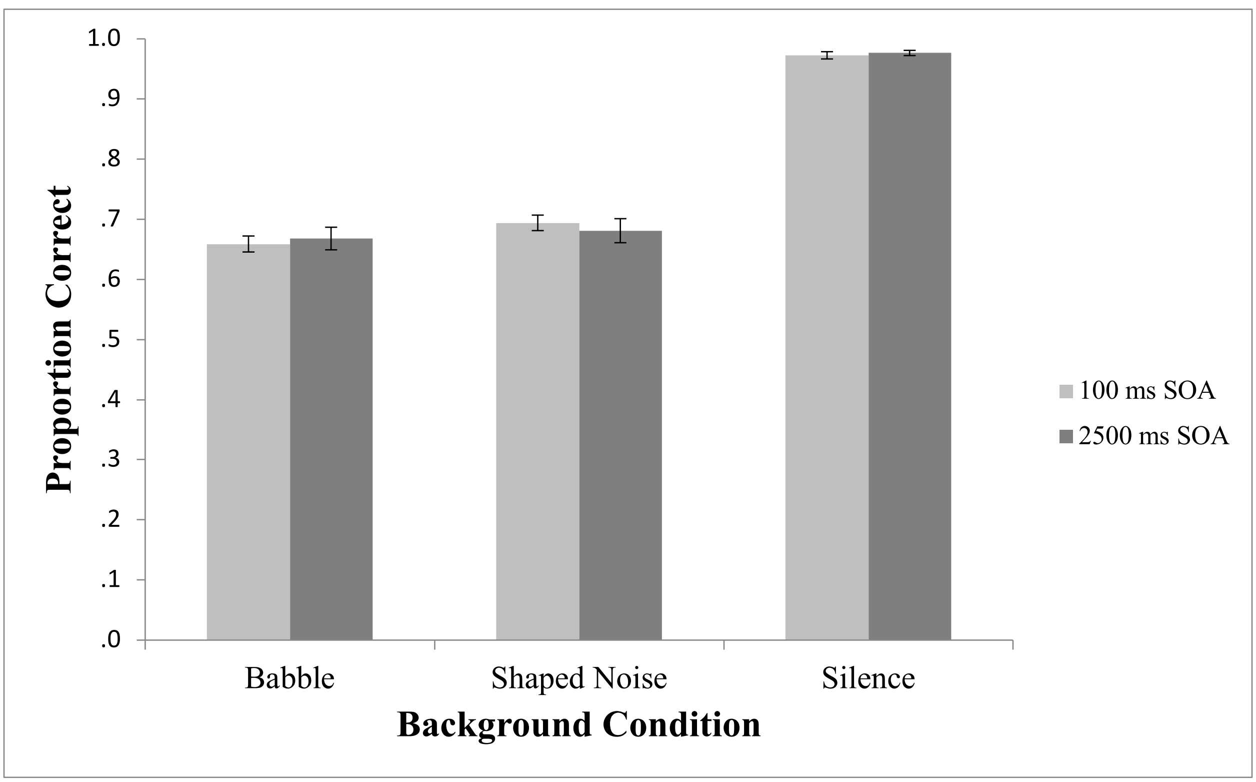
Results

Listening Task: Target Word Identification Accuracy

Figure 8 shows the proportion of correctly identified target words by condition. There was a main effect of Background Condition, $F(2, 46) = 257.5$, $p < .001$, $\eta p^2 = .92$; post-hoc Tukey tests showed that target word identification accuracy was greater for silence (.98) than for shaped noise (.69), $p < .001$, and greater for silence (.98) than for babble noise (.66), $p < .001$. There were no other effects on primary task accuracy.

Figure 8

Word Identification Accuracy



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

Secondary Task: Number Parity Accuracy, Reaction Time, and Inverse Efficiency Score

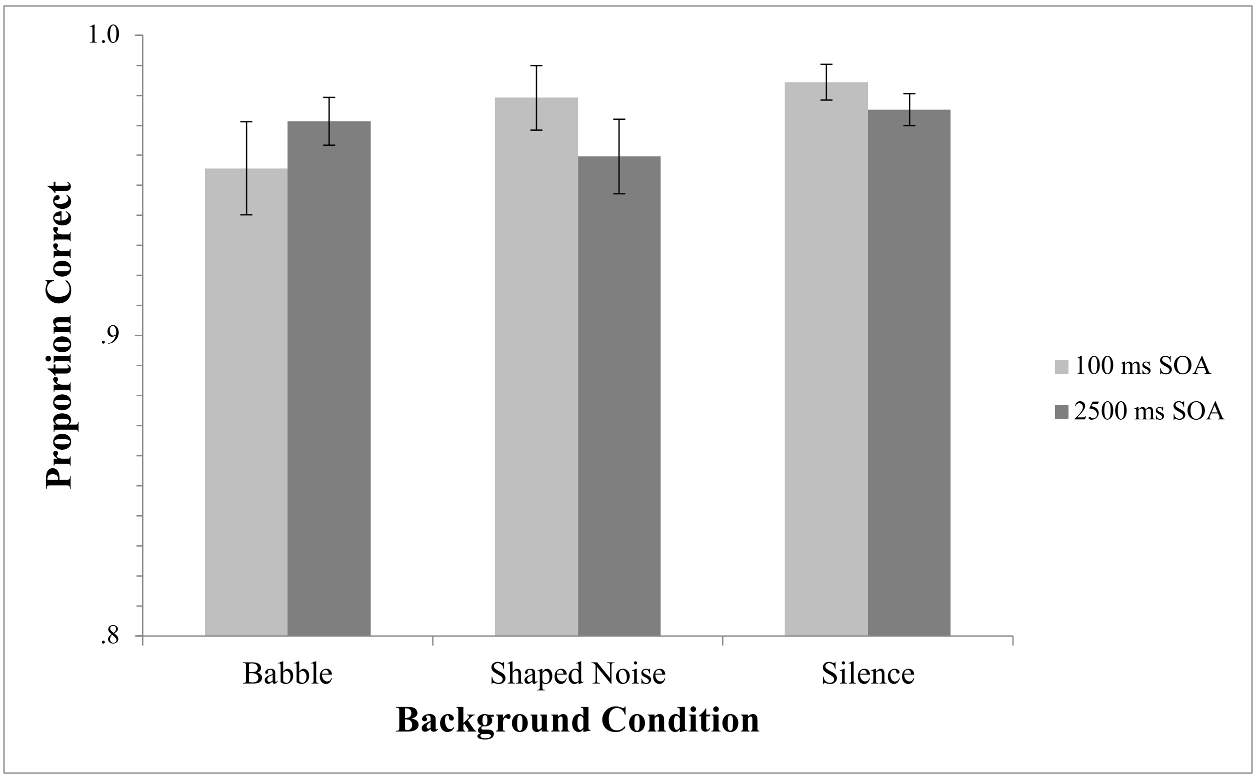
Figures 9 and 10 display proportion correct and mean RT, respectively, for the secondary task.

Accuracy

In the analysis of secondary task accuracy, there were no significant effects.

Figure 9

Secondary Task Accuracy



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

Reaction Time

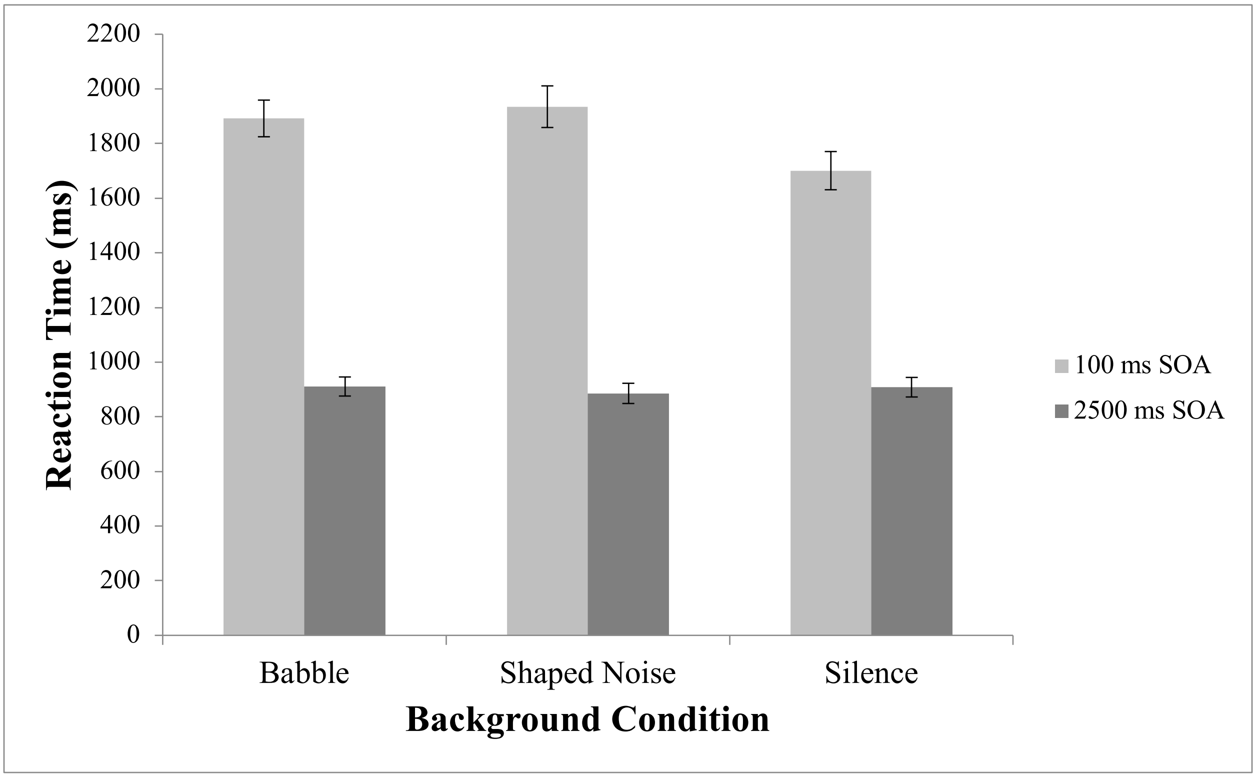
In the analysis of secondary task RTs, there was a significant main effect of Background Condition, $F(2,46) = 22.43$, $p < .001$, $\eta p^2 = .49$; secondary task RTs following both babble noise

(1401 ms) and shaped noise (1410 ms) were higher than following silence (1305 ms), both $p <$ .001. There was a significant main effect of Secondary Task SOA, $F(1,23) = 346.2$, $p < .001$, $\eta p^2 = .94$; secondary task RTs were longer for the 100 ms SOA (1842 ms) than for the 2500 ms SOA (902 ms). There was also an interaction between Background Condition and Secondary Task SOA, $F(2,46) = 36.51$, $p < .001$, $\eta p^2 = .61$. Post hoc comparisons revealed that the difference in secondary task RTs between short SOA and long SOA conditions (i.e., the PRP effect) was significant following the babble noise (981 ms), shaped noise (1049 ms), and silence (793 ms) conditions, all $p < .001$. Moreover, the PRP effect was larger for both shaped noise and babble noise than silence, both $p < .001$, and larger for shaped noise than babble noise, $p < .01$.[3]

[3] Stated differently, for the 100 ms SOA, the RTs for each ‘noise’ condition (1892 ms for babble noise and 1935 ms for shaped noise) were significantly slower than the RTs for silence (1701 ms), both $p < .001$, but not different from each other, $p = .45$. For the 2500 ms SOA, however, there were no significant differences between any Background Conditions.

Figure 10

Secondary Task Reaction Time



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

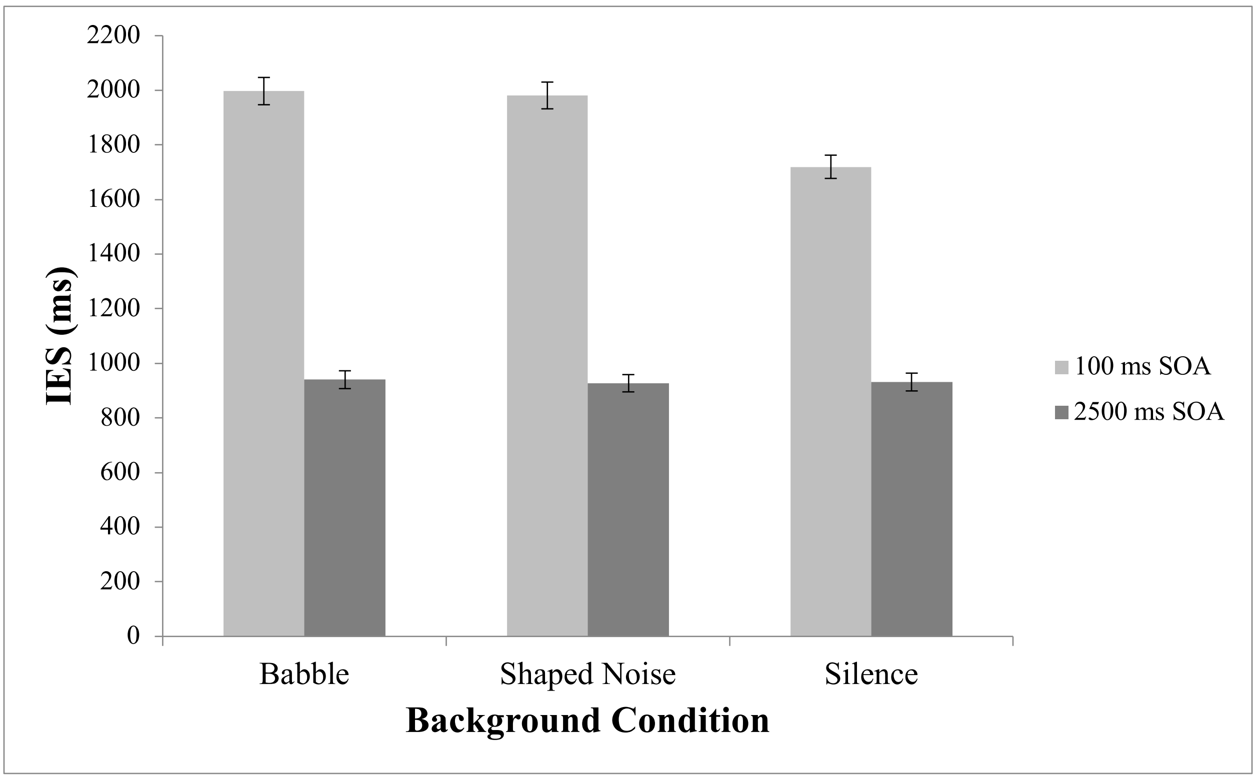
Inverse Efficiency Scores

Figure 11 displays IES for the secondary task. In the analysis of IES, there was a significant main effect of Background Condition, $F(2, 46) = 10.19$, $p < .001$, $\eta p^2 = .31$. Post hoc Tukey tests revealed that secondary task IES was lower following silence (1326 ms) than following both babble noise (1469 ms) and shaped noise (1454 ms), both $p < .01$, but not different for babble and shaped noise, $p = .91$. There was also a significant main effect of Secondary Task SOA, $F(1,23) = 324.2$, $p < .001$, $\eta p^2 = .93$; secondary task IES was higher for the 100 ms SOA (1899 ms) than for the 2500 ms SOA (933 ms). Finally, there was a significant interaction between Background Condition and Secondary Task SOA, $F(2,46) = 16.42$, $p < .001$,

$\eta p^2 = .42$. As seen in the RTs, there was a significant PRP effect for all three background conditions: 1057 ms for babble noise; 1054 ms for shaped noise; 788 ms for silence, all $p < .001$. Moreover, the PRP effect was larger for both shaped noise and babble noise than silence, both $p < .001$, but not different for babble and shaped noise, $p = .48$.

Figure 11

Inverse Efficiency Scores (IES)



Error bars are standard error of the mean, with Morey correction for within-subjects designs applied.

These secondary task results revealed two important findings. First, as in Experiment 1, there was a strong PRP effect in secondary task performance. Second, as in Experiment 1, the

PRP effect was modulated by the input demands of the listening task. In this case the PRP effect was larger for the two noise conditions than for the silence condition.

Listening Effort

In the analysis of subjective listening effort, there was a significant main effect of Background Condition, $F(2, 45) = 146$, $p < .001$, $\eta p^2 = .87$. Post hoc Tukey tests showed that listening effort measured on a 1 to 7 Likert scale differed significantly between all pairs of conditions; effort rating was highest for babble noise (5.66), lowest for silence (2.08), and intermediate for shaped noise (4.66), all $p < .001$.

Working Memory

Reading Span and Operation Span scores significantly positively correlated with each other, $r = 0.87$, $p < .001$. Reading Span and Operation Span scores did not correlate with IES or absolute pupil size in the baseline in any conditions. Operation Span significantly negatively correlated with mean peak pupil dilation amplitude in the babble noise and 2500 ms SOA condition, $r = -.487$, $p < .05$. Appendix B shows all correlations with working memory tasks.

Pupillary Response

Absolute Pupil Size

There was a significant effect of Background Condition on absolute pupil size in the baseline period (-2500 ms – 0 ms), $F(2,34) = 22.72$, $p < .001$, $\eta p^2 = .57$. Post hoc Tukey comparisons revealed a significant difference between all three Background Conditions, all $p <$

.01. Mean absolute pupil size was highest for babble noise (34.0 pixels), lowest for silence (32.3 pixels), and intermediate for shaped noise (33.1 pixels).

Baseline Corrected Pupil Dilation

Figure 12 displays baseline corrected pupil dilation across the trial time course, separated by Background Condition. Tables 2A and 2B display the mean peak dilation amplitude and peak dilation latency separately for each condition. As in Experiment 1, the baseline corrected data are a measure of the change from baseline pupil size to the pupil size observed across the trial.

In the analysis of peak dilation latency, there was a significant main effect of Background Condition, $F(2,34) = 4.1$, $p < .05$, $\eta p^2 = .19$; peak dilation latency was later for silence (2557 ms) than for babble noise (2283 ms), $p = .03$. No other comparisons between Background Conditions were significant.

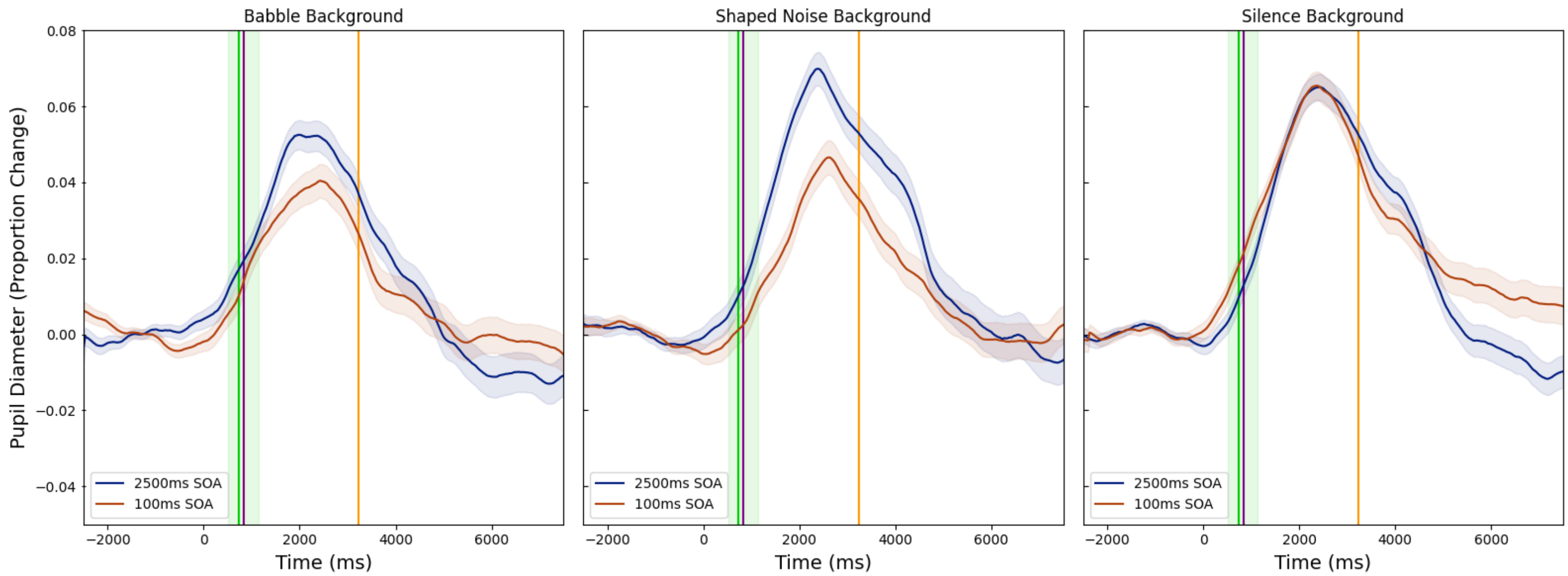
In the analysis of peak dilation amplitude, there was a significant main effect of Background Condition, $F(2,34) = 10.52$, $p < .001$, $\eta p^2 = .38$. Peak dilation amplitude was greater for silence (.065 proportion change from baseline) than for both shaped noise (.053 proportion change from baseline), $p < .05$, and babble noise (.046 proportion change from baseline), $p < .001$. The difference between babble noise and shaped noise was not significant, $p = .26$. There was also a significant main effect of Secondary Task SOA, $F(1,17) = 6.0$, $p < .03$, $\eta p^2 = .26$, where peak dilation amplitude was greater for the 2500 ms SOA (.06 proportion change from baseline) than for the 100 ms SOA (.049 proportion change from baseline), $p < .05$. The interaction between Background Condition and Secondary Task SOA was not significant, $F(2,34) = 3.27$, $p = .065$, $\eta p^2 = .16$, but given that there was a significant interaction between Background Condition and Secondary Task SOA in Experiment 1, I looked closer at this interaction. Peak dilation amplitude was smaller for the short SOA than for the long SOA for shaped noise (.04 vs .06 proportion change from baseline, $p = .04$), but did not differ significantly for short and long SOAs in either the babble noise (.04 vs .05) or silence (.065 vs .064) conditions, both $p > .30$. Post hoc comparisons of the SOA effect across the three background conditions revealed only one significant comparison; the effect of SOA was larger for shaped noise than for silence, $p < .01$.[4]

The results of the baseline corrected pupil data analyses revealed two key findings. First, in accord with the results of Experiment 1, peak amplitude was generally higher for the 2500 ms SOA than for the 100 ms SOA. This result again contradicts my initial prediction that the opposite pattern would occur. Second, this SOA effect on peak amplitude was modulated by input demands of the listening task, being largest for the shaped noise condition and smallest for the silence condition. In accord with the behavioural results, these pupil dilation results point to an interaction between listening task input demands and temporally overlapping dual-task demands.

[4] As seen in secondary task performance, when stated differently, this interaction shows that for the 100 ms SOA, peak dilation amplitude for silence (.065 proportion change from baseline) was greater than for both shaped noise (.042 proportion change from baseline), $p < .01$ and babble noise (.039 proportion change from baseline), $p < .001$; but the difference in peak pupil dilation amplitude between babble and shaped noise was not significant, $p < 1$. For the 2500 ms SOA conditions, there were no significant differences between Background Conditions.

Figure 12

Change in Pupil Dilation from Baseline



Proportion change in pupil dilation from baseline period of -2500 ms to 0 ms. Long secondary task SOA conditions are shown in blue, and short SOAs are shown in red. Shaded areas represent standard error of the mean. The 0 ms timepoint is the start of the auditory stimulus "Say the word…". The vertical green line represents the average time of target word onset (735 ms), with the shaded green section representing the time window in which the target word onset occurs (512 ms to 1136 ms). The vertical purple line represents the average onset of the short SOA secondary task (835 ms), 100 ms after the average onset of the target word; and the vertical orange line represents the average onset of the long SOA secondary task (3235 ms), 2500 ms after the average onset of the target word. The left panel shows long and short secondary task SOA conditions in the babble background condition; the centre panel shows long and short secondary task SOA conditions in the shaped noise background condition; and the right panel shows long and short secondary task conditions in the silence background condition.

Table 2A

Average pupil dilation at peak dilation latency (in proportion change dilation from baseline)

	Babble	Shaped Noise	Silence
100 ms SOA	.039	.041	.065
	(2014 – 2514 ms)	(2290 – 2790 ms)	(2245 – 2745 ms)
2500 ms SOA	.053	.064	.064
	(2052 – 2552 ms)	(2325 – 2825 ms)	(2238 – 2738 ms)

Values in brackets are the time windows (in milliseconds after sentence onset at 0 ms) over which pupil dilation was averaged. This time window was determined as the 500 ms surrounding the peak dilation latency (plus and minus 250 ms); peak dilation latencies are detailed in Table 1B.

Table 2B

Average peak dilation latency (in milliseconds) at short and long SOAs

	Babble	Shaped Noise	Silence
100 ms SOA	2264	2540	2495
2500 ms SOA	2302	2575	2488

Discussion

The aim of this experiment was to combine listening to speech in noise and PRP methods to ask whether the attentional resources that help one understand speech in noise contribute to the overlap in processing when we do two tasks in close sequence. In Experiment 1, I obtained preliminary evidence from both behavioural and pupil measures that input demands when listening to speech in noise interact with processes that produce the PRP effect. I addressed a similar issue in Experiment 2, but with an additional silence condition added to the two noise conditions, and with a primary task that eliminated temporal ambiguity about target onset in the primary listening task.

Listening Input Demands and Listening Task Performance

As in Experiment 1, there was a strong and significant effect of listening task input demands on listening task performance. In particular, performance in the silence condition was substantially better than in the babble noise and shaped noise conditions. However, in contrast to Experiment 1, here I did not observe a difference in listening task performance between the shaped noise and babble noise conditions. At this point, it is unclear to us why this difference was present in Experiment 1 but not in Experiment 2. Nonetheless, the large difference in performance between the silence condition and the two noise conditions allowed me to address my research question: Do processes associated with listening task input demands interact with those associated with temporally overlapping dual-task demands?

Listening Input Demands and Absolute Pupil Size

I observed an effect of Background Condition on mean absolute pupil size in the baseline period. This absolute measure of pupil size during the baseline period was highest for babble noise, intermediate for shaped noise, and lowest for silence. These findings align well with previous studies showing the pupil is sensitive to the cognitive demands of different types of background noise (Koelwijn et al., 2012). These findings also reflect the change in trial design, where having background noise run continuously throughout each block allowed me to measure pupillary response to noise in the baseline period.

Listening Input Demands and Secondary Task Performance

As in Experiment 1, I expected and found the input demands on the listening task to carry over and affect secondary task performance (Gagné et al., 2017). Indeed, I again found typical PRP effects on secondary task performance in both mean RT and IES measures, with better performance for the 2500 ms SOA than for the 100 ms SOA (Pashler, 1992; Pashler 1994). Of particular note, both RT and IES measures revealed an interaction between Background Condition and Secondary Task SOA—the performance difference between the 100 ms and 2500 ms SOA conditions was greater for both babble noise and shaped noise conditions than for the silence condition. This result is consistent with the idea that additional resources required to identify target words in noise (either babble noise or shaped noise) relative to silence can amplify a behavioural measure of the PRP effect.

Listening Input Demands, Secondary Task Performance, and the Pupillary Response

Recall that my initial prediction was that task overlap ought to increase cognitive effort, which should in turn increase pupil dilation. By this view, pupil dilation ought to be greater for the 100 ms SOA than for the 2500 ms SOA. Instead, in Experiment 1, peak dilation amplitude showed the opposite pattern—it was greater for the 2500 ms SOA than for the 100 ms SOA. In Experiment 2, this pattern of results was observed again, with greater peak amplitude for the 2500 ms SOA than for the 100 ms SOA. As in Experiment 1, I was particularly interested in whether this effect of SOA on pupil dilation would be sensitive to the input demands of the listening task. Indeed, the SOA effect on peak amplitude interacted with Background Condition—it was significant in the shaped noise condition, but not significant in both the babble noise and silence conditions. Though the nature of this interaction differed from my original prediction, it was similar across Experiments 1 and 2. Together with the behavioural results of both experiments, these pupil dilation results demonstrate that input demands in the listening task and processes engaged in response to task overlap are not independent.

Listening Input Demands and Subjective Listening Effort

Contrary to the behavioural measures, the subjective ratings of listening effort did reveal differences between all three background noise conditions, with the highest level of subjective effort for babble noise and the lowest for silence. This result supports the idea that background noise manipulations can affect subjective ratings of effort, as well as the idea that subjective ratings do not always align with other measures of effort (Klink et al., 2012; Francis et al., 2016; Krueger et al., 2017).

Working Memory Correlates

As in Experiment 1, WM scores did not correlate with IES on the secondary task or with absolute pupil size in the baseline period. Operation Span did, however, significantly negatively correlate with mean peak dilation amplitude in the babble noise and 2500 ms SOA condition. While this correlation was significant, I am hesitant to identify it as a meaningful relationship given that is the only significant correlation, out of 36, in this experiment. Furthermore, this correlation was only significant for the Operation Span and not for the Reading Span (though the Reading Span correlation followed the same direction). Though this correlation might draw attention to a relationship between working memory capacity and pupil dilation in one particular condition, the correlations did not reach significance in any other conditions and this one should therefore be interpreted with caution.

Summary of Experiment 2

As in Experiment 1, listening input demands affected listening task performance, and carried over to affect secondary task performance. For both the RT and IES measures, the expected PRP effect in secondary task performance was observed, with substantially better performance for the 2500 ms SOA than for the 100 ms SOA. More important, listening task input demands interacted with the PRP effect. For both RT and IES, the PRP effect was greater for both the babble noise and shaped noise conditions than for the silence condition. These results suggest that input demands have a greater effect when the two tasks overlap closely in time, and support the proposal that listening to speech in noise and doing two tasks that overlap temporally tap related processes.

Turning to the pupil dilation results, I again found that pupil dilation tended to be greater for the long SOA condition than for the short SOA condition, rather than the opposite (which I had initially predicted). More important, I was particularly interested in looking at pupil dilation under conditions in which input demands interacted with the PRP effect in behaviour, as occurred for the RT and IES measures. Would there be an accompanying interaction in the peak amplitude measure of pupil size? As in Experiment 1, I did find such an interaction. The effect of SOA on peak amplitude described above was largest in the shaped noise condition and smallest in the silence condition. As for the behavioural results, this interaction supports the proposal that listening to speech in noise and doing two tasks that overlap temporally tap related processes.

General Discussion

This study combined the methods from two areas used to study the role of attention in human cognition. One of these areas focuses how limited capacity attentional resources are used in listening to speech in noise, while the other framework focuses on attentional limits associated with doing two tasks at once. By combining the methods from these two domains, I aimed to address whether the attentional resources needed to understand speech in noise are related to those that are used when we do two tasks simultaneously.

Listening Task Input Demands Affected Secondary Task Performance

A first broad issue addressed by the present results concerns whether input demands of the listening task carried over to affect secondary task performance, with higher primary task input demands resulting in poorer secondary task performance. A more specific objective was to examine the time course of this effect. If the effect of primary task input demands on secondary task performance is driven by the same processes that produce the PRP, then the input demand effect should interact with the PRP effect in secondary task performance; that is, the input demand effect on secondary task performance should be greater for the short SOA condition than for the long SOA condition. In both Experiments 1 and 2, there was strong evidence that primary task input demands affected secondary task performance, with less efficient secondary task performance for greater primary task input demands. In both experiments I found evidence that this effect interacted with secondary task SOA. In Experiment 1, there was a greater PRP effect in the secondary task IES measure for the greater listening task input demands. In Experiment 2, a similar interaction appeared in both the RT and IES measures of the secondary task, with the PRP effect being larger for the two noise conditions than for the silence condition. This effect may have been stronger in Experiment 2 because of the stronger manipulation of input demands in Experiment 2—whereas Experiment 1 compared two noise conditions, Experiment 2 compared those two noise conditions to silence. Overall, I conclude that resource allocation to meet input demands of the listening task is indeed related in some manner to processes that produce a performance bottleneck when two tasks overlap temporally.

Pupil Dilation Was Sensitive to Listening Task Input Demands

A second important issue concerns the effect of listening task input demands on pupil dilation. The design of Experiment 1 did not allow me to properly measure the effect of the type of background noise on absolute pupil size during the baseline period. However, by having the background noise run continuously throughout the blocks in Experiment 2, I was able to show that absolute pupil size during the baseline period did vary as a function of listening task input demands, being largest for babble and smallest for silence.

Pupil Dilation Was Sensitive to the Interaction Between Input Demands and Task Overlap

Given the interaction between listening task input demands and the PRP effect in my behavioural measures, it was of particular interest whether pupil dilation was sensitive to this interaction. In both Experiments 1 and 2, the baseline corrected measure of pupil dilation amplitude did produce an interaction between listening task input demands and secondary task SOA—that is, the effect of task overlap on pupil dilation amplitude, as measured by differences in pupil dilation amplitude for short and long SOA conditions, varied significantly as a function of listening task input demands. Together with the behavioural results, these pupil dilation results provide strong evidence that processes engaged to meet the challenge of listening in noise are related in some manner to attention limits associated with task overlap.

Following prior research showing that pupil dilation varies directly with cognitive effort, and assuming that cognitive effort should be maximal when the tasks overlap temporally, I predicted that pupil dilation ought to be greater for the short SOA than for the long SOA, and particularly so for high listening task input demands. In fact, the observed interaction did not follow this predicted pattern at all. Rather, pupil dilation amplitude was greater for the long SOA than for the short SOA in both Experiments 1 and 2, and this pattern of results was most robust for the shaped noise condition in both experiments. This pattern was the opposite of that predicted. Interpretation of this surprising result poses an interesting and important challenge.

One way to explain the unanticipated direction of the effect of secondary task timing on pupil dilation amplitude relates to the construct of cognitive overload. Many studies have reported that when a task is more difficult than a participant can comfortably complete, the pupil often constricts instead of dilating (Granholm et al., 1996; Zekveld & Kramer, 2014). This pattern has been studied in a variety of cognitive tasks, and pupillary constriction is generally attributed to participants having hit a limit to cognitive resource allocation. In accord with this general idea, in tasks involving listening in noise, pupillary constriction often accompanies very difficult SNRs. Although the subjective ratings of listening effort did not indicate that participants found listening in noise to be extremely difficult, there is evidence in the literature that subjective measures of listening effort do not necessarily correspond with behavioural and physiological measures of listening effort (Francis et al., 2016).

I predicted that the pupils would dilate more for the short secondary task SOA than for the long secondary task SOA on the basis that task overlap for the short SOA would increase the need for cognitive resources. The opposite pattern of results might indicate instead that task overlap elicited cognitive overload, producing resource demands that exceeded what could comfortably be allocated. One consequence of this cognitive overload could be decreased investment in the task and thus a decrease in pupil dilation. It is particularly interesting that this pattern of results (i.e., lower pupil dilation for the short SOA than long SOA) was observed for the shaped noise condition in both experiments, but not for the babble noise condition in either experiment, and with no hint of such an effect for the silence condition in Experiment 2. The absence of this effect in the silence condition and presence of this effect for the shaped noise condition in Experiment 2 both support the idea that it is challenging input demands together with task overlap that jointly leads to cognitive overload. However, the non-significant effect in the babble noise condition contradicts this idea. One possibility is that babble noise on its own, and without task overlap, can produce some amount of cognitive overload, which could mitigate against observing differences in pupil dilation for the short and long SOAs. For this explanation to hold, one would predict lower pupil dilation for the babble noise condition than for both the shaped noise and silence conditions when looking solely at the long SOA. A post-hoc analysis comparing mean peak pupil dilation in only the 2500 ms SOA in Experiment 2 was not significant, $F(2,34) = 2.92$, $p = .067$. However, the pattern seen in Figure 12 aligns well with this prediction, and the pattern is in the direction this theory would predict (.053 proportion change from baseline in babble noise compared to .064 proportion change from baseline in shaped noise and silence). This proposal for thinking about cognitive overload also suggests that there may be a “sweet spot” in input demands necessary to measure the interaction of listening task input demands and task overlap on pupil dilation—input demands that are too low may fail to produce cognitive overload at the short SOA (as in the silence condition), and input demands that are too high may produce cognitive overload at both short and long SOAs (as in the babble noise condition). Further work will be needed to confirm this complex property of the interaction between listening task input demands and task overlap on pupil dilation.

Summary

Taking into account both the behavioural and pupil dilation results, the present study supports the view that putative attentional resources underlying listening to speech in noise are related to those responsible for the decreased performance when tasks overlap observed commonly in studies requiring completion of two temporally overlapping tasks. Additionally, the two experiments did produce a wide array of findings that inform my understanding of: (a) how attentional resources are used and allocated in a dual-task listening environment; and (b) the utility of the pupillary response as a measure of listening and cognitive effort in a dual-task listening environment. Many of these results align with well-known findings when studying listening to speech in noise (FUEL; Pichora-Fuller et al., 2016) and dual-task interference in temporally overlapping tasks (i.e., the PRP effect; Pashler, 1994). In particular, I observed large effects of input demands on performance in the primary listening task in both experiments. I observed large PRP effects in the secondary task in both experiments. Input demand effects in the primary listening carried over to the secondary task in both experiments. Pupil dilation was sensitive to task-evoked activity, with a strong dilation effect occurring in the 4-5 second window following onset of the primary task and was sensitive to 'noise' when compared to silence in Experiment 2. Finally, measures of pupil dilation were sensitive to secondary task SOA, although not in the manner predicted by the view that temporal overlap of tasks should maximize both cognitive effort and pupil dilation. These experiments support previous studies that have demonstrated the effects of input demands on listening performance, secondary task performance, and pupillary response, as well as studies that show a classic PRP pattern in dual-task performance. In combining these two methodologies and using pupillometry as a measure of cognitive resource use, I have shown that, together, these result in measurable interactions.

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International journal of audiology*, *47*(sup2), S53-S71.

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scandinavian journal of psychology*, *50*(5), 371-384.

Auditec of St. Louis. (1971). Four-talker babble. 2515 S. Big Bend Boulevard, St. Louis, MO, 63143–2105.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, *91*(2), 276.

Bennett, K. O. C., Billings, C. J., Molis, M. R., & Leek, M. R. (2012). Neural encoding and perception of speech signals in informational masking. *Ear and Hearing*, *32*(2), 231.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?. *Psychologica Belgica*, *51*(1), 5-13.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing research*, *229*(1-2), 132-147.

Doherty, K. A., & Desjardins, J. L. (2015). The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults. *Frontiers in Psychology*, *6*, 721.

Ellenbogen, R., & Meiran, N. (2008). Working memory involvement in dual-task performance: Evidence from the backward compatibility effect. *Memory & Cognition*, *36*(5), 968-978.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. Memory & cognition, 43(2), 226–236.

Francis, A. L., MacPherson, M. K., Chandrasekaran, B., & Alvar, A. M. (2016). Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. *Frontiers in psychology*, *7*, 263.

Francis, A. L., & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics*, *71*(6), 1360-1374.

Gagne, J. P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in hearing*, *21*, 2331216516687287.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*(4), 457-461.

Hall III, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, *23*(2), 159-165.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190-1192.

Hunter, C. R. (2020). Tracking cognitive spare capacity during speech perception with EEG/ERP: Effects of cognitive load and sentence predictability. *Ear and hearing*, *41*(5), 1144-1157.

Hunter, C. R., & Pisoni, D. B. (2018). Extrinsic cognitive load impairs spoken word recognition in high-and low-predictability sentences. *Ear and Hearing*, *39*(2), 378.

Johnson, J., Xu, J., Cox, R., & Pendergraft, P. (2015). A comparison of two methods for measuring listening effort as part of an audiologic test battery. *American journal of audiology*, *24*(3), 419-431.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585.

Katidioti, I., Borst, J. P., & Taatgen, N. A. (2014). What happens when we switch tasks: Pupil dilation in multitasking. *Journal of experimental psychology: applied*, *20*(4), 380.

Klink, K. B., Schulte, M., & Meis, M. (2012). Measuring listening effort in the field of audiology—A literature review of methods (part 2). *Zeitschrift für Audiol*, *51*(2), 60-67.

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hearing research*, *312*, 114-120.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and hearing*, *33*(2), 291-300.

Krueger, M., Schulte, M., Zokoll, M. A., Wagener, K. C., Meis, M., Brand, T., & Holube, I. (2017). Relation between listening effort and speech intelligibility in noise. *American Journal of Audiology*, *26*(3S), 378-392.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious?. *Perspectives on psychological science*, *7*(1), 18-27.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(2), 189.

Mathôt, S. (2018). Pupillometry: psychology, physiology, and function. *Journal of Cognition*, *1*(1).

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 471.

McCloy, D. R., Lau, B. K., Larson, E., Pratt, K. A., & Lee, A. K. (2017). Pupillometry shows the effort of auditory attention switching. *The Journal of the Acoustical Society of America*, *141*(4), 2440-2451.

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., De Lissa, P., Graham, P., & Lyxell, B. (2017). Objective assessment of listening effort: Coregistration of pupillometry and EEG. *Trends in hearing*, *21*, 2331216517706396.

Mitterer, H., & Mattys, S. L. (2017). How does cognitive load influence speech perception? An encoding hypothesis. *Attention, Perception, & Psychophysics*, *79*(1), 344-351.

Murphy, D. R., Craik, F. I., Li, K. Z., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and aging*, *15*(2), 323.

Neagu, M. B., Dau, T., Hyvärinen, P., Bækgaard, P., Lunner, T., & Wendt, D. (2019). Investigating pupillometry as a reliable measure of individual's listening effort. In *Proceedings of the International Symposium on Auditory and Audiological Research* (Vol. 7, pp. 365-372).

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *The quantitative methods for psychology*, *10*(1), 56-67.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior research methods*, *47*(4), 1343-1355.

Padilla, L. M., Castro, S. C., Quinan, P. S., Ruginski, I. T., & Creem-Regehr, S. H. (2019). Toward objective evaluation of working memory in visualizations: A case study using pupillometry and a dual-task paradigm. *IEEE transactions on visualization and computer graphics*, *26*(1), 332-342.

Park, S. Y., Kim, M. J., Sikandaner, H., Kim, D. K., Yeo, S. W., & Park, S. N. (2016). A causal relationship between hearing loss and cognitive impairment. *Acta oto-laryngologica*, *136*(5), 480-483.

Pashler, H. (1984). Processing stages in overlapping tasks: evidence for a central bottleneck. *Journal of Experimental Psychology: Human perception and performance*, *10*(3), 358.

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science*, *1*(2), 44-48.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Pschophysiology, 11*(5), 559-566.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.

Piquado, T., Cousins, K. A., Wingfield, A., & Miller, P. (2010). Effects of degraded sensory input on memory for speech: Behavioral data and a test of biologically constrained computational models. *Brain research*, *1365*, 48-65.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560-569.

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247-261.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*, *7*, 31.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*, *18*(07), 559-572.

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*(1), 489-499.

Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. *Attention, Perception, & Psychophysics*, *80*(2), 461-484.

Welford, A. T. (1952). The ‘psychological refractory period’and the timing of high-speed performance—a review and a theory. *British Journal of Psychology. General Section*, *43*(1), 2-19.

Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., & Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hearing research*, *369*, 67-78.

Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech*, *37*(3), 221-235.

Wingfield, A., Amichetti, N. M., & Lash, A. (2015). Cognitive aging and hearing acuity: Modeling spoken language comprehension. *Frontiers in Psychology*, *6*, 684.

Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *Journal of the American Academy of Audiology*, *18*(07), 548-558.

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in hearing*, *22*, 2331216518800869.

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*, *101*, 76-86.

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*(3), 277-284.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and hearing*, *31*(4), 480-490.

Zhivomirov, H. (2018). A method for colored noise generation. *Romanian journal of acoustics and vibration*, *15*(1), 14-19.

Zylberberg, A., Oliva, M., & Sigman, M. (2012). Pupil dilation: a fingerprint of temporal selection during the “attentional blink”. *Frontiers in Psychology*, *3*, 316.

APPENDIX A
Working Memory Correlations

Table 1

Working Memory Task Correlations with Secondary Task Inverse Efficiency Scores

	Babble				Shaped Noise			
	High Predictability Sentences		Low Predictability Sentences		High Predictability Sentences		Low Predictability Sentences	
	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA
Operation Span	-.187	-.347	-.175	-.322	-.344	-.235	-.250	-.273
Reading Span	-.233	-.173	-.069	-.163	-.167	-.330	-.255	-.205

Table 2

Working Memory Task Correlations with Peak Pupillary Dilation Amplitude

	Babble				Shaped Noise			
	High Predictability Sentences		Low Predictability Sentences		High Predictability Sentences		Low Predictability Sentences	
	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA
Operation Span	.195	.019	-.256	-.001	-.274	.024	.026	-.059
Reading Span	.373	.217	.032	.334	-.133	.208	.133	.148

Table 3

Working Memory Task Correlations with Absolute Pupillary Dilation Size in the Baseline Period

	Babble				Shaped Noise			
	High Predictability Sentences		Low Predictability Sentences		High Predictability Sentences		Low Predictability Sentences	
	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA
Operation Span	.081	.104	.196	.150	.164	.128	.051	.175
Reading Span	.209	.242	.241	.277	.288	.137	.131	.224

APPENDIX B
Working Memory Correlations

Table 1

Working Memory Task Correlations with Secondary Task Inverse Efficiency Scores

	Babble		Shaped Noise		Silence	
	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA
Operation Span	-.175	-.211	-.142	-.022	-.206	-.046
Reading Span	-.180	-.054	-.127	.052	-.197	.059

Table 2

Working Memory Task Correlations with Peak Pupillary Dilation Amplitude

	Babble		Shaped Noise		Silence	
	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA
Operation Span	.040	-.487*	-.349	-.411	-.377	-.425
Reading Span	-.002	-.435	-.285	-.398	-.350	-.371

*Significant at the level of $p < .05$ (2-tailed)

Table 3

Working Memory Task Correlations with Absolute Pupillary Dilation Size in the Baseline Period

	Babble		Shaped Noise		Silence	
	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA	100 ms SOA	2500 ms SOA
Operation Span	.008	.064	.090	.056	.073	.070
Reading Span	.045	.109	.047	.062	.017	.012

CHAPTER 4: GENERAL DISCUSSION

The main objective of this thesis was to examine the constructs of attention allocation and limited attentional capacity using a unique combination of methods drawn from literatures on divided attention and listening in noise. This combination of methods allowed me to ask theoretically-driven questions about attention by tapping into and measuring attention in diverse ways. The experiments asked participants to listen to speech in noise while also completing resource-demanding secondary tasks. Pupil dilation served as a particularly important measure of attentional resource use.

In Chapter 2, I combined a listening in noise task with an n-back working memory task. The input demands of the listening task were manipulated by varying background noise. Consistent with prior studies that have examined related issues, the input demands affected listening task performance (Hall et el., 2002; Schneider et al., 2007), secondary task performance (Gagné et al., 2017), and the pupillary response (Zekveld et. al, 2014; Zekveld & Kramer, 2014). Listening performance and secondary task performance were worse in response to increased input demands and these additional demands were often reflected in greater pupillary dilation. Also consistent with prior studies, the pupillary response was sensitive to the cognitive load in the n-back task, with greater dilation for the dual-task n-back condition than for the single-task condition (Hjortkjær et al., 2020; Scharinger et al., 2015). Together, these findings constitute strong conceptual replications of prior studies of both listening in noise and dual-task research. Moreover, these findings extend previous knowledge by allowing me to directly examine the interaction between listening in noise and this particular dual-task context.

Combining the listening in noise task with an n-back working memory task revealed that the effects of background noise carried over into n-back performance. This result indicates that the attentional resources used to listen in noise are also used in the working memory task. A working memory task of this nature had not previously been used in prior studies of listening in noise, and therefore the present results add to the literature that uses dual-task methods to infer listening effort in a primary task. This result also aligns with the idea that working memory and attention are related (Cowan et al., 2005) and that maintaining items in working memory uses resources that would otherwise be used to attend to other tasks. In other words, the demands of the listening task and the demands of the working memory task overlap at a processing level, resulting in the sharing of attentional resources when those tasks are at their most demanding.

In using the n-back task to introduce cognitive load, I also discovered that this task presented some challenges in interpreting the pupillary response. While absolute pupillary dilation was larger in the dual-task n-back condition than in the single-task condition, the task-evoked pupillary response to the listening task was harder to interpret with regard to the effect of secondary task. A key result was that the task-evoked response was smaller in the dual-task n-back condition than in the single-task condition. However, there are at least two different interpretations of this effect. On the one hand, this result could imply that fewer attentional resources were available in the dual-task n-back condition than in the single-task condition, and that the smaller task-evoked component in the dual-task condition reflects these limited attentional resources. On the other hand, this result could also be related to the higher absolute pupil dilation for the dual-task n-back condition placing a physical limit on how much task-evoked dilation could possibly occur in the dual-task n-back condition. This issue was addressed in Chapter 3 by using a dual-task method that eliminated the difference in absolute dilation prior to onset of the listening task.

Chapter 3 was similar to Chapter 2 in that its goal was to examine the joint attentional demands of listening in noise and divided attention, though it differed in the nature of the secondary task used. Once again, I found results that constitute close conceptual replications of the effects of input demands on the listening task that have been reported in other studies: input demands on listening affected listening performance, secondary task performance, and the pupillary response. I also found the usual Psychological Refractory Period (PRP) pattern of results in secondary task performance (Pashler, 1992; Pashler 1994), and an effect of temporally overlapping tasks on the pupillary response (Zylberberg et al., 2012). Again, these results served as a foundation for examining directly how listening in noise interacted with this specific dual-task context.

The dual-task context in Chapter 3 differed from that in Chapter 2 in an important way. Specifically, the PRP dual-task method examined closely the effect of temporal overlap in task demands. By using a secondary task that included a condition in which there was maximal temporal overlap in task demands with the primary task, the effect of this dual-task demand became more evident than with the working memory task in Chapter 2. In particular, the PRP method included a condition in which critical resource demanding components of the secondary task had a high degree of temporal overlap with the listening task and a condition in which those resource demanding components of the two tasks did not overlap temporally at all. By randomly intermixing these two types of trials, the pre-task period was equated between the two secondary task conditions in a way that was not possible with the n-back task in Chapter 2. In this way, the PRP method solves one of the challenges introduced in Chapter 2 by equating the baselines with respect to secondary task load and made it possible to compare more meaningfully the task-evoked pupillary response for these two conditions.

This unique property of the PRP dual-task method made possible the most important findings of Chapter 3. First, the effect of listening task demands on secondary task performance was dependent on the degree of overlap between the two tasks, and therefore reflective of the attentional resources required when tasks overlapped and when they did not. Second, the interaction between the two tasks was also observable in the pupillary response, where pattens of pupil dilation differed across secondary task conditions in accord with listening task input demands. This interaction between listening task input demands and single- and dual-task demands was not observed in Chapter 2, but as noted, the different baseline pupil dilations for single- and dual-task conditions in Chapter 2 compromised comparisons of task-evoked pupil dilation across the single- and dual-task conditions.

Taken together, these two chapters highlight how listening in noise and dual-task demands both draw on limited attentional resources. Many of the results in Chapters 2 and 3 constitute close conceptual replications of results reported in prior studies, including the effect of listening task input demands on listening task performance, secondary task performance, and pupillary response, as well as other well-established divided attention effects on performance and pupillary response. These findings ensure that the research I have conducted aligns well with prior studies, and the replication of these results then provide a foundation for the novel and most important aspects of the current thesis. Specifically, by combining listening in noise with dual-task methods I was able to demonstrate that listening task input demands interact with dual-task demands, a result that supports the view that both draw on limited attentional capacity. This finding is consistent with the view that the cognitive resources used when listening in noise (as in the FUEL; Pichora-Fuller et al., 2016) are the same as those that are used in divided attention contexts with constant working memory load and in PRP temporally overlapping tasks (Kahneman, 1973).

Timing of Task Demands

In both empirical chapters, listening in noise was combined with a resource-demanding secondary task, and in both chapters there was evidence that these tasks tap into the same attentional resources. The results did differ between these two chapters, however, and the interaction seen in Chapter 3 was not present in Chapter 2. The most salient difference between the studies of the two chapters concerns the timing of the secondary task, which I argue is the source of the difference in results.

In Chapter 2, the working memory task required participants to remember two numbers from preceding trials throughout the current trial (i.e., while also completing the listening task), to compare a newly presented number on the current trial to one of the two remembered numbers from previous trials, and to update the numbers to be remembered for the following trial. Note that the comparison and updating of the n-back task were cued by a number presented 2500 ms after presentation of the listening task target. In Chapter 3, by contrast, participants were required to categorize a stimulus (either a colour or a number) presented 100 ms or 2500 ms after onset of the listening task target. Both secondary tasks demanded the use of cognitive resources above and beyond those required by the listening task, but the different temporal nature of the task demands may well have played a large part in the extent to which I observed interference between the listening in noise task and the secondary task. According to the FUEL, the demands for available attentional resources are evaluated to determine how attention should be allocated;

when tasks overlap temporally, the demands for attention allocation are higher, and therefore more strain on resource allocation should be evident.

Although in Chapter 2 the working memory task demands to maintain information in memory did overlap with the task demands of the listening task (i.e., the two numbers in the n-back had to be remembered while completing the listening task), the updating of this information did not occur until well after completion of the listening task. In contrast, the nature of the secondary task timing in Chapter 3 meant that the active work done to complete and respond to the secondary task overlapped temporally with the listening task nearly exactly in the short SOA condition. This temporal overlap in task demands means that the work done to complete the listening task and the work done to complete the secondary task required substantial cognitive resources at the same time. Moreover, there is some evidence that the process of updating information in working memory is more costly to attentional resources than the process of maintaining information (Kessler & Oberauer, 2014), which aligns with the view that the overlap in task demands was likely greater in the PRP study of Chapter 3 than the n-back study in Chapter 2. This is not to say that maintaining numbers in working memory in the n-back task did not require cognitive resources, but rather that the peak requirement for resources in the secondary task likely occurred well after the listening task was complete in Chapter 2. As a result, the temporal overlap in task demands and therefore resources would have been significantly greater in the PRP study of Chapter 3 than in the n-back study of Chapter 2. This line of argument suggests that the timing of task demands of the primary and secondary tasks may play an integral role in whether the demands on attention interact between the two tasks—evidence that the tasks tap into the same attentional resources may be most clear when the overlap in task demands is high.

The combination of methods I used in these studies highlighted the utility of different dual-task methods and the potential importance of the timing of task demands in these methods. In Chapter 2, primary task input demands affected secondary task performance, and this was evident in the behavioural data, but not in the pupillary response data—the nature of the n-back task may have compromised the task-evoked response to the listening task across the n-back (dual-task) and no n-back (single-task) conditions. So, although there was an effect of primary task input demands on secondary task performance in the behavioural data, the listening in noise task-evoked pupil response could not easily be compared across n-back (dual-task) and no n-back (single-task) conditions in the physiological data. Therefore, in Chapter 3 a PRP dual-task method was used; this method allowed the task-evoked responses to the listening task to be compared between short and long SOA conditions without a baseline difference in pupil dilation that could potentially compromise interpretation of the task-evoked component. In this study, not only did listening task input demands affect secondary task performance in the behavioural results, but also an interaction between listening input demands and secondary task timing was evident in the task-evoked pupillary response. Thus, it appears that changes in the relative timing of peak resource demands for the primary and secondary tasks from Chapter 2 to Chapter 3 strengthened the results indicating that primary and secondary tasks tapped the same attentional resources.

Limitations

The research presented in this thesis had the broad goal of conducting theoretically-driven studies of divided attention in the area of Cognitive Hearing Science (CHS). The experiments were aimed specifically at the allocation of limited attentional resources in a listening in noise task completed under divided attention demands, with the pupillary response used as a key measure of attentional resource allocation. Although the results of the studies described here provide answers to some questions, they also raised many additional questions that could not be answered in the context of this thesis. The remainder of this section focuses on limitations that are inherent to the research approach adopted in this thesis.

Processing of Speech Information

The research presented here focused on a divided attention approach to understanding attention allocation. As a result, the description of this research may appear to discount or not fully appreciate the complexity of speech processing. The FUEL was created to address the complexity of listening in noise and accounted for many more factors than are accounted for in this thesis. From a psycholinguistic perspective, the stages of perceiving and processing speech are complex. There are mappings of the perceptual representations onto semantic meaning; prelexical and phonological representations of auditory information are processed even before words are identified; syntax and semantics play a role in how the information is represented (Foss & Blank 1980; Frauenfelder & Tyler, 1987; Lieberman, 1963; Marslen-Wilson, 1984). These rich properties of speech perception and processing are important, and yet not addressed in the studies in this thesis. Rather, this thesis focused more on attention-related aspects of the broad issue of listening in noise. In an ideal world, the speech response data from these experiments would have been error coded to determine what kind of speech errors participants make under different cognitive and attentional constraints. By breaking down the responses of participants, and in particular, focusing on the speech errors, more might be understood about *how* different listening conditions and different kinds of cognitive loads affect a participant's ability to understand the presented speech. This kind of analysis would have implications for the clinical aspects of CHS where a primary goal is to understand how individuals with hearing impairments perceive auditory information and, more importantly, how their experience might be improved through a variety of interventions (Campos & Launer, 2020). Such an analysis of the types of errors people make when listening to speech in difficult listening conditions might also provide insight into how phonological information may get lost along the way, how people with hearing loss may compensate for missing information, and how aids can be provided to address the types of information that are lost.

Cross Modal Attention

Another salient aspect of the experiments presented in this thesis is the nature of the modalities used in each task. In all experiments, the primary task used auditory stimuli (sentences or words) and the secondary task used visual stimuli (either colours or numbers). These studies were not designed to address multisensory or crossmodal attention issues, but it cannot be ignored that the use of both auditory and visual information implies that attention was tapped through two sensory modalities. There is a rich research literature that addresses issues related to crossmodal attention, such as selective attention across and between modes (Koelewijn et al., 2010; Spence, 2002), and the back-and-forth interplay between attention and the integration of information across modalities (Talsma et al., 2010). Additionally, studies of crossmodal divided attention can illuminate how attentional resources are limited, and how integration and attention in different modalities can affect how one is able to perform tasks, depending on what modalities are present and how they overlap (Scerra & Brill, 2012). From this perspective, a limitation of the research in this thesis is that it cannot speak to how attention allocation may have differed if the experiments had been conducted solely in the auditory modality. If a secondary task had been presented in the auditory modality rather than the visual modality, would there have been greater dual-task costs? Would these heightened demands on cognitive resources have produced a more pronounced task overlap effect in the pupillary response?

Although research questions aimed at crossmodal attention were not the goal of this thesis, the results of my studies could have been affected by the use of two modalities rather than a single modality; yet the experimental designs used only crossmodal methods and therefore cannot establish whether effects are driven specifically by crossmodal attention processes. On the other hand, attentional demands from more than one modality do reflect realistic environments in which we complete everyday, multisensory tasks; our lives consist of constant incoming information from all modalities and attending to the demands of two modalities at the same time is a common occurrence. Further, from the perspective of CHS, it is important to consider how demands such as listening to a conversation overlap with demands such as doing a tactile-related or visual-related task, and how these overlapping task demands affect listening effort and the everyday experience of those with sensory hearing loss. As such, there are good reasons to justify my use of both auditory and visual modalities in the studies in this thesis. At the same time, not addressing the unique multisensory contribution to the results in the current experiments is a limitation of the study design and perspectives taken on the questions asked. A greater understanding of how divided attention across the same and different sensory modalities affects listening in noise would further benefit the field of CHS.

Controls and Baselines

A further limitation that stems from the design of my experiments is that I did not always include the most appropriate baseline measures or controls for comparisons. Specifically, it would have been ideal to ensure that there was always a silence condition to use as a control when studying input demands, and a single-task baseline to use as a control when studying PRP or n-back dual-task demands. Baseline measures of pupil dilation in silence, for different input demand conditions, and when performing the secondary tasks alone, would also have been useful to compare cognitive loads across conditions in the studies reported here. Furthermore, rigorous control conditions would have provided more opportunities to not only make comparisons across the experiments, but to also isolate effects of the different task demands. There are many different components to the experiments presented in this thesis and simplifying the initial designs and working up to the more complex combinations would have helped provide more basic results on which to build the more theoretically complex ideas. This is not to say that I could not make definitive statements about the effects of the different tasks on cognitive work and listening effort, but that it became a much more complex task to interpret the results. Despite this limitation in the experimental design, the results of these studies do replicate many findings from the literature—such as the effects of background noise on listening accuracy and on pupillary response, as well as secondary task performance sensitivity to the demands of the primary task. Further, these findings address the main questions regarding attention allocation and demonstrate a unique task-evoked pupil response in which input demands in a listening task interact with dual-task demands.

Future Directions

An intuitive next step from this research—in addition to addressing some of the previously mentioned limitations—is to extend these studies to larger demographics. Given that this research is adjacent to the field of CHS, addressing these questions in older adults, people with hearing loss, and people who use hearing aids would be a logical follow up. Both age-related hearing loss and other kinds of hearing loss are related to cognition and greatly affect the ability to hear and understand speech in complex listening environments. Age-related changes in hearing ability make listening in noisy environments much more difficult (Pichora-Fuller & Souza, 2003), age-related hearing loss is also intertwined with cognitive decline (Tun et al., 2012; Wayne & Johnsrude, 2015), and hearing loss in general is also associated with worse performance in memory and executive function measures and dementia (Lin et al., 2011; Lin & Albert, 2014). By extension, there is a variety of research into hearing aid and cochlear implant efficacy in difficult listening conditions, hearing assistance device use in older adults, and how hearing aids or cochlear implants affect cognitive abilities and cognitive effort (Desjardins & Doherty, 2014; Lavie et al., 2014; Lunner, 2003; Lunner et al., 2009; Shehorn et al., 2008; Stewart & Wingfield, 2009). Additionally, physiological measures, such as pupil dilation, show different response patterns in different demographics, allowing inferences about the interactions between task demands and cognitive effort in older adults (Piquado et al., 2010; Winn et al., 2018).

When I first started my PhD work, I had hoped to expand the demographics I worked with to include older adults and those with sensory hearing loss. In fact, I had designed an experiment that was scheduled to be run in the Large Interactive Virtual Environment Lab (LIVELab) at McMaster University. The LIVELab is a research and performance space that allows researchers to have complete control over many acoustic factors (along with control of many other aspects of the environment) in their experiments. The LIVELab is also equipped to collect data from many participants simultaneously and can measure a variety of physiological indices. The planned experiment was an altered version of the experiments presented in this thesis to be run in the LIVELab, which would have allowed me to collect data from multiple groups of participants and to manipulate a variety of acoustic factors of interest. This experiment was to require listening to speech in different types of background noise, and with different types of reverberation, and was to include measures of subjective listening effort, pupillary response, and EEG alpha power—a measure known to be related to cognitive and listening effort (Dimitrijevic et al., 2019; Obleser et al., 2012). Additionally, the same measures of working memory used in the presented studies in this thesis—the Operation and Reading Spans—were to be included. I intended to use the LIVELab to test groups of participants from several different demographics, including normal hearing younger adults, normal hearing older adults, younger adults who use hearing aids, and older adults who use hearing aids. Audiologists were set to be involved and run audiometric tests on all participants leading up to the experiment so there were recent audiograms to indicate hearing ability, which would have allowed for analysis by different levels and types of hearing loss.

Again, this was a large undertaking that was in the process of final scheduling and design when various interruptions led to the experiment being cancelled. The results from this experiment would have complimented the experiments presented in this thesis while fitting better into the area of CHS. I anticipated finding many of the same results—different levels of input demands affecting listening performance, subjective listening effort ratings varying as a function of listening demands, and pupillary response reflecting cognitive work. I was also interested in exploring results that would extend beyond the current thesis—how EEG patterns align with pupil dilation patterns, how different levels of hearing loss or different ages interact with different input demands, how different ages perform on working memory tasks, and how that performance relates to listening effort.

The goals of this experiment were different from those presented in this thesis. Instead of taking a purely theoretically-driven approach to understanding attention allocation when listening in noise under divided attention, this LIVELab experiment had a more clinical and applied focus. The goals were to use this unique lab space to address questions about listening demands, physiological measures of effort, and individual difference factors in a setting that can reflect everyday listening environments. The utility of the LIVELab for this purpose—aside from allowing me to test groups of participants at the same time—is that it is capable of creating sound environments that reflect real-world environments (including various specific locations such as a music hall or a busy café) without sacrificing the control and careful design of an experiment. While the divided attention approach I took in this thesis aimed at a foundational understanding of attentional resource allocation when listening in noise, looking forward it would be great to study applied issues related to listening in noise with a more diverse population than normal hearing undergraduate students. I hope this experiment is completed someday.

Conclusions

This thesis describes the first studies that combine input demands from listening in noise research with two particular dual-task methods, that track resource allocation with the pupillary response. I replicated numerous previous findings from the literature and provided strong new evidence in the pupillary response of an interaction between listening input demands and dual-task demands. These studies contribute to the field of CHS by virtue of their use of multiple divided attention methods to study attention allocation while listening in noise. Further, the results are broadly supportive of the view that listening in noise, divided attention, and the pupillary response all reflect the sharing of limited capacity attentional resources.

References

Desjardins, J. L., & Doherty, K. A. (2014). The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear and hearing*, *35*(6), 600-610.

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2019). Neural indices of listening effort in noisy environments. *Scientific Reports*, *9*(1), 1-10.

Campos, J. L., & Launer, S. (2020). From healthy hearing to healthy living: A holistic approach. *Ear and hearing*, *41*, 99S-106S.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive psychology*, *51*(1), 42-100.

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology*, *12*(1), 1-31.

Frauenfelder, U. H., & Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition*, *25*(1-2), 1-20.

Gagné, J. P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in hearing*, *21*, 2331216516687287.

Hall III, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, *23*(2), 159-165.

Hjortkjær, J., Märcher-Rørsted, J., Fuglsang, S. A., & Dau, T. (2020). Cortical oscillations and entrainment in speech processing during working memory load. *European Journal of Neuroscience*, *51*(5), 1279-1289.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kessler, Y., & Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 738.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta psychologica*, *134*(3), 372-384.

Lavie, L., Banai, K., Attias, J., & Karni, A. (2014). How difficult is difficult? Speech perception in noise in the elderly hearing impaired. *Journal of basic and clinical physiology and pharmacology*, *25*(3), 313-316.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and speech*, *6*(3), 172-187.

Lin, F. R., & Albert, M. (2014). Hearing loss and dementia–who is listening?. *Aging & mental health*, *18*(6), 671-673.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology*, *25*(6), 763.

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International journal of audiology*, *42*, S49-S58.

Lunner, T., Rudner, M., & Rönnberg, J. (2009). Cognition and hearing aids. *Scandinavian journal of psychology*, *50*(5), 395-403.

Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition: A tutorial review. *Attention and performance: Control of language processes*, 125-150.

Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, *32*(36), 12376-12383.

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science*, *1*(2), 44-48.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International journal of audiology*.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560-569.

Scerra, V. E., & Brill, J. C. (2012, September). Effect of task modality on dual-task performance, response time, and ratings of operator workload. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 1456-1460). Sage CA: Los Angeles, CA: Sage Publications.

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology*, *52*(10), 1293-1304.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*, *18*(07), 559-572.

Shehorn, J., Marrone, N., & Muller, T. (2018). Speech perception in noise and listening effort of older adults with non-linear frequency compression hearing aids. *Ear and hearing*, *39*(2), 215.

Spence, C. (2002). Multisensory attention and tactile information-processing. *Behavioural brain research*, *135*(1-2), 57-64.

Stewart, R., & Wingfield, A. (2009). Hearing loss and cognitive effort in older adults' report accuracy for verbal materials. *Journal of the American Academy of Audiology*, *20*(02), 147-154.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in cognitive sciences*, *14*(9), 400-410.

Tun, P. A., Williams, V. A., Small, B. J., & Hafter, E. R. (2012). The effects of aging on auditory processing and cognition.

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing research reviews*, *23*, 154-166.

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in hearing*, *22*, 2331216518800869.

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*, *101*, 76-86.

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*(3), 277-284.

Zylberberg, A., Oliva, M., & Sigman, M. (2012). Pupil dilation: a fingerprint of temporal selection during the “attentional blink”. *Frontiers in Psychology*, *3*, 316.

REFERENCES

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, speech, and hearing services in schools*, *49*(3), 340-355.

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International journal of audiology*, *47*(sup2), S53-S71.

Alletto, S., Palazzi, A., Solera, F., Calderara, S., & Cucchiara, R. (2016). Dr (eye) ve: a dataset for attention-based tasks with applications to autonomous and assisted driving. In *Proceedings of the ieee conference on computer vision and pattern recognition workshops* (pp. 54-60).

Anderson-Gosselin, P. A., & Gagné , J. P. (2011). Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. International Journal of Audiology, 50(11), 786–792.

Arent, S. M., & Landers, D. M. (2003). Arousal, anxiety, and performance: A reexamination of the inverted-U hypothesis. *Research quarterly for exercise and sport*, *74*(4), 436-444.

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scandinavian journal of psychology*, *50*(5), 371-384.

Attar, N., Schneps, M. H., & Pomplun, M. (2016). Working memory load predicts visual search efficiency: Evidence from a novel pupillary response paradigm. *Memory & cognition*, *44*(7), 1038-1049.

Auditec of St. Louis. (1971). Four-talker babble. 2515 S. Big Bend Boulevard, St. Louis, MO, 63143–2105.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. D. (2002). Is working memory still working?. *European psychologist*, *7*(2), 85.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model.

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*(5), 649-666.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, *91*(2), 276.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of psychophysiology*, *2*(142-162).

Bennett, K. O. C., Billings, C. J., Molis, M. R., & Leek, M. R. (2012). Neural encoding and perception of speech signals in informational masking. *Ear and Hearing*, *32*(2), 231.

Broadbent, D. E. (1958). Perception and Communication. New York: Pergamon Press.

Brouwer, W. H., Waterink, W., Van Wolffelaar, P. C., & Rothengatter, T. (1991). Divided attention in experienced young and older drivers: lane tracking and visual analysis in a dynamic driving simulator. *Human factors*, *33*(5), 573-582.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?. *Psychologica Belgica*, *51*(1), 5-13.

Campos, J. L., & Launer, S. (2020). From healthy hearing to healthy living: A holistic approach. *Ear and hearing*, *41*, 99S-106S.

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic bulletin & review*, *6*(2), 204-223.

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, *30*(7), 1086-1095.

Colflesh, G. J., & Conway, A. R. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic bulletin & review*, *14*(4), 699-703.

Conway, A. R. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, *4*(6), 577-590.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why?. *Current directions in psychological science*, *19*(1), 51-57.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive psychology*, *51*(1), 42-100.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing research*, *229*(1-2), 132-147.

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature neuroscience*, *3*(11), 1218-1223.

Desjardins, J. L., & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. Ear and Hearing, 34(3), 261–272.

Desjardins, J. L., & Doherty, K. A. (2014). The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear and hearing*, *35*(6), 600-610.

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2019). Neural indices of listening effort in noisy environments. *Scientific Reports*, *9*(1), 1-10.

Doherty, K. A., & Desjardins, J. L. (2015). The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults. *Frontiers in Psychology*, *6*, 721.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*(1), 53-78.

Dzubak, C. M. (2008). Multitasking: The good, the bad, and the unknown. *The Journal of the Association for the Tutoring Profession*, *1*(2), 1-12.

Ellenbogen, R., & Meiran, N. (2008). Working memory involvement in dual-task performance: Evidence from the backward compatibility effect. *Memory & Cognition*, *36*(5), 968-978.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*(2), 190-193.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309.

Ewing, K., & Fairclough, S. (2010). The impact of working memory load on psychophysiological measures of mental effort and motivational disposition. *Human Factors: A system view of human, technology and organisation. Maastricht: Shaker Publishing*.

Finley, J. R., Benjamin, A. S., & McCarley, J. S. (2014). Metacognition of multitasking: How well do we predict the costs of divided attention?. *Journal of experimental psychology: applied*, *20*(2), 158.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*(5509), 1803-1806.

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology*, *12*(1), 1-31.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. Memory & cognition, 43(2), 226–236.

Francis, A. L., MacPherson, M. K., Chandrasekaran, B., & Alvar, A. M. (2016). Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. *Frontiers in psychology*, *7*, 263.

Francis, A. L., & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics*, *71*(6), 1360-1374.

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in cognitive sciences*, *17*(3), 134-141.

Frauenfelder, U. H., & Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition*, *25*(1-2), 1-20.

Gagné, J. P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in hearing*, *21*, 2331216516687287.

Garland, E. L., Farb, N. A., Goldin, P. R., & Fredrickson, B. L. (2015). The mindfulness-to-meaning theory: extensions, applications, and challenges at the attention–appraisal–emotion interface. *Psychological Inquiry*, *26*(4), 377-387.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*(4), 457-461.

Groner, R., & Groner, M. T. (1989). Attention and eye movement control: an overview. *European archives of psychiatry and neurological sciences*, *239*(1), 9-16.

Hagen, J. W., & Hale, G. A. (1973). The development of attention in children.

Hall III, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, *23*(2), 159-165.

Hatfield, G. (1998). Attention in early scientific psychology. *Visual attention*, *1*, 3-25.

Heinrich, A., Ferguson, M. A., & Mattys, S. L. (2020). Effects of cognitive load on pure-tone audiometry thresholds in younger and older adults. *Ear and hearing*, *41*(4), 907.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, *45*(1), 119-129.

Heitz, R. P., Unsworth, N., & Engle, R. W. (2005). Working memory capacity, attention control, and fluid intelligence. *Handbook of understanding and measuring intelligence*, 61-77.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190-1192.

Hill, L., Williams, J. H., Aucott, L., Milne, J., Thomson, J., Greig, J., ... & Mon-Williams, M. A. R. K. (2010). Exercising attention within the classroom. *Developmental Medicine & Child Neurology*, *52*(10), 929-934.

Himi, S. A., Bühner, M., Schwaighofer, M., Klapetek, A., & Hilbert, S. (2019). Multitasking behavior and its related constructs: Executive functions, working memory capacity, relational integration, and divided attention. *Cognition*, *189*, 275-298.

Hjortkjær, J., Märcher-Rørsted, J., Fuglsang, S. A., & Dau, T. (2020). Cortical oscillations and entrainment in speech processing during working memory load. *European Journal of Neuroscience*, *51*(5), 1279-1289.

Hoeks, B., & Levelt, W. J. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, *25*(1), 16-26.

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. Ear and Hearing, 34(5), 523–534.

Hunter, C. R. (2020). Tracking cognitive spare capacity during speech perception with EEG/ERP: Effects of cognitive load and sentence predictability. *Ear and hearing*, *41*(5), 1144-1157.

Hunter, C. R., & Pisoni, D. B. (2018). Extrinsic cognitive load impairs spoken word recognition in high-and low-predictability sentences. *Ear and Hearing*, *39*(2), 378.

Hurlbert, A., & Poggio, T. (1985). Spotlight on attention. *Trends in Neurosciences*, *8*, 309-311.

Huxhold, O., Li, S. C., Schmiedek, F., & Lindenberger, U. (2006). Dual-tasking postural control: aging and the effects of cognitive demand in conjunction with focus of attention. *Brain research bulletin*, *69*(3), 294-305.

Iqbal, S. T., Ju, Y. C., & Horvitz, E. (2010, April). Cars, calls, and cognition: Investigating driving and divided attention. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1281-1290).

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394-412.

Jha, A. P., Krompinger, J., & Baime, M. J. (2007). Mindfulness training modifies subsystems of attention. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(2), 109-119.

Johnson, J., Xu, J., Cox, R., & Pendergraft, P. (2015). A comparison of two methods for measuring listening effort as part of an audiologic test battery. *American journal of audiology*, *24*(3), 419-431.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual review of psychology*, *37*(1), 43-75.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of experimental psychology: General*, *130*(2), 169.

Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336.

Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil dilation dynamics track attention to high-level information. *PloS one*, *9*(8), e102463.

Katidioti, I., Borst, J. P., & Taatgen, N. A. (2014). What happens when we switch tasks: Pupil dilation in multitasking. *Journal of experimental psychology: applied*, *20*(4), 380.

Kessler, Y., & Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 738.

Klimkeit, E. I., Mattingley, J. B., Sheppard, D. M., Lee, P., & Bradshaw, J. L. (2005). Motor preparation, motor execution, attention, and executive functions in attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *11*(2), 153-173.

Klink, K. B., Schulte, M., & Meis, M. (2012). Measuring listening effort in the field of audiology—A literature review of methods (part 2). *Zeitschrift für Audiol*, *51*(2), 60-67.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta psychologica*, *134*(3), 372-384.

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hearing research*, *312*, 114-120.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and hearing*, *33*(2), 291-300.

Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention wanders: Pupillometric signatures of fluctuations in external attention. *Cognition*, *168*, 16-26.

Krueger, M., Schulte, M., Zokoll, M. A., Wagener, K. C., Meis, M., Brand, T., & Holube, I. (2017). Relation between listening effort and speech intelligibility in noise. *American Journal of Audiology*, *26*(3S), 378-392.

Kuzminykh, A., & Rintel, S. (2020, April). Classification of functional attention in video meetings. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious?. *Perspectives on psychological science*, *7*(1), 18-27.

Lavie, L., Banai, K., Attias, J., & Karni, A. (2014). How difficult is difficult? Speech perception in noise in the elderly hearing impaired. *Journal of basic and clinical physiology and pharmacology*, *25*(3), 313-316.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and speech*, *6*(3), 172-187.

Lin, F. R., & Albert, M. (2014). Hearing loss and dementia–who is listening?. *Aging & mental health*, *18*(6), 671-673.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology*, *25*(6), 763.

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International journal of audiology*, *42*, S49-S58.

Lunner, T., Rudner, M., & Rönnberg, J. (2009). Cognition and hearing aids. *Scandinavian journal of psychology*, *50*(5), 395-403.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(2), 189.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current directions in Psychological science*, *20*(4), 240-245.

Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition: A tutorial review. *Attention and performance: Control of language processes*, 125-150.

Mathôt, S. (2018). Pupillometry: psychology, physiology, and function. *Journal of Cognition*, *1*(1).

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 471.

McCloy, D. R., Lau, B. K., Larson, E., Pratt, K. A., & Lee, A. K. (2017). Pupillometry shows the effort of auditory attention switching. *The Journal of the Acoustical Society of America*, *141*(4), 2440-2451.

McDowd, J. M., & Craik, F. I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of experimental psychology: human perception and performance*, *14*(2), 267.

Meegan, D. V., Purc-Stephenson, R., Honsberger, M. J., & Topan, M. (2004). Task analysis complements neuroimaging: an example from working memory research. *Neuroimage*, *21*(3), 1026-1036.

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., De Lissa, P., Graham, P., & Lyxell, B. (2017). Objective assessment of listening effort: Coregistration of pupillometry and EEG. *Trends in hearing*, *21*, 2331216517706396.

Miller, A. L., Gross, M. P., & Unsworth, N. (2019). Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation. *Journal of Memory and Language*, *104*, 25-42.

Mitterer, H., & Mattys, S. L. (2017). How does cognitive load influence speech perception? An encoding hypothesis. *Attention, Perception, & Psychophysics*, *79*(1), 344-351.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta psychologica*, *27*, 84-92.

Morton, L. L., Kershner, J. R., & Siegel, L. S. (1990). The potential for therapeutic applications of music on problems related to memory and attention. *Journal of Music Therapy*, *27*(4), 195-208.

Murphy, D. R., Craik, F. I., Li, K. Z., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and aging*, *15*(2), 323.

Naber, F. B., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Dietz, C., Van Daalen, E., Swinkels, S. H., ... & Van Engeland, H. (2008). Joint attention development in toddlers with autism. *European Child & Adolescent Psychiatry*, *17*(3), 143-152.

Neagu, M. B., Dau, T., Hyvärinen, P., Bækgaard, P., Lunner, T., & Wendt, D. (2019). Investigating pupillometry as a reliable measure of individual's listening effort. In *Proceedings of the International Symposium on Auditory and Audiological Research* (Vol. 7, pp. 365-372).

Ng, E. H., Rudner, M., Lunner, T., & Rönnberg, J. (2015). Noise reduction improves memory for target language speech in competing native but not foreign language speech. Ear and Hearing, 36(1), 82–91.

Nieborowska, V., Lau, S. T., Campos, J., Pichora-Fuller, M. K., Novak, A., & Li, K. Z. (2019). Effects of age on dual-task walking while listening. *Journal of Motor Behavior*, *51*(4), 416-427.

Ninio, A., & Kahneman, D. (1974). Reaction time in focused and in divided attention. *Journal of Experimental Psychology*, *103*(3), 394.

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *The quantitative methods for psychology*, *10*(1), 56-67.

Ocasio, W. (2011). Attention to attention. *Organization science*, *22*(5), 1286-1296.

Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, *32*(36), 12376-12383.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, *25*(1), 46-59.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior research methods*, *47*(4), 1343-1355.

Padilla, L. M., Castro, S. C., Quinan, P. S., Ruginski, I. T., & Creem-Regehr, S. H. (2019). Toward objective evaluation of working memory in visualizations: A case study using pupillometry and a dual-task paradigm. *IEEE transactions on visualization and computer graphics*, *26*(1), 332-342.

Park, S., Kim, M. S., & Chun, M. M. (2007). Concurrent working memory load can facilitate selective attention: evidence for specialized load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1062.

Park, S. Y., Kim, M. J., Sikandaner, H., Kim, D. K., Yeo, S. W., & Park, S. N. (2016). A causal relationship between hearing loss and cognitive impairment. *Acta oto-laryngologica*, *136*(5), 480-483.

Pashler, H. (1984). Processing stages in overlapping tasks: evidence for a central bottleneck. *Journal of Experimental Psychology: Human perception and performance*, *10*(3), 358.

Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, *21*(4), 469-514.

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science*, *1*(2), 44-48.

Pashler, H. (1994a). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Pashler, H. (1994b). Graded capacity-sharing in dual-task interference?. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 330.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Pschophysiology, 11*(5), 559-566.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International journal of audiology*.

Picou, E. M., & Ricketts, T. A. (2014). The effect of changing the secondary task in dual-task paradigms for measuring listening effort. Ear and Hearing, 35(6), 611–622.

Piquado, T., Cousins, K. A., Wingfield, A., & Miller, P. (2010). Effects of degraded sensory input on memory for speech: Behavioral data and a test of biologically constrained computational models. *Brain research*, *1365*, 48-65.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560-569.

Ponds, R. W., Brouwer, W. H., & Van Wolffelaar, P. C. (1988). Age differences in divided attention in a simulated driving task. *Journal of Gerontology*, *43*(6), P151-P156.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Rao & Starkey (n.d.) Speech Reception Threshold Program [computer software]. Starkey Hearing Industries

Redick, T. S., & Lindsey, D. R. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic bulletin & review*, *20*(6), 1102-1113.

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247-261.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*, *7*, 31.

Ruthruff, E., & Pashler, H. (2001). Perceptual and central interference in dual-task performance. *The limits of attention: Temporal constraints in human information processing*, 100-123.

Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement?. *Psychonomic bulletin & review*, *8*(1), 73-80.

Scerra, V. E., & Brill, J. C. (2012, September). Effect of task modality on dual-task performance, response time, and ratings of operator workload. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 1456-1460). Sage CA: Los Angeles, CA: Sage Publications.

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology*, *52*(10), 1293-1304.

Schmidt, M. E., & Vandewater, E. A. (2008). Media and attention, cognition, and school achievement. *The Future of children*, 63-85.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*, *18*(07), 559-572.

Semple, R. J. (2010). Does mindfulness meditation enhance attention? A randomized controlled trial. *Mindfulness*, *1*(2), 121-130.

Shehorn, J., Marrone, N., & Muller, T. (2018). Speech perception in noise and listening effort of older adults with non-linear frequency compression hearing aids. *Ear and hearing*, *39*(2), 215.

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*(1), 489-499.

Shipstead, Z., Lindsey, D. R., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116-141.

Sibley, C., Coyne, J., & Baldwin, C. (2011, September). Pupil dilation as an index of learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 237-241). Sage CA: Los Angeles, CA: SAGE Publications.

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., ... & Schooler, J. W. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PloS one*, *6*(3), e18298.

Somberg, B. L., & Salthouse, T. A. (1982). Divided attention abilities in young and old adults. *Journal of Experimental Psychology: human perception and performance*, *8*(5), 651.

Spence, C. (2002). Multisensory attention and tactile information-processing. *Behavioural brain research*, *135*(1-2), 57-64.

Stewart, R., & Wingfield, A. (2009). Hearing loss and cognitive effort in older adults' report accuracy for verbal materials. *Journal of the American Academy of Audiology*, *20*(02), 147-154.

Sullivan, L. (1976). Selective attention and secondary message analysis: A reconsideration of Broadbent's filter model of selective attention. *Quarterly Journal of Experimental Psychology*, *28*(2), 167-178.

Svetina, M. (2016). The reaction times of drivers aged 20 to 80 during a divided attention driving. *Traffic injury prevention*, *17*(8), 810-814.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in cognitive sciences*, *14*(9), 400-410.

Tang, Y. Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews Neuroscience*, *16*(4), 213-225.

Tombu, M., & Jolicœur, P. (2002). All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. *Psychological research*, *66*(4), 274-286.

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 3.

Tomporowski, P. D., & Audiffren, M. (2014). Dual-task performance in young and older adults: speed-accuracy tradeoffs in choice responding while treadmill walking. *Journal of aging And physical activity*, *22*(4), 557-563.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248.

Treisman, A. M. (1964). Selective attention in man. *British medical bulletin*, *20*(1), 12-16.

Treisman, A., & Geffen, G. (1967). Selective attention: Perception or response?. *Quarterly Journal of Experimental Psychology*, *19*(1), 1-17.

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological review*, *95*(1), 15.

Tun, P. A., Williams, V. A., Small, B. J., & Hafter, E. R. (2012). The effects of aging on auditory processing and cognition.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757-765.

Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. *Attention, Perception, & Psychophysics*, *80*(2), 461-484.

Valentine, E. R., & Sweet, P. L. (1999). Meditation and attention: A comparison of the effects of concentrative and mindfulness meditation on sustained attention. *Mental health, religion & culture*, *2*(1), 59-70.

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review*, *25*(6), 2005-2015.

Van Gerven, P. W., Paas, F., Van Merriënboer, J. J., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167-174.

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing research reviews*, *23*, 154-166.

Welford, A. T. (1952). The 'psychological refractory period'and the timing of high-speed performance—a review and a theory. *British Journal of Psychology. General Section*, *43*(1), 2-19.

Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., & Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hearing research*, *369*, 67-78.

Wickens, C. D. (1976). The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 1.

Wilhelm, O., Hildebrandt, A. H., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it?. *Frontiers in psychology*, *4*, 433.

Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech*, *37*(3), 221-235.

Wingfield, A., Amichetti, N. M., & Lash, A. (2015). Cognitive aging and hearing acuity: Modeling spoken language comprehension. *Frontiers in Psychology*, *6*, 684.

Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *Journal of the American Academy of Audiology*, *18*(07), 548-558.

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in hearing*, *22*, 2331216518800869.

Wu, Y. H., Stangl, E., Zhang, X., Perkins, J., & Eilers, E. (2016). Psychometric functions of dual-task paradigms for measuring listening effort. Ear and Hearing, 37(6), 660–670.

Xia, J., Nooraei, N., Kalluri, S., & Edwards, B. (2015). Spatial release of cognitive load measured in a dual-task paradigm in normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America, 137(4), 1888–1898.

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*, *101*, 76-86.

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*(3), 277-284.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and hearing*, *31*(4), 480-490.

Zekveld, A. A., Kramer, S. E., Rönnberg, J., & Rudner, M. (2019). In a concurrent memory and auditory perception task, the pupil dilation response is more sensitive to memory load than to auditory stimulus characteristics. *Ear and hearing*, *40*(2), 272.

Zhivomirov, H. (2018). A method for colored noise generation. *Romanian journal of acoustics and vibration*, *15*(1), 14-19.

Zylberberg, A., Oliva, M., & Sigman, M. (2012). Pupil dilation: a fingerprint of temporal selection during the "attentional blink". *Frontiers in Psychology*, *3*, 316.