

THE IRRELEVANT SOUND EFFECT:
SIMILARITY OF CONTENT OR SIMILARITY OF PROCESS?

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
The Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Zachary Adam Schendel, M.A.

The Ohio State University

2006

Dissertation Committee:

Dr. Neal Johnson, Adviser

Dr. David Huron

Dr. Richard Jagacinski

Dr. Mari Jones

Approved By

Adviser

Graduate Program in Psychology

Copyright by
Zachary Adam Schendel
2006

ABSTRACT

The purpose of this dissertation was twofold: The general purpose was to further investigate the properties of the phonological loop by extending the irrelevant sound effect paradigm to music. The specific purpose was to investigate the similarity of content (Salame & Baddley, 1982; 1989) and similarity of process (Jones & Macken, 1993; 1995) hypotheses of working memory in order to determine which could best account for the irrelevant sound effects in both language and music. Experiment I consisted of serial recall of visually presented nine-digit (selected from the digits 1-9) or nine-note (selected from the notes C4-E5) sequences in the presence of silence, instrumental music, vocal music, or Arabic speech. Digit results supported the similarity of content hypothesis: the more speech-like the irrelevant sound was, the worse recall was, which replicated Salame and Baddeley (1989). Note results provided mixed results. It was hypothesized that visual notes are memorized in a variety of ways which may be influenced by musical experience and the memorization context (Schendel & Palmer, in press).

Experiment IIA used a six-digit or four-tone standard-comparison task. The digit trials were set up with a visual-standard/auditory-comparison procedure and the tone trials were set up as an auditory-standard/auditory-comparison procedure so as to force participants to rehearse in an auditory form. The same irrelevant sound conditions as in Experiment I were used. Both digit and tone results supported the similarity of content

hypothesis: reater overlap between the irrelevant sound and the rehearsed information resulted in greater performance decrements. Experiment IIB used only digits and was the same as the digit half of Experiment IIA, but the irrelevant sound conditions were a between participant factor. Results of Experiment IIB replicated the results of Experiment IIA.

Experiment III used an auditory-standard/auditory-comparison procedure with both six-digit and four-tone sequences. The irrelevant sound conditions were silence, low-overlap, and high-overlap irrelevant sound. For the digit trials, high-overlap irrelevant sound consisted of a random series of spoken digits in the same range as the stimuli (1-9), while low-overlap irrelevant sound consisted of a series of nine random spoken words that did not rhyme with the digits 1-9. For the tone trials, high-overlap irrelevant sound consisted of a random series of auditory piano tones in the same range as the stimuli (C4-E5), while low-overlap irrelevant sound consisted of a random series of nine piano tones which were members of the C# Major scale and outside of the stimulus range. In both stimulus conditions the low-overlap condition resulted in significant interference compared to the silent control, and the high-overlap condition resulted in significantly more interference than the low-overlap condition. The results supported the similarity of content hypothesis.

It was concluded that the main theory behind the similarity of content hypothesis should be central to any theory of auditory working memory: an increased acoustic overlap between the to-be-remembered and to-be-rehearsed information results in

increased performance decrements. The results of all experiments also supported one of the components of the similarity of process hypothesis: serial auditory working memory is particularly vulnerable to irrelevant sound. Finally, the results supported the new idea that working memory for language and working memory for music are both governed by a single acoustic loop. One such new model was proposed (the Acoustic Overlap hypothesis) which includes key components of the similarity of content and similarity of process hypotheses as well as new ideas based on these experiments.

Dedicated to my Mom, my Dad, and Amy

ACKNOWLEDGMENTS

I would like to thank Dr. Neal Johnson for taking me on as his graduate student in 2003 and helping me get to the finish line. I would also like to thank Dr. Caroline Palmer for her continuing support from Montreal. You both have been a great inspiration. A big thanks goes out to my dissertation committee, Dr. Neal Johnson, Dr. David Huron, Dr. Richard Jagacinski, and Dr. Mari Jones. Thanks for your insightful comments on my dissertation.

A sincere thanks also goes out to the Sequence Production Lab at McGill University for their help. Especially Sean Hutchins, Janeen Loehr, as well as Christine Capota, Anjali Bhatara, Ricky, Julian, and Bubbles for their sanity-inducing hilarious takes on sharp fonts, didgeridoos, bush pies, and lab alert levels. Also, thanks to my domestic friends for their years of support in and outside of Columbus. Especially the Jungers, Erik Tracy, Rob Ellis, the soon-to-be Burkes, and the basketball crew. Thanks go to Andy Webb and Corey White, my roommates, for the times we could squeeze in various nonacademic enjoyable activities such as video games, pizza, sports, and watching really bad movies. And, thanks to my friends outside of Ohio – the Beilsteins, the soon-to-be Allens, and the Bowens for continuing to be great friends despite the distance.

I would also like to thank my family for putting up with me for 27 years, reading every draft of anything I ever wrote, and pushing me to continue doing all of this when I

wanted to quit. But, most importantly, thanks for always having time to play catch and to help finish the crossword puzzle. Thanks to the extended Watts' family. I am so glad I met your daughter/cousin/niece/sister Amy, it has been wonderful getting to know you and becoming a fixture at your gatherings. A special thanks goes out to Deward, Linda, and Andrew Watts for welcoming me into their family – I could not ask for a better second family to be a part of. Another special thanks goes out to Tootsie Bond for being such a great friend even if that involves cutting your grass and rewiring your TV every week.

To Scott, Rosa, Radek, Caesar, and Nietzsche Jankowski – thanks for all the cookbooks, dinners, and putting up with watching 24. Scott, dude you rule. Even though we both are jobless coming out of grad school, at least we have a bunch of million dollar ideas to fall back on.

Thanks to Fritz for at least demonstrating the proper way to relax. If only sleeping 18 hours a day was an option. Finally, thanks to me dearest Amy Watts for believing in me every step of the way. Your constant love has always amazed me. I cannot wait for September, 16th.

VITA

October 13, 1978.....Born – Alexandria, Virginia
2001.....B.A. Psychology, St. Olaf College
2001.....B.A. Music, St. Olaf College
2001- presentGraduate Research and Teaching Assistant,
The Ohio State University (OSU)
2003.....M.A. Cognitive Psychology, OSU

PUBLICATIONS

Research Publications

1. Schendel, Z.A. & Palmer, C. (in press). Suppression effects on musical and verbal memory. *Memory & Cognition*.
2. Schendel, Z. A. (1999). Tests of eyewitness recall and recognition. *Journal of Undergraduate Sciences*, 5, 70-74.

FIELDS OF STUDY

Major Field: Cognitive Psychology

TABLE OF CONTENTS

	<u>Page</u>
Abstract.....	ii
Dedication.....	v
Acknowledgments.....	vi
Vita.....	viii
List of Figures.....	xi
Chapters:	
1. Introduction:	
.....	1
1.1 Working Memory and Language.....	2
1.2 The Articulatory Suppression Effect and Language.....	3
1.3 The Irrelevant Sound Effect and Language.....	4
1.3.1 The Similarity of Content Hypothesis.....	4
1.3.2 The Similarity of Process Hypothesis.....	9
2. Working Memory and Music	18
2.1 The Articulatory Suppression Effect and Music.....	20
2.2 The Irrelevant Sound Effect and Music.....	23
2.3 Purpose.....	2
5	
3. Experiment I.....	27
3.1 Participants.....	28
3.2 Materials and Apparatus.....	28
3.3 Design.....	29
3.4 Procedure.....	30
3.5 Results.....	30

3.6	Discussion.....	37
4.	Experiment IIA.....	40
4.1	Participants.....	41
4.2	Materials and Apparatus.....	42
4.3	Design.....	43
4.4	Procedure.....	43
4.5	Results.....	44
4.6	Discussion.....	50
5.	Experiment IIB.....	53
5.1	Participants.....	53
5.2	Materials and Apparatus.....	54
5.3	Design.....	54
5.4	Procedure.....	54
5.5	Results.....	54
5.6	Discussion.....	57
6.	Experiment III.....	58
6.1	Participants.....	59
6.2	Materials and Apparatus.....	60
6.3	Design.....	61
6.4	Procedure.....	61
6.5	Results.....	62
6.6	Discussion.....	66
7.	General Discussion.....	68
7.1	Similarity of Content Hypothesis.....	70
7.2	Similarity of Process Hypothesis.....	73
7.3	The Acoustic Overlap Hypothesis: Towards an Integrated Theory.....	76
7.4	Summary.....	79
	Bibliography.....	80
	Appendix A: Background Questionnaire.....	84

Appendix B: Music Pre-test.....86

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
3.1	Experiment I: Mean percent correct on a nine-item serial recall task with visual digits and notes in the presence of silence, instrumental-music, vocal-music, and Arabic-speech irrelevant sound.....	32
3.2	Experiment I: Mean percent correct on a nine-item serial recall task with visual digits and notes in the presence of silence, instrumental-music, vocal-music, and Arabic-speech irrelevant sound.....	36
4.1	Experiment IIA: Mean percent correct on a visual-auditory digit and auditory-auditory tone recognition task in the presence of silence, instrumental-music, vocal-music, and Arabic-speech irrelevant sound.....	46
5.1	Experiment IIB: Mean percent correct on a visual-auditory digit recognition task in the presence of silence, instrumental-music, vocal-music, and Arabic-speech irrelevant sound.....	55
6.1	Experiment III: Mean percent correct on auditory standard/comparison task with digits and tones in the presence of silence, low-overlap irrelevant sound, and high-overlap irrelevant sound.....	63

CHAPTER 1

INTRODUCTION

Research in the field of working memory has been primarily concerned with the language domain (Murray, 1967). Some research has begun to investigate how working memory for music may relate to working memory for language (Berz, 1995; Pechmann & Mohr, 1992; Schendel & Palmer, in press). The current research was grounded in the working memory model of Baddeley and Hitch (1974). The general purpose of this project was to use working memory research in language as a foundation and to further explore its applications in the musical domain. The specific purpose was to research the effect of irrelevant sound on the working retention of language and music.

First, Baddeley and Hitch's (1974) working memory model will be described. A discussion of two experimental paradigms - the articulatory suppression effect paradigm and the irrelevant speech effect paradigm - used to flesh out the functions of the phonological loop will follow (Colle & Welsh, 1976; Murray, 1967; 1968). Then, additional research will be cited that has begun to shed some light on the concept of a working memory model that is responsible for both linguistic and musical materials (Berz, 1995; Deutsch, 1970; Jones & Macken, 1993; Salame & Baddeley, 1982; 1989; Schendel & Palmer, in press).

Finally, a series of experiments that will further investigate the hypothesized function of the phonological store component of the working memory model from a language and music perspective will be proposed. Specifically, the experiments will investigate whether the characteristic breakdown patterns found in serial memory for language and music in the presence of irrelevant sound are a result of process similarities (Beaman & Jones, 1997; Jones & Macken, 1993; Jones, Macken, & Nicholls, 2004) or content similarities (Elliot & Cowan; 2005; Salame & Baddeley, 1982; 1989).

WORKING MEMORY AND LANGUAGE

Baddeley and Hitch's (1974) working memory model is a multi-component system. A central executive controls the system and delegates short-term storage information into two slave systems: a visuo-spatial sketchpad and the phonological loop (Baddeley, 1999). The visuo-spatial sketchpad specializes in working memory for visual information while the phonological loop specializes in working memory for auditory information (Baddeley & Hitch, 1994).

The phonological loop has two components: the articulatory control process and the phonological store (Baddeley, 1990). The functions of the articulatory control process are to translate a visual signal (the written word) into a phonological signal (articulated word), and to refresh the memorial signal through rehearsal. Visual-verbal material only has access to the phonological store through the articulatory control process. Auditory information does not need to be translated; it has obligatory direct access to the

phonological store. In order for auditory information to remain longer than two seconds in working memory, rehearsal is required (Baddeley, 1986; 1992). The purpose of rehearsal is to refresh a fading auditory signal and allow material to be kept in the phonological store for longer periods.

THE ARTICULATORY SUPPRESSION EFFECT AND LANGUAGE

The articulatory suppression effect is seen when participants are required to speak while memorizing lists (Levy, 1975). Murray (1967; 1968) extensively explored articulatory suppression in two experiments using both visual and auditory presentation of to-be-remembered lists. He found that recall of sequences in both presentation modalities was impaired by concurrent suppression, but recall of visual sequences were more impaired than recall of auditory sequences. Baddeley (1990) subsequently interpreted these results in terms of the working memory model. According to Baddeley, articulatory suppression ties up the articulatory control process which impairs the translation from visual code to phonological code and rehearsal of to-be-remembered information. While articulatory suppression also impairs performance on auditory presentation trials, the interference is not as severe since the auditory information already has direct access to the phonological store.

THE IRRELEVANT SOUND EFFECT AND LANGUAGE

The irrelevant sound effect occurs when participants' recall of lists is impaired by the presence of to-be-ignored irrelevant sound (Colle & Welsh, 1976; Salame & Baddeley, 1982; 1989). Colle and Welsh (1976) found that when participants were required to memorize random, non-rhyming letter sequences in the presence of silence or irrelevant auditory German text, their immediate serial recall was significantly impaired in the German text condition. It was further discovered that irrelevant auditory noise did not impair recall in the same task (Colle, 1980). Therefore, the irrelevant sound effect was not attributed to general auditory distraction.

THE SIMILARITY OF CONTENT HYPOTHESIS

Salame and Baddeley (1982; 1989) utilized all types of irrelevant sound in order to discover what other types of auditory information would result in the irrelevant speech effect. Using a series of nine-digit immediate serial recall tasks in the presence of irrelevant sound, Salame and Baddeley (1982; 1989) found that monosyllabic English words, monosyllabic nonsense words, five-syllable English words, and Arabic text all resulted in similar performance decrements. In another experiment, they found similar disruption by digits and monosyllabic nonsense syllables, which were phonologically equivalent to the digits (e.g., tun, gnu, tee). In this experiment, phonologically dissimilar

disyllabic words (e.g., tennis, jelly) also impaired performance more than silence, but not as severely as digits or the monosyllabic nonsense words that were phonologically equivalent to digits. This led them to conclude that the irrelevant sound effect was not a result of meaning or semantics because nonsense items have no meaning, and participants did not have any experience with Arabic (Salame & Baddeley, 1982; 1989).

Salame and Baddeley (1989) went on to investigate whether irrelevant non-speech auditory stimuli like music or noise would have similar effects on the serial recall of digits. To test for this effect, they used an immediate visual nine-digit serial-recall task in the presence of silence, instrumental, and vocal music. Serial recall of digits in the presence of vocal music and instrumental music was statistically worse than silence, and serial recall of digits in the presence of vocal music was significantly worse than instrumental music. Furthermore, Salame and Baddeley (1989) conducted another experiment in which Arabic speech, instrumental music, and silence were used as the to-be-ignored background conditions. Similar to the first experiment, both Arabic speech and instrumental music resulted in significantly worse recall when compared to the silent-control, but Arabic speech was significantly worse than instrumental music. The more speech-like (1989) or digit-like (1982) the to-be-ignored stimulus was, the more it interfered with the memorization of digit sequences.

This breakdown pattern seems to be specific to speech or speech-like stimuli. Salame and Baddeley (1982; 1987; 1989) found that silence, noise bursts presented before each visual to-be-remembered item, and continuous pink noise did not cause significantly different amounts of interference compared to control conditions while all caused

significantly less interference than any speech-like stimuli. Additionally, Salame and Wittersheim (1978) found that interference caused by continuous pink noise (like white noise, but lower frequency bands are more powerful resulting in a lower-pitched hiss), noise present with each digit presentation, and noise presented between each digit presentation in a serial recall task was only found in early serial positions. Salame and Wittersheim (1978) hypothesized that noise coinciding with early-list presentation affected attention, not working memory. If the noise affected working memory, the noise would consistently impair performance across all serial positions.

Irrelevant speech specifically affects the phonological store component of the working memory model. Evidence that irrelevant speech is specific to working memory as opposed to short-term or long-term memory comes from Salame and Baddeley (1990). They employed a free-recall procedure with 16-word lists in silence or in the presence of irrelevant speech. They found that primacy and recency effects indicative of short-term and long-term memory systems remain intact in the presence of irrelevant speech. They also found no difference between performance in silence as compared to performance in the presence of irrelevant speech. Salame and Baddeley (1990) concluded that it was not the semantic information that was affected by irrelevant speech, but the code of the to-be-remembered input while it was in working memory. This conclusion was supported by Miles, Jones, and Madden (1991) who found that irrelevant speech during either visual digit presentation or a rehearsal interval resulted in the same performance decrement; this result could only have occurred if the locus of the disruption was on memory rather than on encoding.

Salame and Baddeley (1989) developed two hypotheses to account for their irrelevant speech effect findings: the Noise Filter and Speech Detector hypotheses. The Noise Filter hypothesis is a two-stage hypothesis in which less speech-like auditory information is blocked from entering the phonological store by a filter in stage one. In stage two, the more phonological features the to-be-remembered and to-be-ignored information have in common, the more interference occurs. The speech-detector hypothesis states that the phonological store is more sensitive to sounds that have features in common with speech. The more phonological features the irrelevant sounds have in common with speech, the more interference will occur.

Elliot and Cowan (2005) conducted an extensive series of experiments investigating the irrelevant sound effect and individual differences in working memory span. Six experiments ($n = 205$) were run in which participants immediately recalled visually-presented eight-digit serial lists. They performed the task in silence or in the presence of irrelevant tones or words (e.g., red, blue, tall). In averaging the effects across all the experiments, Elliot and Cowan (2005) found that participants performed most accurately in silence, significantly more poorly in the presence of irrelevant tones, and significantly still more poorly in the presence of irrelevant speech. This result is similar to Salame and Baddeley's (1989) finding in which instrumental music impaired performance but not as severely as Arabic speech.

Elliot and Cowan (2005) went on to conduct a series of experimental correlations. They found that the correlation of the performance decrement in the tone condition (performance in silence minus performance in tone condition) with the

performance decrement in the speech condition (performance in silence minus performance in speech condition) was significant ($r = .59$). This correlation also occurred independent of the participant's span level indicating it was not a function of working memory capacity. They concluded that this finding "suggests a role for a common mechanism of disruption regardless of the nature of the irrelevant sound" (p. 674). This conclusion seems to support the notion that greater overlap between the to-be-remembered and the to-be-ignored stimuli results in greater performance decrements. It also raises questions about the hypothesized "phonological" nature of the "phonological" store since irrelevant speech and irrelevant music both interfered with similar processes.

In summary, the similarity of content hypothesis was developed based entirely on experiments in which participants were memorizing speech sequences in the presence of irrelevant sound. The similarity of content hypothesis is based on the phonological loop model and reads as follows: "It is the degree of phonological similarity between the irrelevant material and the memory items that underlies the irrelevant speech effect" (Gathercole and Baddeley, 1993, p. 13). As will be later discussed, serious theoretical problems have arisen based on working memory experiments in which music is used as the to-be-remembered stimulus (Schendel & Palmer, in press). The validity of this hypothesis will be further tested in the current experiments by investigating the effects of different irrelevant-sound conditions on performance decrements in the memorization of musical sequences.

THE SIMILARITY OF PROCESS HYPOTHESIS

Recently, the similarity of content hypothesis has been a hotbed for theoretical debate in the literature (Baddeley & Larsen, 2003; Larsen & Baddeley, 2003; Macken & Jones, 2003). In the last 15 years, a series of experiments have been conducted which have been claimed to support a hypothesis alternative to the similarity of content hypothesis proposed by Salame and Baddeley (1982; 1989). Four lines of experiments will be discussed in this section: (1) Jones and Macken 1993; (2) Jones and Macken, 1995; (3) Hughes, Tremblay, and Jones, 2005; and (4) Bridges and Jones, 1996. These experiments are all said to be the seminal research supporting an alternative acoustic working memory theory called the Object-Oriented Episodic Record (O-OER) hypothesis otherwise known as the similarity of process hypothesis.

First, Jones and Macken (1993) ran an experiment in which serial recall of visual seven-consonant sequences was tested in the presence of a single repeating irrelevant tone or four changing irrelevant tones. The four-tone condition produced significant interference, and the single tone condition was not significantly different from silence. Jones and Macken (1993) went on to demonstrate that the performance decrement on the task in the presence of a repeating four-tone condition, while greater than a silent condition, was no different statistically than that of a single syllable "ah" spoken on different pitches. They concluded that it was not a differential phonological overlap between the to-be-remembered and to-be-ignored stimulus that resulted in the irrelevant

sound effect, but rather the changing state of the auditory sequence (the changing-state hypothesis).

The key difference between the Salame and Baddeley (1989) and Jones and Macken (1993) experiments lies in the results of the post hoc analyses. Salame and Baddeley (1989) found post-hoc statistical differences between their interference conditions using a Newman-Kuels procedure, while Jones and Macken (1993) did not find post-hoc differences using the same procedure. While there may not have been statistically significant post hoc differences between each and every interference condition in these two experiments, it is also important to take into account the general trends of interference. The similarity of content explanation would predict that an increased phonological similarity between the to-be-remembered and to-be-ignored stimuli would result in increased interference. The changing state hypothesis would predict an interference pattern which impairs performance to the degree of dissimilarity among the to-be-ignored stimulus items. It is unclear whether the stimuli used in this experiment were dissimilar enough to result in differences in interference.

When one looks at every experiment from Salame and Baddeley (1989) and Jones and Macken (1993) the relative impact of irrelevant sound was much more characteristic of the similarity of content explanation. Depending on the nature of the to-be-remembered stimulus, there was a consistent interference continuum, where greater overlap resulted in greater interference. While not all of these differences were statistically significant in the post hoc analyses, it is important to note that the interference was not random. Instead it showed a marked and predictable trend in both sets of experiments. It

is unclear whether Jones and Macken's (1993) experiments resulted in null effects of interference conditions because the irrelevant-sound conditions were truly equipotent or because the conditions were so acoustically similar that no performance differences would logically be predicted.

Second, Jones and Macken (1995) further investigated the similarity of process hypothesis in a series of serial recall experiments with various irrelevant-sound conditions. In Experiment I, participants were required to serially recall visually-presented seven-item letter lists in the presence of an irrelevant random spoken list of the same seven letters, or an irrelevant spoken random list of seven different letters. These two conditions were compared to a silent-control condition. They found no differences in the two auditory irrelevant-sound conditions until the last three serial positions (five, six, and seven). In positions six and seven, the same-letters irrelevant sound resulted in greater interference than the different-letter irrelevant sound, and both resulted in greater interference than the silent-control.

In the same study, Jones and Macken (1995) ran an additional experiment using a visually-presented nine-digit serial-recall task in the presence of irrelevant sound. The irrelevant-sound conditions in this experiment were items phonologically similar to digits (Salame & Baddeley, 1982: e.g. tun, gnu, tee...), phonologically dissimilar two-syllable words (e.g., tennis, jelly, double...), and phonologically dissimilar one-syllable words (e.g., bed, sap, pick...). They found the same results as in Experiment I: the two irrelevant-sound conditions were only different from each other in the last three serial positions. The phonologically similar nonwords resulted in the greatest performance

decrements in the least three serial positions, followed by the phonologically dissimilar words, and the silent-control.

Based on the finding that only the information in the recency portion of the serial recall curve is affected by different types of irrelevant sound, they conclude that “phonological similarity...plays only a minor role in the irrelevant-[sound] effect.” (Jones and Macken, p. 114). The results and conclusions of these experiments are subject to major theoretical criticisms in light of historical memory findings.

The typical j-shaped recall curve has three components – a primacy portion, a middle portion, and a recency portion regardless of list length (Murdock, 1962). The three components are the direct functions of the interaction between rehearsal and different memory systems. Rehearsal of to-be-remembered items has two functions: to maintain information in a rehearsal buffer, and to increase the likelihood that the rehearsed information is transferred to long-term-memory for more permanent storage (Ashcraft, 1994; Atkinson & Shiffrin, 1968). The first few items (where there was no effect in the Jones & Macken, 1995 studies) in a list are rehearsed the most, and therefore, transfer to long-term memory occurs. This transfer to long-term memory results in the primacy effect. There is no change in the primacy effect even after long interference-laden retention intervals which are known to interfere with short-term memory (Glanzer & Cunitz, 1966). Also, allowing an increased amount of time for rehearsal between item presentations significantly improves the primacy effect (Glanzer, & Cunitz, 1966). The middle portion of the serial-position curve is not altered by these manipulations. Very

little of the middle section is recalled on average because of the significant decrease in rehearsals per item (Rundus, 1971).

The recency portion of the serial-position curve, where the effects were found in the Jones and Macken (1995) studies, is the short-term memory portion of the curve. If participants are required to do mental arithmetic during a retention interval following list presentation, the recency effect disappears (Glanzer & Cunitz, 1966). Also, giving participants extra time to rehearse between list items does not affect the recency portion (Glanzer & Cunitz, 1966). But, mental arithmetic does not remove the primacy effect while allowing for increased rehearsal time does increase the primacy effect. Irrelevant attention-demanding rehearsal does not impair long-term memory, and increasing rehearsal time allows for a stronger long-term memory representation. The recency portion of the serial-recall curve would theoretically be the most important part for Jones and Macken (1995) to investigate. On the other hand, the primacy portion is not important for the theoretical conclusions they are attempting to form since they are attempting to counter a claim by Baddeley et. al. concerned with working memory. Based on this information, it could be concluded that their findings might actually support the similarity of content hypothesis.

Nonetheless, following these two initial experiments, Jones and Macken (1995) have indicated that a different model was needed to explain their results. They developed an alternative to the similarity of content hypothesis which they termed the Object-Oriented Episodic Record (O-OER) model. For the purposes of this discussion, the O-OER model will be referred to as the similarity of process hypothesis. The similarity of

process hypothesis states that the to-be-remembered sequence is stored in working memory in terms of object linkages. The to-be-ignored stream is also represented as object linkages. They conclude that the irrelevant sound effect occurs because of competition between order cues bonding the to-be-remembered stimuli and order cues present in the changing state irrelevant sound. Indeed, memory for serial order has been shown to be particularly impaired by irrelevant sound (Beaman & Jones, 1997; Jones & Macken, 1993; Salame & Baddeley, 1990).

Jones and Macken (1995) believe that a phonological *dissimilarity* effect is the best explanation for the interference by an irrelevant stream. The greater the degree to which objects in the irrelevant sound stream are distinguishable from each other, the greater interference will result. The items that make up the irrelevant sound also have a specific order, and each component or object is bound with serial order linkages. It is the confusion that occurs between the to-be-remembered object linkages and the irrelevant linkages that form between the to-be-remembered sequence and to-be-ignored irrelevant sound sequence that result in order confusions. The more distinguishable the objects that make up the irrelevant sound are, the stronger the bonds between the objects. Stronger bonds interfere more with the to-be-remembered objects' bonds, and more errors result.

This theory was demonstrated in a third set of experiments by Hughes, Tremblay, and Jones (2005) who have recently begun to investigate the nature of different changes in stimuli and how these changes act as interference in irrelevant sound experiments. They found that a repeated consonant-vowel-consonant (CVC) sequence (e.g., vik, vik, vik...) heard as irrelevant sound is sufficient to cause performance decrements in the recall of

nine-letter sequences. Changes in the first or third consonant of the CVCs (e.g., first consonant vik, gik, mik...) also result in performance decrements that are equivalent to a repeated CVC. Changes in the entire CVC (e.g., vik fuv, zat...) or just in the vowel of the CVC (e.g., vik, vak, vuk...) result in significantly worse performance decrements compared to the repeated, initial-consonant-change, or final-consonant-change CVCs. It had earlier been found that CVC stimuli that only change in vowel position (e.g., vik, vak, vuk...) were more easily recalled in order than stimuli that change in the initial consonant (e.g., vik, gik, mik...) (Cole, 1973). Based on their experiments, Hughes, Tremblay, and Jones (2005) concluded that verbal-serial recall relies heavily on changing vowel information; therefore, recall of letter sequences in the presence of irrelevant vowel-change CVCs results in an increased competition for verbal resources and a decrease in task performance.

The results of these experiments are subject to theoretical criticisms based on the similarity of content hypothesis. In Experiment I and II, Hughes, Tremblay, and Jones (2005) found a significant interference difference between the silent-control and the repeated CVC as well as a significant interference difference between the repeated CVC and the CVC with all components changing (Experiment I) or the vowel-changing CVC (Experiment II). In both of these experiments, the most speech-like condition (all components changing and vowel changing) did interfere with the to-be-remembered letter sequences. The less speech-like condition (the repeating of a single CVC) resulted in significantly less interference. Therefore, these results can have a very different explanation that actually supports the similarity of content theory.

Finally, Bridges and Jones (1996) investigated the similarity of content hypothesis' validity in Experiments IV and V. They ran a replication of Salame and Baddeley (1982) in which participants were asked to recall nine-digit sequences in the presence of spoken digits, phonologically similar nonwords (e.g., tun, gnu, tee...), phonological dissimilar words (e.g., tennis, jelly, double...) and silence. They failed to replicate Salame and Baddeley (1982). The irrelevant speech conditions all resulted in similar amount of interference, and they were not even in the order predicted by the similarity of content explanation. Though the conditions were not statistically different from one another, even the trend was random. Listening to digits (the greatest phonological overlap) actually resulted in the least interference. Experiment V replicated these results with digits, phonologically dissimilar monosyllabic words (e.g., bad, sap pick...), and phonologically dissimilar two-syllable words (e.g., tennis, jelly, double...). The three types of interference were indistinguishable from one another. These results seem to be the first that are completely different from what would be predicted by the phonological loop model through the similarity of content hypothesis.

In summary, the similarity of process hypothesis states that the irrelevant sound effect occurs because order cues present in the to-be-remembered stimulus are confused with order cues present in the to-be-ignored stimulus. The strength of order cues present in the ignored stimulus influences the effect of irrelevant sound: stronger cues (more distinguishable objects) result in more interference. The theoretical problems with the similarity of process hypothesis have been laid out in this section. Again, the validity of this hypothesis will be further tested in the current experiments by investigating the effects

of different irrelevant-sound conditions on performance decrements in the memorization of musical sequences.

CHAPTER 2

WORKING MEMORY AND MUSIC

Based largely on Salame and Baddeley (1989), Berz (1995) proposed an alternative scheme to account for the working rehearsal and retention of musical material. Salame and Baddeley (1989) state, “the fact that we can hear and remember sounds that are very unlike speech means that there must be some – presumably additional – form of acoustic storage system capable of dealing with such material” (p.120). Berz (1995) argued that working memory for verbal material operates independently of working memory for musical material. He went so far as to propose an additional slave system to Baddeley’s working memory model – a music memory loop. The music memory loop would encode music in a separate store that he indicated could actually overlap in certain areas with the phonological loop.

Seemingly supporting this idea that interference for music and speech is modality specific, Deutsch (1970) found that interpolated, to-be-ignored spoken digit sequences did not interfere with recognition memory for a single tone nearly as much as to-be-ignored interpolated tonal sequence. Furthermore, she found that this was not an artifact of an increased attention to the interpolated tone sequences. Requiring participants to recall the interpolated digit sequences did not increase tone recognition errors, and the same amount

of digits were recalled regardless of whether participants were required to complete the tone comparison task.

Pechmann and Mohr (1992) extended this paradigm to include five interpolated conditions between the standard/comparison test tones: attended verbal material, unattended verbal material, unattended tonal material, attended visual material, and unattended visual material. The attended verbal condition required participants to indicate whether the last two items of the six-item interpolated sequence rhymed, whereas the attended visual condition required participants to say whether the last two visual matrix patterns of the six-item visual interference sequence were the same.

Pechmann and Mohr (1992) found that non-musicians performed the best in the silent-control condition, significantly worse in the attended verbal condition, and significantly worse still in the tonal condition. Musicians only showed interference of the tonal condition. Although they feel that their results do not point to one particular explanation, they suggested two possible hypotheses for their results. The first was a sister loop similar to a music memory loop (Berz, 1995), called the tonal loop, in which there are tonal-storage and tonal-rehearsal components separate to language storage and rehearsal. The second hypothesis stated that a single acoustic store existed which dealt with both speech and tonal information while having specialized rehearsal mechanisms for language and music.

An alternate and simpler explanation of the results of Deutsch (1970) and Pechmann and Mohr (1992) can be developed based on the Salame and Baddeley (1989) experiment which was discussed earlier. The differential overlap of to-be-ignored and to-

be-remembered stimuli causes differential interference. In the Salame and Baddeley (1989), Deutsch (1970), and Pechmann and Mohr (1992) studies, the more the to-be-remembered and to-be-ignored stimuli had in common, the more interference occurred. Also, the Deutsch (1970) and Pechmann and Mohr (1992) experiments dealt with the retention of a single tone. It is difficult to believe that this is enough to tax working memory. Instead participants may have been mentally “singing” the to-be-remembered interpolated digits on the to-be-remembered standard pitch in order to keep them both fresh in memory. An alternative explanation, as Jones and Macken (1993) have suggested, is that the initial tone could have been grouped more readily with interpolated tones than interpolated digits making the tone less likely to be differentiated in the standard/comparison task. Based on this perceptual grouping explanation, Jones and Macken (1993) concluded that a “separate storage for speech and nonspeech in memory is untenable” (p. 379).

THE ARTICULATORY SUPPRESSION EFFECT AND MUSIC

A recent study (Schendel & Palmer, in press) has also disputed the need for an additional music memory loop for musical working memory (Eliot & Cowan, 2005; Jones & Macken, 1993) instead proposing that both music and language are dealt with by the phonological loop. The Schendel and Palmer experiments utilized the articulatory suppression paradigm discussed earlier (Murray, 1967; 1968) and a musical paradigm used by Logie and Edworthy (1985). Logie and Edworthy (1985) required participants to make same/different judgments about standard/comparison auditorily-presented short

melodies with different secondary tasks. Participants were required either to use articulatory suppression (“the the the”), complete a homophone judgment task (e.g., Are these two words homophones: cloke and cloak), or match visual symbols as secondary tasks. There was a significant drop in recognition accuracy when articulatory suppression or homophone judgment tasks were the secondary tasks, with no drop when visual symbol matching was the secondary task.

Murray (1967; 1968) had already demonstrated the effect of verbal suppression on the working retention of visually and auditorily-presented letter sequences. Material presented in both modalities is impaired by suppression when compared to a no-suppression baseline, but visually-presented material is impaired more than auditorily-presented material. The purpose of Schendel and Palmer (in press) was to find out if this same pattern of results would occur when music is used as both the stimulus and the articulatory suppressor.

Schendel and Palmer (in press) had participants make same/different judgments about standard/comparison digit or note sequences. Experiment I used visual presentation of sequences, and Experiment II used auditory presentation of sequences. During the presentation of the standard and the retention interval before the comparison, participants either spoke “the the the” (verbal suppression; Murray, 1967), sang “la la la” on middle C (musical suppression), or remained silent. Schendel and Palmer (in press) found no difference between types of suppression: singing and speaking had similar effects on notes and numbers respectively, and this finding was not without precedent. Saito (1997; 1998) found that any intermittent use of the articulators (including whistling intermittently) acted

as an equivalent suppressant. Also, there was a consistent main effect of stimulus type: digits were recognized more accurately than notes in all conditions (Roberts, 1986). That would be expected since everyone has more experience memorizing number sequences than note sequences.

Digit results in Experiment I replicated those of Murray (1967; 1968). He found that visually-presented digits were more impaired by articulatory suppression than auditorily-presented digits. But, note results from Experiment I were in the opposite direction – auditorily-presented notes were more impaired by suppression than visually-presented notes. In Experiment II, it was found that visual interference impaired performance on standard/comparison visual note trials but did not impair performance on auditory note, visual digit, or auditory digit trials. Since visual interference impairs visually-coded information and not phonologically coded information (Cocchini, Logie, Della Salla, MacPherson, & Baddeley, 2002; Logie & Edworthy, 1985), it was determined that participants were not translating the visual notes in Experiment I into an auditory signal for rehearsal. If no translation to a phonological code occurs, then suppression will have no effect on working memory since suppression does not impair visual memorial representations.

In order to be sure both digits and notes were being rehearsed in an auditory form, a mixed modality visual-standard/auditory-comparison experiment was devised for Experiment III. Participants were forced to translate the visual signal into an auditory code to succeed at the task. Following this procedural manipulation, music results mirrored those of the digits results from Experiment I: both visual and auditory

presentation of digit and note sequences were impaired by suppression, but the visual modality was impaired more than the auditory modality. Schendel and Palmer (in press) concluded that when forced to translate notes into an auditory form, both language and music show the same breakdown pattern characteristic of the phonological loop.

THE IRRELEVANT SOUND EFFECT AND MUSIC

Sloboda (1976) and Martin, Wogalter, and Forlano (1988) extended the irrelevant sound effect to music; they investigated the retention of musical information in the presence of irrelevant sound. Sloboda (1976) used visual presentation of one to six note sequences while to-be-ignored auditory speech, tonal music, or atonal music was playing as interference. He found no effect of interference with any background sounds in the recall of the visual note sequences with musicians. Sloboda's (1976) results can be explained in terms of the working memory model, and they do not require the addition of a music memory loop. To-be-ignored auditory background material does not interfere with visually based material (e.g., Cocchini, et. al., 2002). Visual interference degrades working rehearsal and retention of visually-based information. It has subsequently been found that visual interference significantly impairs the retention of visually-presented notes but not auditorily-presented notes, visual numbers, or auditory numbers (Logie & Edworthy, 1985; Schendel & Palmer, in press). Therefore, Sloboda's (1976) results were not that surprising.

Martin, Wogalter, and Forlano (1988) required participants to identify popular songs from their written musical notation in the presence of different irrelevant-sound conditions. Essentially, participants would be required to translate the visual musical symbols into an auditory form in order to determine what the notation “sounded like” to succeed at the task. With language, this is most equivalent to reading. The three to-be-ignored background interference conditions used were silence, listening to a recording of spoken prose, and a musical jazz piece. They found an interference continuum: silent performance was well above that of performance with the irrelevant speech background, and performance with the irrelevant speech background was well above performance with the irrelevant music background.

In summary, research on working memory and music is just beginning. Early research has debated whether the phonological loop model can account for musical working memory research results (Berz, 1995; Pechmann & Mohr, 1992). Based on experiments involving articulatory suppression, it has been hypothesized that working memory for language and music are both governed by similar if not the same processes (Schendel & Palmer, in press). Indeed Bridges and Jones (1996) state: “at the very least, the assumption that speech is the only material to gain automatic access to the store and that the irrelevant speech effect is the result of phonological confusions between heard and seen items would have to be dropped” (p. 938). The current research should help clear up the similarities and differences between working memory for language and working memory for music.

PURPOSE

The general purpose of the current research was to further investigate the effects of irrelevant sound on serial memory in order to better understand the similarities and differences between working memory in language and music. This was accomplished by using the irrelevant sound effect paradigm, but the current experiments went beyond asking participants to retain serial language sequences. On some trials, participants had to retain musical sequences. Requiring participants to hold musical information in working memory in the presence of irrelevant sound allowed for a new test of the working memory model. Is there a need for an additional music memory loop to account for musical working memory?

The specific purpose of the current research was to investigate the differences between the similarity of content hypothesis and the similarity of process hypothesis. Is it the dissimilarity between irrelevant sound stream objects that will result in higher error rates (Jones & Macken, 1993)? Or, will it be shown that greater overlap between the to-be-remembered and to-be-ignored stimuli will result in greater performance decrements (Salame & Baddeley, 1989)? It was hypothesized that when participants were required to remember digit sequences, results would replicate those of Salame and Baddeley (1982; 1989). There would be a performance continuum with the most speech-like irrelevant sound producing the most interference. It was hypothesized that when participants were required to remember musical sequences, an interference continuum would also occur in which the most music-like irrelevant sound would produce the most interference.

The key statistical comparisons in the experiments were made using planned comparisons (a series of one-tailed t -tests) instead of post hoc analyses. It was known going into the experiments which cell means were to be compared as well as the predicted direction of performance to differentiate between the similarity of process and similarity of content hypotheses. In each experiment, it was predicted that the order of performance among the irrelevant-sound conditions would be a result of the acoustic overlap between the irrelevant sound and to-be-remembered stimuli. Greater overlap would yield lower performances. For example, imagine one were memorizing a digit sequence in the presence of silence, irrelevant tones, and irrelevant speech. In this example, performance in the silent-control condition would be predicted as the most accurate condition, followed by a drop in performance in the irrelevant-tone condition, and a further drop in performance in the irrelevant-speech condition. The silent-control would then be compared to the irrelevant-tone and irrelevant-speech conditions, and the irrelevant-tone and irrelevant-speech conditions would also be compared using one-tailed t -tests.

CHAPTER 3

EXPERIMENT I

Salame and Baddeley (1989) ran a within-subject experiment in which each participant serially recalled visually-presented nine-digit sequences in the presence of to-be-ignored irrelevant sound. The irrelevant sound used was instrumental music, Arabic speech, and a silent-control. They found an interference continuum, where the more speech-like the irrelevant sound was, the more it interfered with immediate serial recall. In the current set of experiments, Experiment I was designed to replicate Salame and Baddeley (1989). Along with the visually-presented nine-digit sequences, participants would also memorize visually-presented nine-note sequences on some trials. The purpose of this experiment was to assess whether the immediate serial recall of visually-presented digits and music were differentially impaired by to-be-ignored music and non-music stimuli. A potential problem exists because visual notes can be memorized in a variety of ways: as note names, visual information such as lines and spaces, or tonal information. The way in which visual notes are memorized will impact the effect of irrelevant sound. Notes memorized as note names will demonstrate speech-like interference. Notes memorized as tones will demonstrate an interference pattern in the opposite direction as speech. And notes stores as visual information will not be impaired by irrelevant sound.

Therefore, it was hypothesized that digits would show a speech-like interference continuum and notes would demonstrate a variety of interference patterns.

PARTICIPANTS

Twenty-four adult students from Introductory Psychology at The Ohio State University with at least 6 years of private instruction on a musical instrument participated in the study (mean = 8.21 years lessons; range = 6-14 years lessons). In order to participate, they were required to complete a short Music Pretest (Appendix A) to make sure they could read music in the treble clef. Participants received partial class credit for their participation.

MATERIALS AND APPARATUS

Forty-five, nine-item digit sequences were randomly generated from the digits 1-9. Forty-five, nine-item melodic sequences were randomly generated from nine pitches in the C Major diatonic scale (D4 through E5) to avoid notated accidentals (sharps and flats). The music notes were presented visually one at a time as quarter notes on a treble clef musical staff. The digits were presented visually one at a time in the center of a box; the box was created to be the same size as the musical staff bar in the music trials. Both sequences were presented in the same space. Each sequence did not contain any sequentially repeating items or any intervals of an octave or larger (digit equivalent of 1-8,

1-9, or 2-9). All digits and notes were visually-presented using Presentation software on a PC with 750-ms inter-onset intervals (IOIs).

The irrelevant sound was presented auditorily in stereo at an average level of 65 dB over AKG headphones. The choices made for irrelevant sound were based on Salame and Baddeley (1989) – silent-control, vocal music, instrumental music, and Arabic speech. Vocal music: arias from Mozart’s “Magic Flute,” Schubert’s “Die Schöne Müllerin,” Mozart’s “The Marriage of Figaro,” and Rossini’s “Barber of Seville.” Two of these arias are sung by a female, and two are sung by a male. Instrumental music: Ravel’s “Bolero,” Berlioz’s “Rakoczi March,” Alford’s “Bridge over the River Kwai,” and Offenbach’s “French Cancan.” The Arabic speech came from a "How to Speak Arabic" CD.

DESIGN

A 2 (stimulus type) x 4 (irrelevant sound type) within-participant design was used. The two levels of the stimulus type variable were digits and music: nine-item digit sequences or nine-item note sequences. The four irrelevant sound types were described above in the materials and apparatus section. The experiment was split into eight blocks. Irrelevant sound type was blocked, and stimulus type was blocked within each irrelevant sound block. There were five practice trials before the first trials of each stimulus type, and each block contained ten trials. The order of stimulus type and irrelevant-sound-type blocks was counterbalanced across participants.

PROCEDURE

After obtaining informed consent, participants were asked to take the music pretest and were then seated in front of the display monitor. On all trials, participants completed an immediate serial-recall task. The beginning of each trial was signaled by a visually orienting “+” on the computer screen. A visual digit or note nine-item sequence was then presented at the rate of one item/750-ms. Once the recall cue “+” was given, the participants recorded the sequence on the provided answer sheets. After they finished a trial, they pressed the space bar to begin a new trial. Participants were given short breaks between blocks during which time they were read the instructions for the next block. During six of the eight blocks, the participants wore headphones and listened to irrelevant sounds. Participants were instructed to ignore the background sounds that occurred on any block. The irrelevant sound began 3-secs before the first trial of each block so as to minimize any effect of surprise. The irrelevant sound continued throughout the block without interruption until all 10 trials had been completed. After the fourth block, participants filled out the Musical Experience Questionnaire (Appendix B).

RESULTS

A 2 (stimulus type) x 4 (irrelevant sound type) x 9 (serial position) ANOVA on response accuracy (percent correct) was conducted. There were main effects of stimulus

type: $F(1,23) = 84.29, p < .01$, irrelevant sound type: $F(3,69) = 11.68, p < .01$, and serial position: $F(8,184) = 46.09, p < .01$. As can be seen in Figure 3.1, there was a stimulus-type by irrelevant-sound-type interaction: $F(3,69) = 3.7, p < .05$. There was a speech-like interference continuum for both stimuli, but it was more pronounced for digits: instrumental music resulted in the least interference, followed by greater interference in the vocal-music condition, and the Arabic-speech condition showed the most interference. The decline in performance accuracy as the interference became more speech-like was much steeper with digit stimuli (silent-control condition minus Arabic-speech condition = 15.13%) than with note stimuli (silent-control condition minus Arabic-speech condition = 2.97%).

A series of planned comparisons were run between the performances in the irrelevant-sound-type conditions of each stimulus-type. As for the digit trials, performance in the silent-control trials was significantly better than performance in any of the irrelevant-sound conditions [Control (C) – Instrumental music (I): $t(23) = 2.09, p < .05$; C – Vocal music (V): $t(23) = .5, p < .05$; C – Arabic speech (S): $t(23) = .77, p < .01$]. Performance in the instrumental-music and vocal-music conditions did not differ. But, performance in both the instrumental and vocal-music conditions was better than performance in the Arabic-speech condition [I – S: $t(23) = 2.57, p < .01$; V – S: $t(23) = 1.36, p < .05$]. As for the note trials, planned comparisons indicated that performance in the silent-control condition was significantly better than performance in any of the irrelevant-sound conditions [C – I: $t(23) = 2.35, p < .05$; C – V: $t(23) = .54, p < .05$; C –

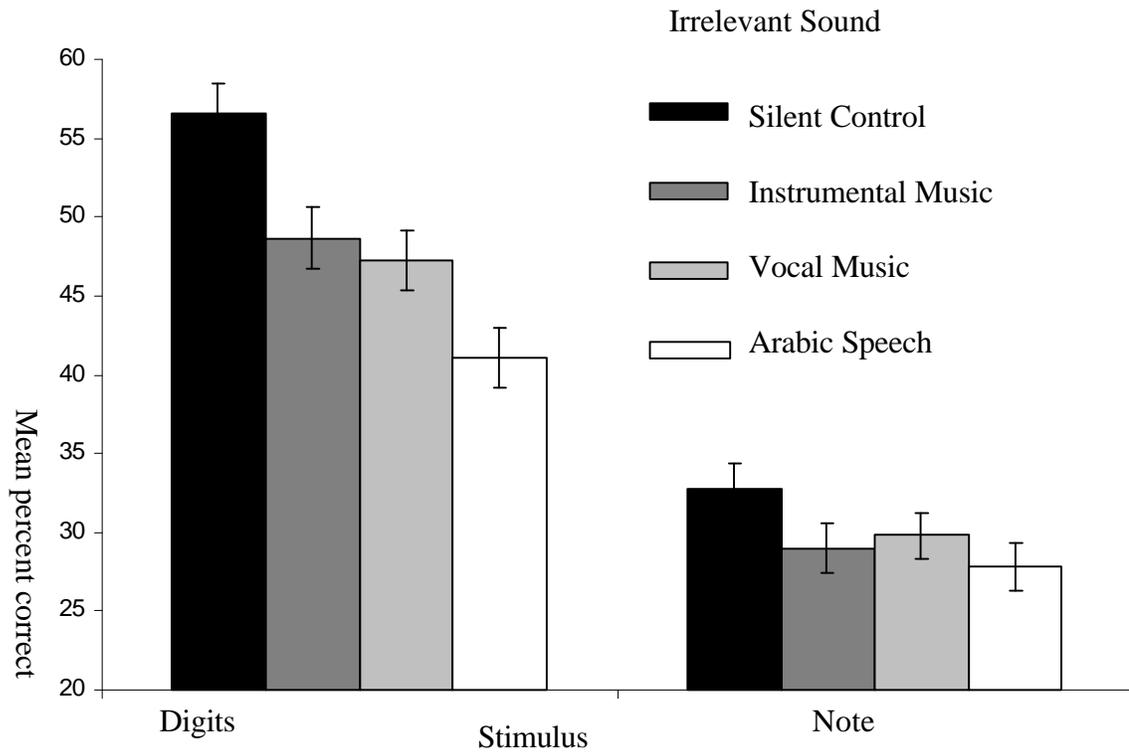


Figure 3.1: Experiment I: Mean percent correct on a nine-item serial recall task with visual digits and notes in the presence of silence, instrumental music, vocal music, and Arabic speech irrelevant sound.

S: $t(23) = 2.96, p < .01$], but there were no differences between any of the irrelevant-sound conditions.

To further understand the null effect of interference type on note trials, additional analyses were run. When examining the individuals' results from the note trials of Experiment I, it was apparent why there were null differences among the interference conditions. Five of the 24 participants performed at chance on at least two of the conditions, so their results were removed from further analysis. Of the remaining 19 participants, 12 had the largest drop in performance in the presence of Arabic-speech (the Arabic speech group), six had the largest drop in performance in the presence of instrumental music (the instrumental music group), and one had the largest drop in performance in the presence of vocal music. The error patterns for the Arabic-speech group and the instrumental-music group were as follows: Arabic-speech group: 37% control, 36% instrumental music, 33% vocal music, 30% Arabic speech; Instrumental-music group: 30% control, 30% Arabic speech, 28% vocal music, 22% instrumental music. There were not enough participants to run planned comparisons on the different groups, but there was clearly a bi-modal interference pattern. The Arabic-speech group also performed better over all conditions (34%) than the instrumental-music group (28%). In fact seven of the top eight performers in the silent-control condition were in the Arabic speech group. Even considering this difference, based on the different group error patterns, visually-presented music did not act the same as visually-presented speech. The participants in the Arabic-speech group performed like the participants in the digit half of

the experiments, but the participants in the instrumental-music group had the opposite interference pattern.

There seems to be a divide between people who performed in the same pattern as the digit results and those who performed in the opposite pattern as the digit results. It is possible that the two participant groups chose to memorize the note sequences in different manners. There were no musical training differences between the two groups (Arabic-speech: 8.54 years, instrumental-music, 8.58 years). Perhaps those who performed the worst in the Arabic-speech condition memorized the sequences primarily as note letter names while those who performed the worst in the instrumental-music condition memorized the stimuli as tone sequences? An additional exit interview question probing the participant to describe how they went about memorizing and rehearsing the to-be-remembered visual notes would be an interesting follow-up experiment and could shed light on the differences described here.

In the original 2 (stimulus type) x 4 (irrelevant sound type) x 9 (serial position) ANOVA on percent correct, there was also a stimulus-type by serial-position interaction: $F(8,184) = 2.77, p < .01$. Both digits and notes showed similar primacy effects, but note recency effects were more pronounced; accuracy in the note trials began to increase on the seventh serial position, and accuracy in the digit trials only increased on the ninth serial position. As discussed in the introduction, a recency effect is indicative of short-term memory. Digit positions seven, eight, and nine were run in a separate 4 (irrelevant sound type) x 3 (serial position) ANOVA on percent correct. Because note results were bimodal, only the digit trials were run in this analysis. As can be seen in Figure 3.2, There

was a main effect of irrelevant sound type: $F(3,69) = 3.78, p < .05$. The order of performance from most accurate to least accurate was as follows: silent-control, instrumental music, vocal music, and Arabic speech. Planned comparisons indicated significant differences between the silent-control condition and all of the irrelevant-sound conditions [C – I: $t(71) = 2.59, p < .01$; C – V: $t(71) = 4.65, p < .01$; C – S: $t(71) = 4.36, p < .01$]. There was also a significant difference between the instrumental-music condition and the Arabic-speech condition [I – S: $t(71) = 1.98, p < .05$]. The difference between instrumental and vocal-music conditions approached significance [I – V: $t(71) = 1.27, p = .1$], and the vocal music and Arabic-speech conditions were not significantly different from one another.

Looking more closely at each serial position, it was seen that on position seven, the only planned comparisons that were significantly different were the silent-control and the vocal-music condition [N – V: $t(71) = 2.88, p < .01$] as well as the silent-control and the Arabic-speech condition [N – S: $t(71) = 1.99, p < .05$]. In position eight, all of the irrelevant-sound conditions were different from the silent-control condition [C – I: $t(71) = 1.89, p < .05$; C – V: $t(71) = 2.23, p < .05$; C – S: $t(71) = 2.41, p < .05$]. In position nine, the silent-control was significantly different from the vocal-music condition and the Arabic-speech condition [C – V: $t(71) = 2.84, p < .01$; C – S: $t(71) = 3.08, p < .01$]. The instrumental-music condition was significantly different from the Arabic-speech condition [I – S: $t(71) = 2.38, p < .05$]. Position nine, the only position in which a recency effect was evident in the digit trials, was the only position that had significant differences among the irrelevant sound types indicative of the similarity of content

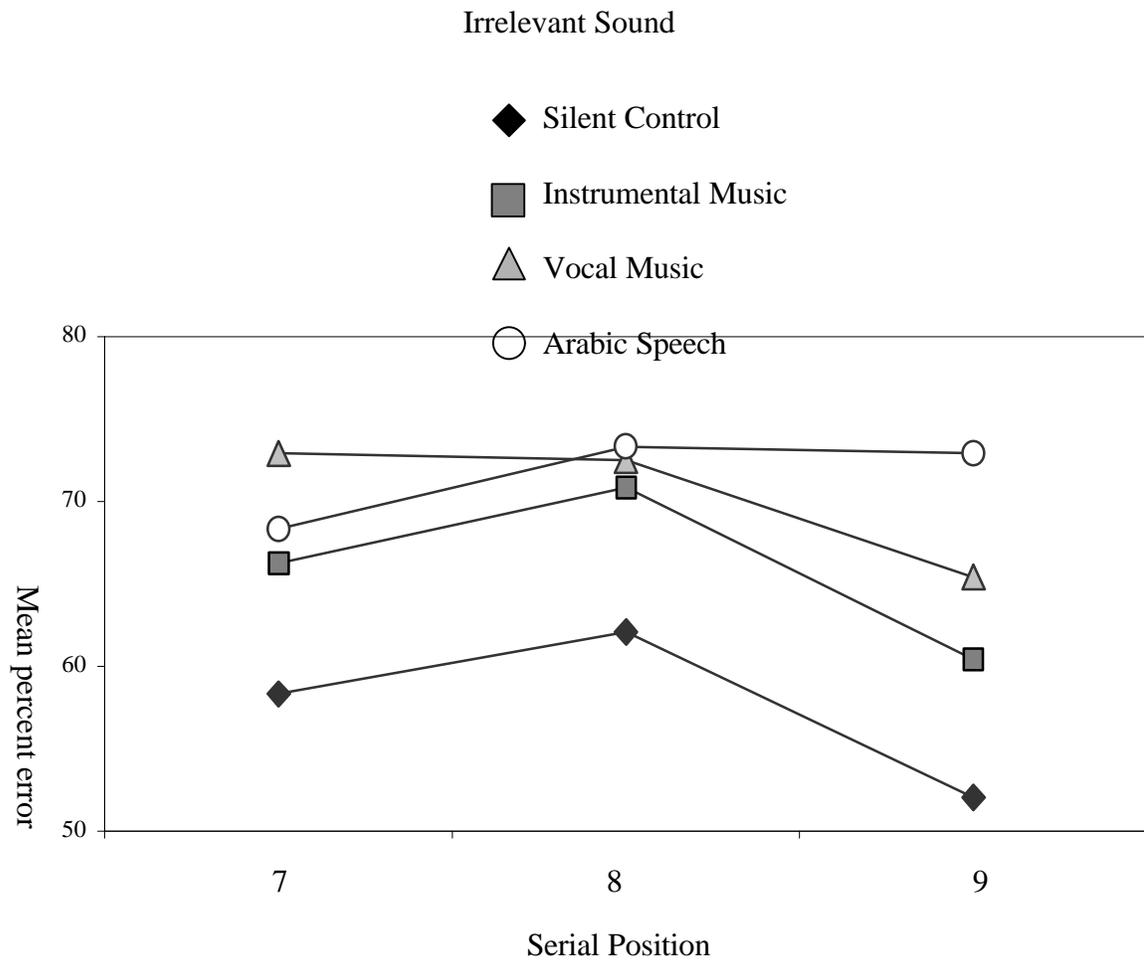


Figure 3.2: Experiment I: Mean percent errors in positions seven, eight, and nine on a nine-digit serial recall task in the presence of silence, instrumental music, vocal music, and Arabic speech irrelevant sound.

hypothesis. This result replicates Jones and Macken (1995) – the recency portion of the serial recall curve demonstrates the predicted interference pattern indicative of the similarity of content hypothesis.

DISCUSSION

The digit trials replicated the results of Salame and Baddeley (1989). The error patterns supported the hypothesis that a greater phonological overlap between the to-be-remembered and to-be-ignored stimuli results in greater interference. The serial position digit results also replicated the pattern found in Jones and Macken's (1995) Experiments I and II. In those experiments, the more speech-like interference resulted in worse performance on the positions indicating a recency effect.

The note trials showed a very different pattern. Across all participants, there was a null effect of interference-type. All three changing-state interference conditions resulted in similar performance decrements. But, looking more closely at the note results, it was evident that there were distinct groups of participants who performed in the opposite direction from each other. This result can be explained in the same manner as Schendel and Palmer's (in press) experiments in which they found that visually-presented notes are a special case. Participants can choose to remember visually-presented notes in a variety of forms – note letter names, tones, or visual lines and spaces – some of which are not impaired by irrelevant sound. It was hypothesized that different rehearsal strategies would be impaired by irrelevant sound in different manners. Rehearsing visual notes as note names is akin to speech, and participants who chose this rehearsal strategy would have

interference patterns similar to the digit trials. There were participants with this error pattern. Rehearsing visual notes as tones is akin to singing the sequence, and participants who chose this rehearsal strategy would have interference patterns in the opposite direction as the digit trials. There were also participants with this error pattern. There was no exit interview which could have detailed the rehearsal strategy chosen by each participant for Experiment I. Future experiments will address this problem of multiple rehearsal strategies as well as attempt to learn more about the participants' thought processes in exit interviews.

One interesting result from Experiment I appeared in the different performance decrements of the digit trials. While the pattern of results was in the predicted direction, all of the conditions did not result in statistically different performance decrements. For example, performance in both instrumental music and vocal music conditions were not significantly different from each other even though performance in both conditions was in the predicted direction. This seemed predictable since both instrumental music and vocal music have similar orchestral content. But, vocal music uses the voice as an instrument on top of the typical orchestral instruments that hypothetically should impair performance slightly more on the digit trials than instrumental music.

Perhaps vocal music and instrumental music are too similar to result in differential performance decrements. This brings up an interesting hypothesis about Jones and Macken's (1993) results that was discussed in the introduction. Even though the results they found were all in the predicted direction based on the similarity of content hypothesis, the similarity of process hypothesis was based partially on non-significant statistical

differences between the different irrelevant-sound conditions. Perhaps the irrelevant sounds they used were too similar to result in a statistical effect that would be predicted by the similarity of content hypothesis. One example would be the changing four-tone sequence compared to the spoken four-letter sequence (Experiment II). Perhaps if spoken digits were used instead of spoken letters (thus the to-be-remembered and to-be-ignored stimuli were the same) there would have been a difference between the changing tones and changing words. While this conclusion seems promising, there are also problems with the similarity of content theory; it does not make any specific predictions about the nature of the overlap between the to-be-remembered and to-be-ignored stimuli and how this overlap may directly affect performance. The drawbacks of both hypotheses will be discussed further in the general discussion.

CHAPTER 4

EXPERIMENT IIA

Earlier results have demonstrated that, when memorizing visually-presented note sequences, musicians do not rely solely on translating visually-presented notes into an auditory form for phonological loop rehearsal (Logie & Edworthy, 1985; Schendel & Palmer, in press). Instead, musicians rely on a combination of auditory and visual cues in memorizing visually-presented note sequences, and the individual's specific combination may be influenced by musical training (Sloboda, 1976) and the task context (Schendel & Palmer, in press). Visual cues are not impaired by irrelevant sound. Experiments IIA and IIB were designed to investigate the effect of irrelevant sound on auditory representations of language and music. For the digit trials, the standard was presented visually, and the comparison was presented auditorily. Schendel and Palmer (in press) and pilot testing found this visual-standard/auditory-comparison paradigm to be particularly difficult for note trials, often resulting in chance performance in all conditions unless the participant had more than 10 years of musical experience. A large population of musicians of this caliber was not readily available, so the music trials were slightly different from the digit trials. In the music trials, the standard and the comparison were both presented auditorily. The tone trials were theoretically the same as the digit trials since both stimuli must be

rehearsed in an auditory form in order for the participant to succeed at the task. It was hypothesized that the more similar the irrelevant sound was to the to-be-remembered stimuli, the more interference would occur. While Experiment I was a good place to begin this line of research, Experiment IIA provided a more exacting test of the original hypothesis because participants had to rehearse the stimuli in an auditory form in order to succeed at the tasks. Rehearsing the stimuli in an auditory form makes the memorial code susceptible to irrelevant speech. The hypothesis was that greater content similarities between the to-be-remembered and to-be-ignored stimuli would result in worse performance on the standard/comparison task. For the digit trials, the more speech-like the to-be-ignored stimulus was, the worse performance would be. For the tone trials, the more music-like the to-be-ignored stimulus was, the worse performance would be.

PARTICIPANTS

Forty-eight students were run in the experiment. Twenty-four had the same musical experience qualifications as in Experiment I, and they were tested in the music half of the experiment (mean = 7.5 years lessons, range = 6-12 years lessons). No participant from Experiment I participated in Experiment IIA. Students received partial class credit for their participation.

MATERIALS AND APPARATUS

Sixty-nine, auditory four-item note sequences were randomly generated from nine pitches in the C Major diatonic scale (D4 through E5). The auditory tones were sampled at 44 kHz with a piano timbre using Cakewalk Professional with a duration (from onset to offset) of 450-ms. The tone sequences were created by splicing together single recordings of each tone with 700-ms IOIs. This was done to avoid possible timbral variations which could occur if a different recording of each sequence was used on each trial. With many different recordings, it would have been possible for some of the recorded sequences to have additional acoustic cues which could aid the participant in sequence memorization. The comparison sequences were the same as the standard 50% of the time, and they differed by one unit (either a whole step or a half step) on 50% of the trials. Half of the pitch changes occurred in position two, and half occurred in position three. Half of position two and three changes went upward and half changed downwards.

Sixty-nine, visual six-item standard digit sequences were randomly generated from the digits 1-9. The digit standard and comparison sequences had two more items than the tone sequences because pilot test performance on four-item digit trials was at ceiling in all interference conditions. Each sequence did not contain any sequentially repeated digits. The auditory digit comparison sequences were created from a recording of a female voice speaking the digits one through nine. The duration (from onset to offset) of the digits was 400-ms (the slightly shorter duration of the spoken digits than the musical tones was

judged more natural). The comparison digit sequences were created by splicing together the recordings of the single digits with 700-ms IOIs. The comparison sequences were the same as the standard 50% of the time, and they differed by one digit on 50% of the trials. A quarter of the digit changes occurred in position two, 25% in position three, 25% in position four, and 25% in position five. Half of position two, three, four, and five changes went upward and half changed downwards.

All stimuli were presented on a PC running Presentation software. Participants heard the auditory stimuli through computer speakers. Participants heard the irrelevant sounds over AKG headphones at an average level of 65 dB. The same irrelevant sound recordings were used in Experiment IIA as were used in Experiment I: silence, instrumental music, vocal music, and Arabic speech.

DESIGN

The design of Experiment IIA was the same as Experiment I – 2 (stimulus type) x 4 (irrelevant sound type) – but stimulus type was a between subject factor. There were four blocks of 16 trials and five practice trials.

PROCEDURE

Participants were seated in front of the display monitor wearing the AKG headphones. For the music participants, each trial was a standard/comparison trial in

which a four-tone auditory standard sequence (700-ms IOIs) was played over the computer speakers and followed by a 4900-ms retention interval. The purpose of the retention interval was to induce rehearsal such that confusions between the to-be-remembered and to-be-ignored stimuli could occur. The four-tone auditory comparison sequence played immediately following the retention interval over the computer speakers. They were then asked to indicate whether the two sequences were the same or different and how confident they were in their answer (three-point confidence scale). Three of the four trial blocks were completed in the presence of irrelevant sound interference. On each of these trials, the irrelevant sound was played over the AKG headphones during the retention interval. It began after the last item of the standard sequence ended and ended right before the first item of the comparison sequence began.

The digit trials were the same as the tone trials aside from two changes. First, the standard sequence was presented visually. And, second, the irrelevant sound played during the presentation of the standard as well as during the retention interval. It ended immediately before the first item of the comparison sequence began.

RESULTS

There were differences in the number of items participants were required to memorize in each stimulus condition; participants were required to memorize six-digit sequences and four-tone sequences. Also, stimulus type was a between participant factor. Because of these differences, it was unknown whether the memory loads in each stimulus

condition were equivalent. Following the initial ANOVA including both stimulus types, each further analysis only includes a separate single-factor (irrelevant sound type) ANOVA for each stimulus type.

A 2 (stimulus type) x 4 (irrelevant sound type) ANOVA was run on the percent correct scores. There was a main effect of stimulus type: $F(1,30) = 13.91, p < .01$, a main effect of irrelevant sound type: $F(3,90) = 15.55, p < .01$, and, as can be seen in Figure 4.1, a stimulus type by irrelevant sound type interaction: $F(3,90) = 6.47, p < .01$. The error pattern for the digit trials was such that the more speech-like the interference was, the more it interfered with the task. Performance was most accurate in the silent-control condition, followed by instrumental music, then the Arabic speech and vocal music. As for the tone trials, the silent-control condition was still the most accurate, but the interference pattern was in the opposite direction: the instrumental-music condition resulted in the most interference, followed by the vocal-music condition, and then the Arabic-speech condition. Planned comparisons are described below following the two individual stimulus-type ANOVAs on percent correct.

Separate single-factor (irrelevant sound type) ANOVAs were run on the percent correct scores within each stimulus type. For the digit trials, there was a main effect of irrelevant sound type: $F(3,45) = 3.07, p < .05$. Planned comparisons on the digit trials indicated that the difference between performance in the silent-control condition and the instrumental-music condition approached significance [N – I: $t(15) = 1.64, p = .06$]. Performance in the silent-control condition was significantly better than performance in

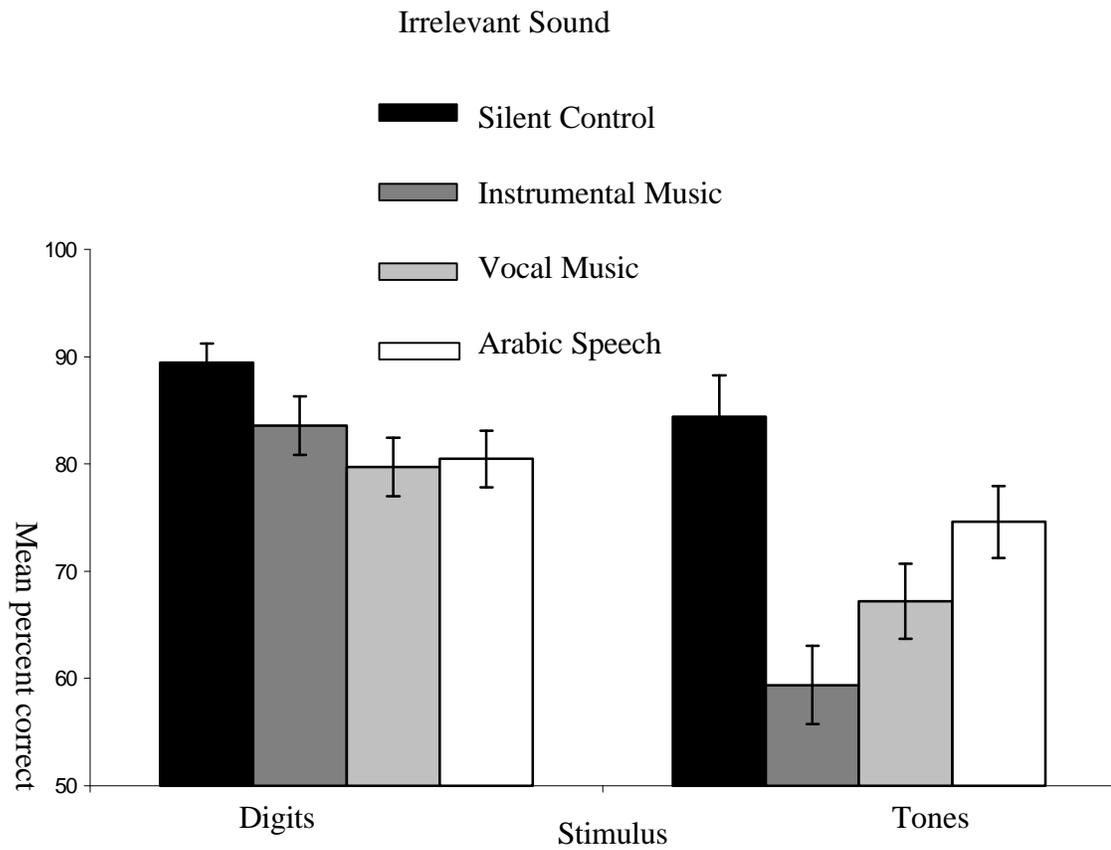


Figure 4.1: Experiment IIA: Mean percent correct on a visual-auditory digit and auditory-auditory tone recognition task in the presence of silence, instrumental music, vocal music, and Arabic speech irrelevant sound.

the vocal music and Arabic-speech conditions [$N - V: t(15) = 2.85, p < .01$; $N - S: t(15) = 2.59, p < .05$]. The difference in performance between the instrumental music and vocal-music conditions also approached significance [$I - V: t(15) = 1.35, p = .09$]. All other comparisons between irrelevant-sound conditions were not significant.

The single-factor (irrelevant sound type) tone ANOVA on percent correct also resulted in a main effect of irrelevant sound type: $F(3,45) = 19.94, p < .01$. Planned comparisons indicated that there was a significant difference between the silent-control condition and each of the interference conditions [$N - I: t(15) = 6.68, p < .01$; $N - V: t(15) = 5.21, p < .01$; $N - S: t(15) = 2.78, p < .01$]. There was also a significant difference between the instrumental-music and vocal-music conditions [$I - V: t(15) = 2.05, p < .05$]. Finally, there was a significant difference between the Arabic-speech condition and both the instrumental-music and vocal-music conditions [$I - S: t(15) = 5.7, p < .01$; $V - S: t(15) = 2.41, p < .01$].

Additionally, a single-factor (irrelevant sound type) ANOVA was run on the d' -prime scores for each stimulus type. A hit was defined as correctly identifying a “different” trial when the standard and comparison were actually different. A correct rejection was defined as correctly identifying a “same” trial when the standard and comparison were the same. There was not a main effect of irrelevant sound type for the digit ANOVA, but the order of sensitivity was in the predicted direction (control, instrumental music, vocal music, and Arabic speech). Planned comparisons indicated there were no significant differences in sensitivity between any of the irrelevant-sound conditions. The tone ANOVA demonstrated the opposite sensitivity pattern (silent-

control, Arabic speech, vocal music, instrumental music), but again there was not a significant main effect of irrelevant sound type. Planned comparisons indicated that there were no significant differences in sensitivity between any of the irrelevant-sound conditions.

Two single-factor (4-irrelevant sound types) ANOVAs were run on the three-item confidence ratings for each stimulus type. Confidence ratings for every trial were used in this analysis. The digit ANOVA on confidence ratings resulted in a main effect of irrelevant sound type: $F(3,45) = 5.02, p < .01$. The highest to lowest confidence ratings were in the following order: control, vocal music, instrumental music, Arabic speech. Planned comparisons indicated significantly higher confidence ratings for the silent-control compared to each of the irrelevant-sound conditions [N – I: $t(15) = 2.33, p < .05$; N – V: $t(15) = 2.74, p < .01$; N – S: $t(15) = 3.36, p < .01$]. The comparison between instrumental music and Arabic speech approached significance [I – S: $t(15) = 1.43, p < .08$] while the vocal-music condition had significantly better confidence ratings than Arabic speech [V – S: $t(15) = 1.98, p < .05$]. The tone ANOVA also resulted in a main effect of irrelevant sound type: $F(3,45) = 10.23, p < .01$. The results from highest to lowest confidence were in the following order: control, Arabic speech, vocal music, instrumental music. Planned comparisons indicated a significantly higher confidence rating for the silent-control compared to each of the irrelevant-sound conditions [N – S: $t(15) = 5.28, p < .01$; N – V: $t(15) = 4.17, p < .01$; N – I: $t(15) = 6.21, p < .01$], and a significantly higher confidence rating for the Arabic-speech condition compared to the instrumental-music condition [S – I: $t(15) = 1.96, p < .05$].

The confidence rating differences replicated results found in exit interviews. When participants were asked to indicate the easiest and hardest conditions, 16/16 of the participants in the tone-stimulus trials indicated the silent-control was the easiest. Ten of the 16 participants indicated that the instrumental-music condition was the hardest. Four participants indicated the vocal-music condition was the hardest and two indicated the Arabic-speech condition was the hardest. Of the six that did not indicate instrumental music was the hardest condition, four actually performed the worst in the instrumental-music condition, and two performed the worst in the vocal-music condition.

Finally, a series of correlations were run comparing the results of the music experience questionnaire with participant data from the tone half of the experiment. There were significant positive correlations between the total years of musical training and performance in the silent-control condition ($r = .54, p < .05$), performance in the instrumental-music condition ($r = .64, p < .01$), and the average performance across all three auditory irrelevant-sound conditions ($r = .58, p < .05$). There was also a positive correlation between the total years participant has played instruments and performance in the vocal-music condition ($r = .51, p < .05$). Other conditions approached significance, but the general trend of the data indicated that more musical training or performing led to better performance in the various tasks.

DISCUSSION

The key difference between Experiment I and Experiment IIA was that both digit and tone trials were set up such that they had to be rehearsed in an auditory form in order for the participant to succeed at the trial. And, as discussed in the introduction, the locus of the irrelevant sound effect is on rehearsal not encoding since including the irrelevant sound during sequence presentation or only during a retention interval result in similar performance decrements. In Experiment I, visually-presented note trials allowed participants to choose their own representation strategy. As a result of this change, the different irrelevant sounds had a differential impact on performance across stimulus types. The digit results of Experiment IIA replicated those of Salame and Baddeley (1989) as well as the pattern found here in Experiment I. Arabic speech and vocal music resulted in the greatest performance decrements and were not different from each other. This was followed by instrumental music, and the silent-control condition was the most accurate. This pattern would be predicted by the similarity of content hypothesis.

Tone results showed a very different interference pattern. Instrumental music showed the greatest performance decrement, followed by vocal music, and then Arabic speech. The silent-control condition was the most accurate. Again, the data are best predicted by the similarity of content hypothesis. This result also provides additional support that the phonological loop may be responsible for the working rehearsal of both

language and music since both language and music showed similar, predictable breakdown patterns formerly only attributed to working memory for language.

One interesting result from Experiment IIA was found in the vocal-music conditions. In the digit trials, vocal music impaired performance to the same degree as the Arabic speech. This is equivalent to saying that vocal music interfered as much as the condition with the predicted greatest content similarity. This pattern is slightly different from the results of the digit trials in Experiment I in which the instrumental music and vocal-music conditions resulted in similar performance decrements, but performance in both of these conditions was better than in the Arabic-speech condition. A different pattern was seen with tone results. Instrumental music resulted in the greatest performance decrement, and the vocal music was significantly better. But, the vocal music was still significantly worse than the Arabic-speech condition.

Just how does vocal music interfere with speech and music? Salame and Baddeley (1989) never ran an experiment in which all three irrelevant-sound conditions were used as interference within the same experiment. But, they found in one experiment that vocal music impaired serial recall of a nine-digit sequence more than instrumental music, and they also found in a different experiment that Arabic speech impaired performance on the same task more than instrumental music. In this series of experiments, vocal music was designed to be a middle-ground interference composed of both music and speech. The arias combine both music and speech at some points, and at other points there is only singing or only instrumental music. The hypothesis was that during the vocal sections, digits would be impaired more, and during the musical sections, tones would be impaired

more. This led to the prediction that over a large group of participants and a large number of trials, both digits and notes would be impaired the most by Arabic speech and instrumental music respectively with vocal music resulting in an intermediate level of interference. The results of Experiment IIA were not entirely consistent with the hypothesis that vocal music would result in an intermediate level of interference in digit trials. However, this hypothesis may still hold in tone trials.

CHAPTER 5

EXPERIMENT IIB

Experiment IIB was the same as Experiment IIA, but only digit trials were run, and irrelevant sound type which had been a within participant factor in Experiment IIA was now a between participant factor. Experiment IIB was designed as a further test of the same hypotheses tested in Experiment IIA. Also of interest was the error pattern found in the digit trials of Experiment IIA. The interference conditions did not result in as sharp of statistically significant differences as the tone conditions did in Experiment IIA, so Experiment IIB was also developed as a continuation of IIA. Finally, Experiment IIB was designed to further test whether vocal music resulted in an intermediate level of interference (as found in Experiment I) or as equivalent interference to Arabic speech (as found in Experiment IIA).

PARTICIPANTS

Forty-eight students were tested in the experiment. No participant from Experiment I or IIA participated in Experiment IIB. Students received partial class credit for their participation.

MATERIALS AND APPARATUS

The materials and apparatus for Experiment IIB were the same as in Experiment IIA.

DESIGN

Experiment IIB only used the six-item digit trials, and irrelevant sound type was a between-participant factor. Participants had five practice trials and 16 experimental trials.

PROCEDURE

Experiment IIB's procedure was the same as the digit half of Experiment IIA, except that irrelevant sound type was a between-participant factor.

RESULTS

A single-factor (irrelevant sound type) ANOVA was run on the percent correct scores. As can be seen in Figure 5.1, there was a main effect of irrelevant sound type: $F(3,44) = 4.93, p < .01$. There were differences in the impact of the type of irrelevant sound on performance accuracy. The order of performance was in the predicted direction (control and instrumental music were the same, followed by vocal music, and Arabic

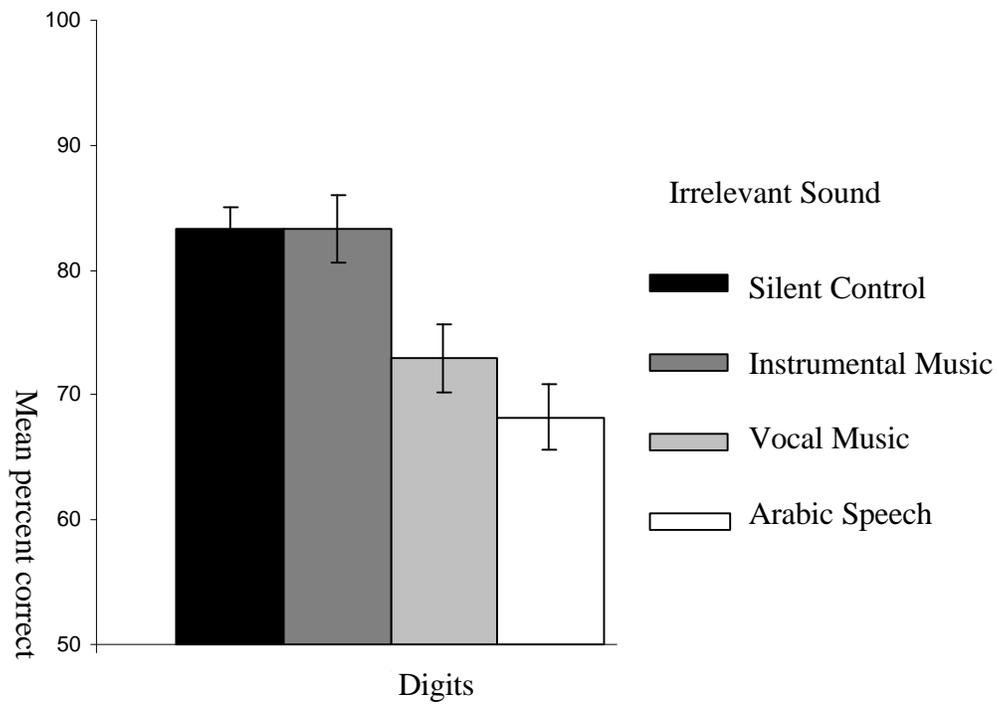


Figure 5.1: Experiment IIB: Mean percent correct on a visual-auditory digit recognition task in the presence of silence, instrumental music, vocal music, and Arabic speech irrelevant sound.

speech). Planned comparisons indicated that there was no difference between the control condition and the instrumental-music condition, but there was a difference between the control and vocal-music condition [$N - V: t(11) = 2.31, p < .05$], the control and the Arabic-speech condition [$N - S: t(11) = 3.56, p < .01$], the instrumental-music condition and the vocal-music condition [$I - V: t(11) = 2.28, p < .05$], and the instrumental-music condition and Arabic-speech condition [$I - S: t(11) = 3.45, p < .01$]. No other comparison was significant.

An additional single-factor (irrelevant sound type) ANOVA was run on d-prime scores. There was not a main effect of irrelevant sound type. Planned comparisons indicated that participants were more sensitive in the control condition compared to the Arabic-speech condition [$N - S: t(11) = 2.26, p < .05$]. The comparison between the instrumental-music condition and the vocal music condition approached significance [$I - V: t(11) = 1.61, p = .06$], and the comparison between the instrumental-music and the Arabic-speech conditions was significant [$I - S: t(11) = 2.73, p < .01$].

Finally, a single-factor (irrelevant sound type) ANOVA was run on the 3- point confidence ratings. There was not a main effect of irrelevant sound, but the confidence ratings were in the predicted order (control, instrumental, vocal, Arabic-speech conditions). Planned comparisons indicated that there were no significant differences in confidence ratings between any of the irrelevant-sound conditions.

DISCUSSION

The results of Experiment IIB replicated the digit results of Experiment IIA. The overall error pattern supported the similarity of content hypothesis, but all differences were not statistically significant. Statistically, there was again no difference between the silent-control and instrumental-music condition, and both vocal music and Arabic speech resulted in equivalent interference. Vocal music and Arabic speech results were both different from instrumental music and the silent-control condition results. As in Experiment IIA, vocal music resulted in equivalent interference to that of Arabic speech. This result supports the conclusion from Experiment IIA that for digits, vocal music probably does not act as an intermediate level of interference between instrumental music and Arabic speech.

The results of Experiments I, IIA, and IIB all show error patterns predicted by the similarity of content hypothesis. While the error percentages were in the predicted direction, some of the individual error percentage differences were not statistically significant. It could be the case that the different interference conditions were not different enough from each other to result in statistical significant differences. Experiment III was designed to investigate whether attempting to maximize the difference between interference conditions would result in predictable error patterns that were also statistically significant from each other. It was also designed to begin testing what exactly “overlap” between the ignored and remembered stimuli could mean.

CHAPTER 6

EXPERIMENT III

The purpose of Experiment III was to further test the hypothesis that differential overlap between the to-be-remembered and to-be-ignored stimuli significantly influenced performance in a serial recall task. Specifically, Experiment III investigated whether changes in pitch (tonal or phonemic pitch) overlap would result in different performance decrements. Experiment III was based on Salame and Baddeley's (1982) Experiment IV. In a nine-digit immediate serial recall task, they found similar disruption by irrelevant auditory digits and irrelevant auditory monosyllabic nonsense syllables, which were phonologically equivalent to the digits (e.g., tun, gnu, tee). These two stimuli both caused significantly more interference than phonologically dissimilar disyllabic words. This experiment was run to replicate the speech results and to extend them to music including auditory tone sequences as the to-be-remembered and to-be-ignored stimuli. The high-overlap speech interference consisted of a female voice speaking the same digit stimuli as the to-be-remembered stimuli. The low-overlap speech interference consisted of a female voice speaking random words that did not rhyme with the digits (Salame & Baddeley, 1982). The high-overlap tone interference consisted of a series of piano tones from the same set found in the to-be-remembered stimuli while the low-overlap tone interference

consisted of a series of piano tones with different frequencies outside the musical range of the to-be-remembered stimuli. It was hypothesized that both music and speech would demonstrate the characteristic breakdown pattern predicted in Experiments IIA and IIB - greater pitch overlap would result in greater interference.

Experiment III was also a test of the similarity of content and similarity of process hypotheses. The high-overlap stimuli were designed to be more homogenous thus increasing the phonological similarity among to-be-ignored items. The low-overlap stimuli were designed to be less homogenous thus decreasing the phonological similarity among to-be-ignored items. This is especially true with the tone stimuli. The C Major scale notes D4-E5 are in a small musical range and all part of a commonly heard Major scale. The notes outside of that range are not only in two completely separate musical ranges but make up a rarely heard Major scale. If the similarity of content hypothesis holds, then the high overlap should interfere more than the low overlap. If the similarity of process hypothesis holds, then at the very least, there should be no difference between the interference of each irrelevant sound type.

PARTICIPANTS

Thirty-two students participated in a between subject design. Sixteen were musically trained (mean experience = 9.17 years; range = 6-15 years) with the same musical experience requirements as in Experiment I, and they participated in the music half of the experiment. The other 16 participant's musical experience was irrelevant, as they

were participating in the digit half of the experiment. No participant from Experiments I, IIA, or IIB participated in Experiment III. Students received partial class credit for their participation.

MATERIALS AND APPARATUS

Fifty-three of the four-tone auditory standard/comparison stimuli from Experiment II were used again. Fifty-three of the six-digit trials from Experiment II were also used, but in this experiment both the standard and comparison sequences were presented auditorily. There were four irrelevant sound type conditions: high-overlap tones, low-overlap tones, high-overlap speech, and low-overlap speech. Care was taken to keep the word-dose (number of syllables in each interference condition), tone-dose (number of tones in each interference condition), and presentation rate constant between the irrelevant-sound conditions since it has been shown that an increase in irrelevant sound objects per unit time results in greater performance decrements (Bridges & Jones, 1996). The high-overlap tone recording was a randomly generated series of piano tones from the same set as the stimuli (nine tones of the C Major scale between D4 and E5) presented with 700-ms IOIs. The low-overlap tone recording was synthesized in a similar manner with 700-ms IOIs, but it was composed of nine random tones of the C# Major scale outside of the musical range of D4-E5 (e.g., C#4 and F#5). None of the tones in the low-overlap conditions are heard in the stimuli. The high-overlap speech recording was made by splicing together a random series of digits (between 1-9) spoken by a female speaker

with 700-ms IOIs. The low-overlap speech recording was composed of a series of nine random words (e.g., hike, dip) spoken by a female speaker spliced together with 700-ms IOIs. To match the nine digits, eight of the random words were monosyllabic, and one of the random words was disyllabic.

DESIGN

A 2 (stimulus type) x 3 (irrelevant sound type) mixed-factorial design was used. The two levels of the stimulus type variable were digits and tones. The three levels of the irrelevant sound type variable were silence, high-overlap irrelevant sound, and low-overlap irrelevant sound. The musically trained half of the participants only received the note trials paired with the silent, high-overlap tone, and low-overlap tone irrelevant-sound conditions. The other half of the participants only received the digit stimuli paired with the silent, high-overlap speech, and low-overlap speech irrelevant-sound conditions. The experiment was split into three irrelevant sound blocks. Each block contained 16 trials and was separated by short breaks (total of 48 trials). The other five trials were practice trials. The order of blocks and trials within a block was counterbalanced across participants.

PROCEDURE

The same auditory standard/comparison procedure from the Experiment IIA note trials was used, but the irrelevant-sound conditions were different.

RESULTS

It was unknown whether the high and low-overlap conditions within each stimulus type were equivalent forms of interference. Also, there were more digit items to memorize per trial (six digits) than tone items (four tones). Therefore, following the initial between factor ANOVA, all other analyses were run within each stimulus type.

A 2 (stimulus type) x 3 (irrelevant sound type) mixed-factorial ANOVA on percent correct was run. There was a main effect of stimulus type: $F(1,34) = 16.64, p < .01$, and a main effect of irrelevant sound type: $F(2,68) = 27.19, p < .01$. As can be seen in Figure 6.1, there was also a stimulus type by irrelevant sound type interaction: $F(2,68) = 3.6, p < .05$. There was a much greater drop in performance during the tone trials (difference between control and high-overlap conditions = 21.18%) than during the digit trials (difference between control and high-overlap conditions = 11.1%). In both digit and tone trials, planned comparisons indicated that performance in the silent-control condition was significantly better than the performance in the low-overlap condition (L), and that performance in the low-overlap condition was significantly better than performance in the high-overlap condition (H). [Digits: C – L: $t(17) = 1.96, p < .05$; L – H: $t(17) = 1.76, p < .05$; Tones: C – L: $t(17) = 5.03, p < .01$; L – H: $t(17) = 1.83, p < .05$].

Two additional single-factor (irrelevant sound type) ANOVAs were run on the percent correct scores separating the digit and tone subjects. The digit ANOVA showed a

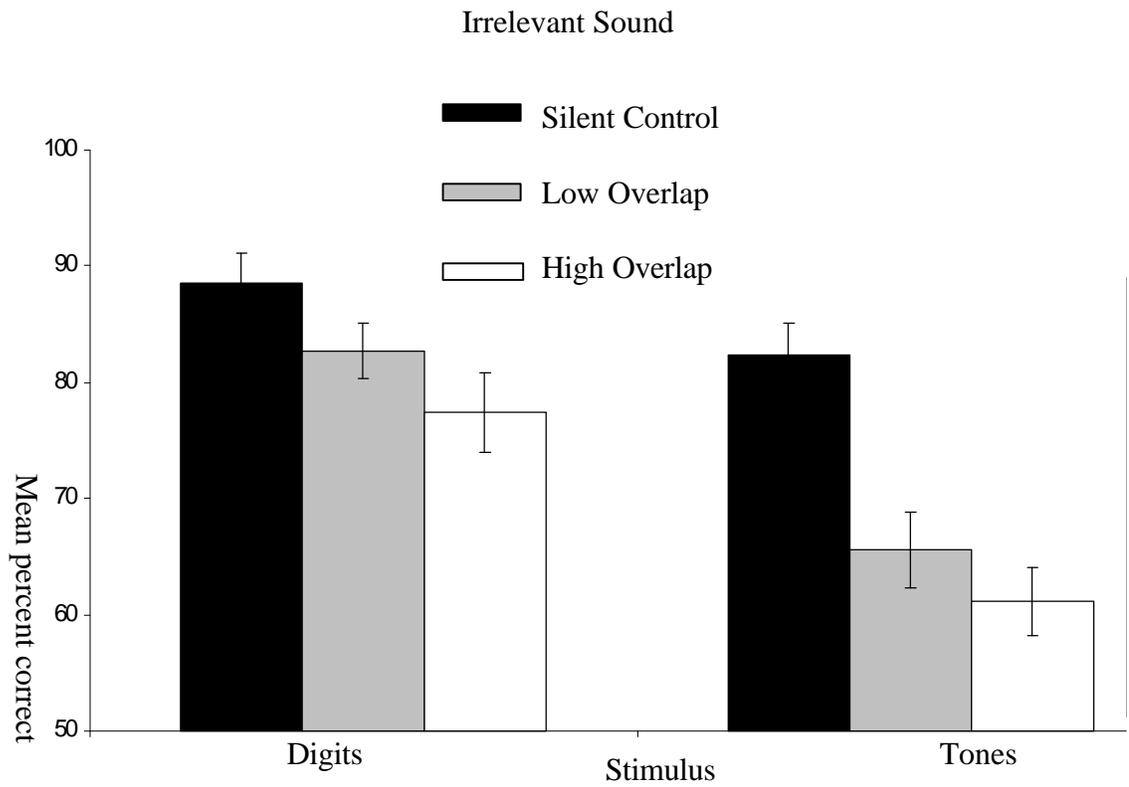


Figure 6.1: Experiment III: Mean percent correct on auditory standard/comparison task with digits and tones in the presence of silence, low overlap irrelevant sound, and high overlap irrelevant sound.

main effect of irrelevant sound type: $F(2,34) = 5.29, p < .01$. The tone ANOVA also showed a main effect of irrelevant sound type: $F(2,34) = 29.27, p < .01$. As indicated above, in both type of stimulus trials, the silent-control trials resulted in the top performance followed by the low-overlap condition, and then the high-overlap condition.

Two additional single-factor ANOVAs (irrelevant sound type) were conducted on d-prime scores. Each ANOVA resulted in a main effect of irrelevant sound type: Digits: $F(2,34) = 4.77, p < .05$; Tones: $F(2,34) = 18.95, p < .01$. In both ANOVAs, participants demonstrated the highest sensitivity in the silent-control condition, followed by the low-overlap and then high-overlap interference conditions. Planned comparisons on the d-prime scores for the digit results indicated that the difference between the silent-control and the low-overlap conditions approached significance [N – L: $t(17) = 1.5, p = .07$], while the difference in sensitivity between the low-overlap and high-overlap conditions was significant [L – H: $t(17) = 1.74, p = .05$]. Planned comparisons on the d-prime scores for the tone results indicated that the difference between the silent-control and the low-overlap conditions was significant [N – L: $t(17) = 4.34, p = .01$], while the difference in sensitivity between the low-overlap and high-overlap conditions approached significance [L – H: $t(17) = 5.62, p = .07$].

Two single-factor (irrelevant sound type) ANOVAs were conducted on participants' three-point confidence ratings given after each trial. Both ANOVAs resulted in main effects of irrelevant sound type: Digits: $F(2,34) = 11.58, p < .01$; Tones: $F(2,34) = 9.8, p < .01$. The digit results demonstrated the highest confidence ratings in the silent-control, followed by the low-overlap, then the high-overlap interference conditions. The

note results also demonstrated the highest confidence ratings in the silent-control condition, but that was followed by the high-overlap and then low-overlap conditions. Planned comparisons on the digit results indicated that there was a significant difference between the silent-control and both overlap interference conditions [N – L: $t(17) = 4.34$, $p = .01$; N – H: $t(17) = 5.62$, $p = .01$]. Planned comparisons on the tone results also indicated a significant difference between the silent-control and both overlap interference conditions [N – L: $t(17) = 3.36$, $p = .01$; N – H: $t(17) = 3.39$, $p = .01$]. There was no difference in confidence ratings for the low-overlap and high-overlap conditions in either stimulus type.

Following the experiment, participants were asked to indicate which section was the most difficult. In the digit experiment, 14 participants indicated the high-overlap condition was the most difficult. The other two participants indicated that the low-overlap condition was the most difficult, one of whom actually performed worst in the low-overlap condition. In the tone experiments, nine of the 18 participants indicated the high-overlap condition was the most difficult, five indicated the low and high-overlap condition were equally difficult, and four indicated the low-overlap condition was the most difficult. Of the five participants who indicated the low and high-overlap conditions were equally difficult, four performed worst in the high-overlap condition. Of the four that indicated the low-overlap condition was the most difficult, all actually performed worst in the high-overlap condition.

Of particular interest is that only six of the participants in the tone experiment could correctly indicate what the actual difference was between the low and high-overlap

conditions. Overall percent correct performance for these six participants was the same as the performance by the overall group: silent-control = 84.38%, low-overlap = 71.43%, and high-overlap = 65.63%. The theoretical implications for this result will be further discussed in the discussion section.

Finally, a series of correlational analyses on the musicians' performances indicated that the age at which a musician participant began learning their primary instrument correlated positively with the performance decrement found in the low-overlap trials ($r = .49, p < .05$). Participants who began performing at a later age had larger differences between performance on the control trials and performance on the low-overlap trials. The lack of other correlations seems to indicate a truncated sample problem; the musicians in this experiment were highly trained, and it is likely that more correlations would appear if a wider range of musically experienced participants were tested.

DISCUSSION

The results of the digit and tone trials from Experiment III replicated those attained by Salame and Baddeley (1982) in Experiment IV. A greater pitch overlap between the two-be-remembered stimuli and the to-be-ignored irrelevant sound resulted in larger performance decrements. Planned comparisons indicated that both digits and tones demonstrated this distinct pattern. The results of Experiment III support the similarity of content hypothesis.

The similarity of process hypothesis is predicated on the idea that increased phonological dissimilarity among the items of the irrelevant stream will result in greater interference. As indicated by the participant exit interviews, Experiment III's tone overlap conditions were practically indistinguishable from one another. And, when a participant correctly discriminated between the two overlap conditions, there was no difference between overall group performance in the three conditions and the performances of the six participants who correctly identified the difference. And, even if all participants could tell the tone conditions were different, but could not explain what the difference was, the results are still not indicative of the similarity of process hypothesis. One would predict that the order of tones in a sequence from a particular common key in a restricted range (e.g., the high-overlap condition nine C Major scale tones from D4 to E5) would be more homogenous than a tone sequence from a less restricted range (e.g., the low-overlap condition nine C# Major scale tones – four lower than C4 and five higher than F5). But, the low-overlap condition did not result in greater interference as would be predicted by the similarity of process hypothesis, instead the high-overlap condition resulted in the greatest interference as would be predicted by the similarity of content hypothesis.

CHAPTER 7

GENERAL DISCUSSION

The results of Experiment I's digit trials replicated those of Salame & Baddeley (1989). Instrumental music, vocal music, and Arabic speech all impaired performance, but Arabic speech impaired performance more than either music condition. Looking closely at individual patterns of data in the note condition, some participants also showed this same interference pattern while others showed the exact opposite interference pattern. This result was explained in terms of the various ways in which visual notes can be rehearsed. It was hypothesized that notes rehearsed as note names would result in interference patterns similar to the digit patterns while notes rehearsed as tones would result in the opposite interference pattern.

Experiment IIA and IIB demonstrated that when participants rehearsed digits and tones in an auditory form, greater acoustic overlap was the best predictor of performance. In the digit trials, Arabic speech and vocal music impaired performance the most. Instrumental music also impaired performance, but to a lesser degree. In tone trials the order of interference was the opposite: instrumental music, vocal music, and then Arabic speech.

There could be a problem with the experimental design that was chosen for Experiment IIA. In Experiment IIA, the digit trials were six digits long while the tone trials were four tones long. Also, musicians participated in the tone half of the experiment while musicianship was not a factor for the digit half of the experiment. Finally, the digit standard sequence was presented visually and the irrelevant sound occurred during sequence presentation and the retention interval, but the tone sequence was presented auditorily and the irrelevant sound only occurred during the retention interval. It is possible that one or more of these differences could have influenced the pattern of results found in Experiment IIA. Further experiments in which musicians run in the digit half of the experiment, the standard sequences are presented auditorily, and irrelevant sound only occurs during the retention interval should clear these potential confounds up. The digit sequences will remain six digits in length to account for the ceiling effect found when digit sequences are only four digits in length.

Experiment III showed that when one maximized the pitch overlap difference between the irrelevant-sound conditions, not only is the predicted pattern found, but each condition is statistically different from each other condition. Digit results were a replication of Salame and Baddeley (1982) – greater pitch overlap resulted in greater interference. Tone results also replicated this pattern.

Given the similarity of content and similarity of process theories, which theory can best account for the results of the current experiments? Both theories make valid contributions to account for the current results, but at the same time, both theories are partially based on ideas that are inconsistent with the current results. Each theory's

strengths and weaknesses are discussed below with an eye towards a new, integrated theory termed the Acoustic Overlap hypothesis.

THE SIMILARITY OF CONTENT HYPOTHESIS

The basic idea behind the similarity of content hypothesis (Salame & Baddeley, 1989) most accurately accounts for the results of these experiments. When there was more content overlap between the to-be-remembered and to-be-ignored stimuli, there was more interference. While it was found in the overall error patterns in Experiments I, IIA, and IIB, the effect was especially evident in Experiment III for both digit and tone trials because the experiment was set up to create maximal difference between the interference conditions. Unfortunately, two problems are evident with the similarity of content hypothesis based on the current experiments' extension of the irrelevant sound effect paradigm to music. First, the theory was developed based solely on the interference-based breakdown patterns found in language stimuli. Secondly, it only makes predictions about what "differential overlap" means in a speech context, and these predictions are not very specific.

The similarity of content hypothesis is based solely on experiments using language as stimuli. This is also one of Jones and Macken's (1993) main criticisms of the theory. The two previously reviewed hypotheses about the functions of the phonological store posed by Salame and Baddeley (1989) (the Noise Filter hypothesis and the Speech Detector hypothesis) both claim that the phonological loop is only affected by language.

They proposed that the phonological store works either by filtering out non-speech like sounds or by being specifically tuned to detecting speech-like sounds in the auditory signal. Jones and Macken, Salame and Baddeley, and the current experiments all found that irrelevant non-speech also impaired performance in these tasks.

These findings support the hypothesis of Schendel and Palmer (in press) who have theorized that the phonological loop is more than “phonological.” They found that music rehearsed in an auditory form was impaired in the same characteristic ways as speech in the presence of articulatory suppression. The results of the current experiments support the Schendel and Palmer conclusion since they have further demonstrated that music and language show similar breakdown patterns in the presence of a different standard working memory paradigm: the irrelevant sound effect.

The similarity of content hypothesis only makes predictions about what “overlap” means in a speech context, and these predictions are not specific. Outside the realm of speech, what does it mean to say that the differential overlap between the to-be-remembered and to-be-ignored stimuli results in differential interference? The original hypothesis was formed (Salame & Baddeley, 1989) based on the conclusion that the more speech-like the interference was, the more overlap and interference would occur when memorizing speech sequences. But, given the results of the current experiments, this is only half of the explanation since the more speech-like interference did not impair performance in the auditory tone trials. Instead, more similar background sounds like instrumental music and vocal music resulted in the greatest performance decrements.

Differential overlap could be in the form of pitch (tonal or phonemic), rhythm, timbre, or perhaps a combination of factors in the overall acoustic signal. During the tone trials in Experiment III, it was discovered that high and low tonal pitch overlap was one element of the acoustic signal that can be manipulated in order to result in significant interference differences. Experiment III's digit trials, a replication of Salame and Baddeley (1982), found that high and low tonal phonemic overlap also resulted in significant interference differences. A series of experiments by Saito (1993; 1994) found that syncopated tapping was more disruptive to serial recall of letter sequences than steady state tapping, because generation of syncopated rhythms are presumed to be akin to speech generation (Saito, 1993; 1994). This is also true with irrelevant sound. That is, syncopated irrelevant sound was found to be more disruptive than steady-state irrelevant sound (Jones & Macken, 1995, Larsen & Baddeley, 2003). This could be the beginning of research investigating a rhythmic overlap between the to-be-remembered and to-be-ignored stimuli.

Other factors listed above may also be involved, and a series of experiments could help shed light on the influence these factors have on performance in irrelevant sound tasks. For example, these experiments could be conducted in order to investigate the effect timbre has on performance. One could have participants memorize auditory piano sequences in the presence of either high or low-overlap interference (as defined in Experiment III) played by another piano or an instrument with a different timbre such as a bassoon. Memorizing auditory bassoon sequences in the presence of piano and bassoon interference could provide an interesting comparison. In the same vein, one could have

participants memorize auditory digit sequences spoken by a female with high and low-overlap interference spoken by the same female or a male.

Also, it is possible that rehearsal of acoustic information other than speech and music could occur. For example, imagine a participant were trying to remember a bird song or the way a car's engine sounded over a short retention interval. If the participant were forced to listen to other bird songs or other car engines, it is possible that greater acoustic overlap could interfere in these situations also.

At the very least, these experiments have demonstrated that an accurate model of working memory must take into account both speech and music findings. Whether speech and music are governed by separate working memory systems that show parallel interference patterns in the presence of articulatory suppression (Schendel & Palmer, in press) and irrelevant sound (the current experiments), or are both governed by the same phonological or acoustic loop, remains a difficult question to answer experimentally. But, given the result that music does interfere with speech and speech does interfere with music, it is most likely the case that a single phonological or acoustic store is responsible for the working rehearsal of both language and music.

THE SIMILARITY OF PROCESS HYPOTHESIS

The main theory behind the similarity of process hypothesis (Jones & Macken, 1993) can also account for most of the findings in the current experiments. The hypothesis was developed to account for the finding that changing-state, irrelevant non-

speech also interfered with working memory for speech. And, most importantly, order confusions in the current experiments brought on by overlapping to-be-remembered and to-be-ignored stimuli during the retention interval could have resulted in an increase in errors. The Object-Oriented Episodic Record hypothesis can account for decreases in performance in serial recall tasks in the presence of irrelevant speech. Jones and Macken (1995) proposed that memorial objects representing the items in a to-be-remembered serial sequence are formed and bonds between the objects remain strong through rehearsal. Irrelevant sound also has objects which are bound together in a serial order. The confusion between the correct order bonds and the bonds that form by irrelevant sound object intrusions into the original object structure result in an increase in errors. As discussed in the introduction, this result has been widely supported (Jones & Macken, 1995; Jones, Macken, & Harries, 1997).

Future experiments could test this theory of the O-OER hypothesis by analyzing the types of errors participants make in a serial recall task in the presence of irrelevant sound. If it is true that confusions between to-be-remembered and to-be-ignored stimuli bonds result in an increase in errors, then it should be that there is an increase in intrusion errors from the items in the to-be-ignored stimulus in recall. If there were just an increase in order errors among the items in the to-be-remembered stimulus, then it would not provide support for the O-OER hypothesis.

Unfortunately, the similarity of process hypothesis is based on the changing state hypothesis, which was later molded into the phonological dissimilarity effect. The changing-state hypothesis was not supported in any of the current experiments. Not only

did different types of irrelevant sound that changed in state interfere with the memory tasks, but in most experiments the irrelevant sound interfered in predictable patterns. As noted in the introduction and discussion section of Experiment II, a lack of statistical differences between conditions could have partially led Jones and Macken (1993) to conclude their results were the result of process similarities. Knowing that the statistical difference between performance in irrelevant-sound conditions depended on what was used as irrelevant sound could have led Jones and Macken (1993) to a different conclusion.

The phonological dissimilarity effect states that an increase in dissimilarity between the objects that make up the irrelevant stream will result in an increase in errors in serial recall tasks. This hypothesis was tested in Experiment III when the phonological dissimilarity effect of the similarity of process hypothesis was compared to the phonological overlap effect of the similarity of content hypothesis. In Experiment III, the high-overlap tone condition had much smaller and less dramatic changes compared to the low-overlap condition. The similarity of process explanation proposes a phonological dissimilarity effect in which sounds within the irrelevant stream that are more different from one another will result in more interference than sounds that are more similar to each other. But, it was found that the low-overlap tones produced significantly less interference than the high-overlap sounds supporting the similarity of content hypothesis.

THE ACOUSTIC OVERLAP HYPOTHESIS: TOWARDS AN INTEGRATED THEORY

It is apparent that these two existing theories of working memory investigated with the current experiments cannot completely account for all of the results. Both theories have ideas that were supported while both theories also have ideas that were not supported. The following is an outline for an integrated theory that will be referred to as the Acoustic Overlap hypothesis.

The current experiments suggest that the term “phonological” loop or “phonological” store may be misleading; the term “phonological” implies language or speech. An acoustic loop or acoustic store are probably more appropriate terms given that music has demonstrated the same pattern of results as speech in a series of experimental working memory replications and extensions. The “articulatory control process” seems to still be a valid term given the articulatory nature of rehearsals of both language and music. Keeping the auditory memorial representation fresh through the rehearsal of both speech and music requires articulation of what the to-be-remembered items would sound like.

Converging evidence from articulatory suppression experiments (Schendel & Palmer, in press) and these irrelevant sound effect experiments suggest that a single acoustic loop with a single rehearsal mechanism is the best possible explanation to account for both language and music results. If there were separate music memory (Berz, 1995) or tonal (Pechmann & Mohr, 1992) loops for the rehearsal of music separate from

language, then singing (Schendel & Palmer, in press) or listening to music would most likely not impair performance in language working memory tasks as severely as they did. Also, it seems that hypothesizing separate rehearsal mechanisms within the same loop (Pechmann & Mohr, 1992) is not the best explanation for results. If it were the case that separate rehearsal mechanisms for language and music existed, then participants would be able to simultaneously rehearse language and music sequences with minimal performance loss for either stimulus. Future experiments can investigate this idea more thoroughly; the current experiments suggest that if both the hypothesized “articulatory” and “tonal” rehearsal mechanisms were impaired simultaneously with vocal music, performance should be significantly worse when compared to only instrumental music or Arabic speech.

Similarly, defining the properties of the acoustic store in terms of speech experiment results, while a good start, is not the entire story. The similarity of content hypothesis is supported in the current experiments with both digit and tone results, and the main hypothesis behind the similarity of content hypothesis should remain as the central component to any theory behind the workings of the acoustic store. The similarity of process hypothesis, while important for other reasons discussed later, is partially based on the changing state hypothesis which has repeatedly been disproved in this series of experiments. Therefore, it is concluded that the differential overlap between the to-be-remembered and to-be-ignored stimuli results in differential interference.

How this differential interference translates to errors is best described by the similarity of process hypothesis. Order confusions (seen in decreased accuracy scores) are best explained by the similarity of process hypothesis. The Speech Detector and the

Noise Filter hypotheses are no longer valid representations of the functions of the acoustic store. It appears as if all acoustic information has access to the acoustic store, not just acoustic information that has features common to speech. It is hypothesized that all acoustic information gains access to the acoustic store, but the irrelevant acoustic information that shares more features with the rehearsed information leads to an increase in order confusions.

In summary, the Acoustic Overlap hypothesis is based on a single acoustic store with a single rehearsal mechanism in charge of the working rehearsal of language and music. All auditory information has access to the acoustic store. The increased acoustic overlap between irrelevant auditory information and rehearsed auditory information results in an increased impairment. The overlap results in impairment as the result of order confusions that arise during the rehearsal process.

Several questions remain to be resolved for the integrated Acoustic Overlap hypothesis to hold. For example, what exactly does “acoustic overlap” mean? Experiment III has shown that pitch overlap is sufficient in significantly impairing working memory representations of language and music. Saito (1993; 1994) has demonstrated a possible rhythmic overlap. New experiments that were above have shown ways in which timbre can be manipulated to discover if its systematic manipulation could result in differential overlap and differential interference. And overlap does not have to stop with the acoustic components of the sound. Perhaps it is the meaning of words chosen in an irrelevant sound experiment that would impair performance. For example, if one were holding a series of words like “doctor, window, sleep, hair” in working memory, and

irrelevant sound contained the words “nurse, pane, rest, beard.” How well would one perform at that task when compared to listening to irrelevant words like “bridge, lady, grass, spoon?” Finally, how would these semantic changes interact with a change in timbre?

SUMMARY

It was shown through a series of experiments that memory for digits and memory for tones were impaired by irrelevant sounds. The specific patterns of interference found in all experiments were representative of the similarity of content hypothesis: the greater acoustic overlap between the to-be-remembered and to-be-ignored stimuli resulted in greater performance decrements. The contributions of the phonological loop (similarity of content) and the O-OER hypotheses (similarity of process) to the explanation of these new results were discussed. A new hypothesis termed the Acoustic Overlap Hypothesis was proposed drawing on previous research findings with language and extending them to music to account for the current results.

BIBLIOGRAPHY

Ashcraft, M. H. (1994). *Human Memory and Cognition*. New York: Harper Collins College Publishers.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its controlled processes. In W. K. Spence & D. T. Spence (Eds.). *The psychology of learning and motivation: Advances in research and theory*. New York, Academic Press.

Baddeley, A. D. (1986). *Working Memory*. New York: Oxford University Press.

Baddeley, A. D. (1990). *Human Memory: Theory and Practice*. Hove: Erlbaum.

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

Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation*, New York: Academic Press.

Baddeley, A. D. & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485-493.

Baddeley, A. D. & Larsen, J. D. (2003) The disruption of STM: A response to our commentators. *The Quarterly Journal of Experimental Psychology*, 56A, 1301-1306.

Beaman, C. P. & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 459-471.

Berz, W. A. (1995). Working memory in music: A theoretical model. *Music Perception*, 12, 353-364.

Bridges, A. M. & Jones, D. M. (1996). Word dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing state? *The Quarterly Journal of Experimental Psychology*, 49A, 919-939.

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

- Cole, R. A. (1973). Different memory functions for consonants and vowels. *Cognitive Psychology*, 4, 39-54.
- Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior*, 19, 722-735.
- Colle, H. A. & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, 15, 17-31.
- Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, 168, 1605-1605.
- Elliot, E. M. & Cowan, N. (2005). Coherence of the irrelevant-sound effect: Individual profiles of short-term memory and susceptibility to task-irrelevant materials. *Memory & Cognition*, 33, 664-675.
- Gathercole, S. E. & Baddeley, A. D. (1993). *Working Memory and Language*. Hillsdale, NJ: Erlbaum.
- Glanzer, M. & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5, 351-360.
- Hanley, J. R. & Broadbent, C. (1978). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology*, 78, 287-297.
- Hughes, R. W., Tremblay, S., & Jones, D. M. (2005). Disruption by speech of serial short-term memory: The role of changing-state vowels. *Psychonomic Bulletin & Review*, 12, 886-890.
- Jones, D. M. (1993). Objects, streams, and threads of auditory attention. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control*. Oxford, UK: Clarendon Press.
- Jones, D. M. & Macken, W. J. (1993). Irrelevant tones produce and irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 369-381.
- Jones, D. M. & Macken, W. J. (1995). Phonological similarity in the irrelevant speech effect: Within- or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 103-115.

Jones, D. M., Macken, W. J., & Harries C. (1997). Disruption of short-term recognition memory for tones: Streaming or interference? *Quarterly Journal of Experimental Psychology*, 50A, 337-357.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 656-674.

Larsen, J. D. & Baddeley, A. D. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source? *The Quarterly Journal of Experimental Psychology*, 56A, 1249-1268.

Levy, B. A. (1975). Vocalization and suppression effects in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 304-316.

Macken, W. J. & Jones, D. M. (2003). Reification of phonological storage. *The Quarterly Journal of Experimental Psychology*, 56A, 1279-1288.

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory & Language*, 27, 382-398.

Miles, C., Jones, D. M., & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. *Experimental Psychology: Learning, Memory, and Cognition*, 17, 578-584.

Murdock, B. B. Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64, 482-488.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, 21, 263-267.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 78, 679-684.

Pechmann, T. & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory and Cognition*, 20, 314-320.

Roberts, L. A. (1986). Modality and suffix effects in memory for melodic and harmonic materials. *Cognitive Psychology*, 18, 123-157.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, 89, 63-77.

Salame, P. & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 150-164.

Salame, P. & Baddeley, A. D. (1987). Noise, unattended speech and short-term memory. *Ergonomics*, 30, 1185-1194.

Salame, P. & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, 41A, 107-122.

Salame, P. & Baddeley, A. D. (1990). The effects of irrelevant speech on immediate free recall. *Bulletin of the Psychonomic Society*, 28, 540-542.

Salame, P. & Wittersheim, G. (1978). Selective noise disturbance of the information input in short-term memory. *Quarterly Journal of Experimental Psychology*, 30, 693-704.

Saito, S. (1993). The disappearance of phonological similarity effect by complex rhythmic tapping. *Psychologia*, 36, 27-33.

Saito, S. (1994). What effect can rhythmic tapping have on the phonological similarity effect? *Memory & Cognition*, 22, 181-187.

Saito, S. (1997). When articulatory suppression does not suppress the activity of the phonological loop. *British Journal of Psychology*, 88, 565-578.

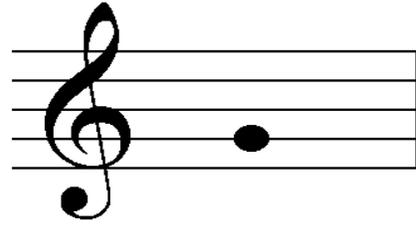
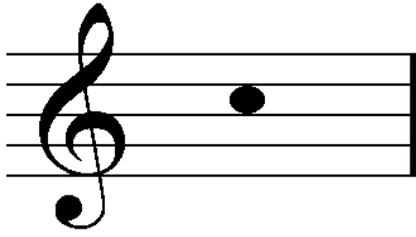
Saito, S. (1998). Phonological loop and intermittent activity: A whistle task as articulatory suppression. *Canadian Journal of Experimental Psychology*, 52, 18-23.

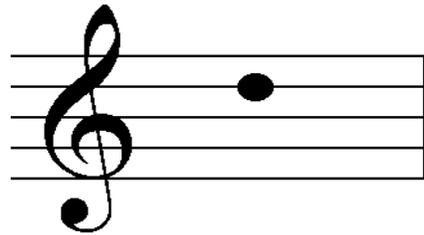
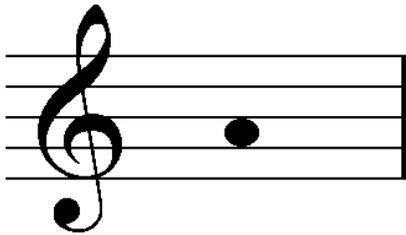
Schendel, Z. A. & Palmer, C. (in press). Suppression effects on musical and verbal memory. *Memory & Cognition*.

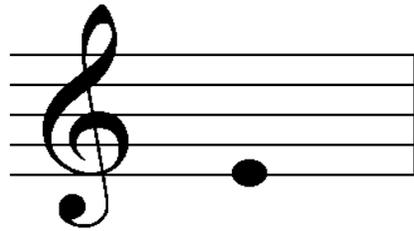
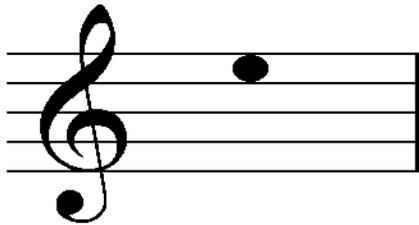
Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, 28, 1-16.

APPENDIX A

MUSIC PRE-TEST







APPENDIX B

BACKGROUND QUESTIONNAIRE

Background Questionnaire

Please take a minute to answer the questions below about your musical background and training. Thank you.

- 1) What musical instrument(s) do you play? How many years have you been playing each instrument?

- 2a) How many years of private instruction on each instrument have you had?

- 2b) At what age did you start lessons?

- 3) Do you still play these instruments? If so, which ones and how many hours a week?

- 4) How often do you take music lessons?

- 5) Do you have any hearing problems?

- 6) Do you have perfect pitch?

Zachary Schendel
Dr. Neal Johnson, Principal Investigator
Psychology Department
Ohio State University
Phone: 292-3170