USE OF PHONOLOGICAL AND LEXICAL CUES IN THE RESOLUTION OF PERCEPTUAL-LEXICAL AMBIGUITIES BY LISTENERS DIFFERING IN WORKING MEMORY SPAN

by

VALERIE LOUISE LLOYD

B.A., The University of British Columbia, 1991
M.A., The University of British Columbia, 1994

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Psychology)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

31 August, 2001

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date 20 August 2001

DE-6 (2/88)

ABSTRACT

The purpose of this research is to examine why there are differences in language processing between listeners with high and low working memory spans (WMS). Possible reasons include: differences in auditory perception of the speech signal, differences in coding the signal at the phonological, lexical and/or sentence levels, interaction between processing at different levels, differences in rehearsal capacity for speech, and/or, differences in generating and/or evaluating alternative interpretations to ambiguous signals. With regard to the matter of how processing at different levels might interact, the following questions are examined: Which aspects of spoken language processing interact with each other? To what extent do they interact with each other? and, Does the pattern of interaction differ for groups that vary in working memory limitations?

This research makes new contributions in two ways: 1) the relationship between working memory span and the domain of listening is examined, (rather than reading, on which most of the research has been focussed to date), and 2) the linguistic signal presented is made ambiguous through perceptual degradation in order to examine the relationship between perception, working memory and language comprehension.

In the first three experiments, target words which varied in terms of a minimal pair contrast were embedded in a variety of sentences and presented in varying degrees of background noise to high and low WMS adult listeners. As the background noise (and consequently the acoustical confusions of speech sounds) increased, the target words were rendered increasingly ambiguous. The listeners were asked, in each experiment, to identify the target word. The sentences varied in the degree to which they contained contextual information at the lexical level in Experiment 1, and at the sentence level in Experiments 2 and 3. The manner in which high and low WMS listeners used phonological, lexical and sentence level information processing components was compared.

The fourth experiment employed a gating task (following Grosjean, 1980), in which the initial portion of the word increased in duration on each trial until the word was correctly identified

by the listener. Thus it was possible to directly assess whether listeners generate multiple alternatives to ambiguous words (Carpenter et al., 1995) or evaluate alternatives differently (e.g., through inhibitory mechanisms; Stoltzus, Hasher, & Zacks, 1996). In Experiments 2-4, temporal auditory processing (via gap-detection scores) was also compared to comprehension performance in order to evaluate whether it could help to explain group differences in processing ambiguous spoken language.

The results of the first and fourth experiments provide evidence that high WMS listeners discriminate phonological cues better, generate more alternatives to ambiguous words, are better at inhibiting incorrect responses, and, overall, process phonological cues relatively independently of lexical knowledge. The second, third and fourth experiments showed that sentence level knowledge interacts with processing of phonological cues similarly for both WMS groups and that high WMS listeners demonstrate superior rehearsal capacity for phonological information. Temporal auditory processing as measured by gap detection does not appear to be related to spoken language processing for either group. Therefore, the results show that the differences in language processing between high and low WMS groups appear to lie in the interaction between lexical and phonological processing, and generation and evaluation of alternatives, but fail to provide evidence that there are differences in sentence processing or temporal auditory processing. The results of this research contribute to relatively unexplored aspects of existing WMS models.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Dr. Kathy Pichora-Fuller (Ph.D. Supervisor) and my other Ph.D. committee members: Susan Butt, Jeff Small, and Lawrence Ward for their scholarly guidance and support. I am also indebted to the undergraduate students at the University of British Columbia who took the time to participate in these studies. Many thanks are also extended to Trudy Adam, Ann-Marie Bowes, Arlene Carson, Lisa Dillon, John Gilbert, Natalie Haubert, Carolyn Johnson, Ruth Kirson, Kelly Kovach, John Nicol, Hollis Pass, Christian Richard, Kristen Roodenburg, Rushen Shi, Christiane Spanik, Glynnis Tidball, Michelle Wiley, and Anita Wong for their friendship, encouragement, and theoretical, methodological and technical contributions to this research. Thanks also to Sue Bryant and Michelle McCaughran from the School of Audiology and Speech Sciences and Rose Tam from the Psychology Department at UBC for administrative support.

In addition, I would like to thank my family, many friends, students and colleagues for their enduring support over the years that this dissertation was in progress. Most of all, I would like to thank my son, Jaman Lloyd, for his patience while his mom spent endless hours "at the lab" throughout his childhood.

This research was supported by a Natural Sciences & Engineering Research Council of Canada (NSERCC) Collaborative Grant to M.K. Pichora-Fuller, an NSERCC PGS-B Doctoral Scholarship and several research grants from Langara College to V. Lloyd.

I INTRODUCTION

1.1 Purpose

The purpose of this research is to examine why there are differences in spoken language processing between listeners with high and low working memory spans (WMS). Are the differences related to auditory perception of the speech signal, processing the signal at the phonological, lexical or sentence level, interaction between processing at different levels, rehearsal capacity, and/or, generation and/or evaluation of alternative interpretations to ambiguous signals? With regard to the matter of how processing at different levels might interact, the following questions are examined: Which aspects of spoken language processing interact with each other? To what extent do they interact with each other? and, Does the pattern of interaction differ for groups that vary in working memory limitations? All of these questions are addressed in four experiments.

1.2 Rationale

Over the past 20 years there has been extensive research conducted on the relationship between language comprehension and working memory. The relationships between the components of working memory and their operation during spoken language comprehension, however, are still not fully understood. There are differing views on whether and how the subcomponents of working memory interact and whether capacity limitations lie in general processing mechanisms or solely within subcomponent processing. Although the models and definitions of "working memory" are numerous, one general definition that has evolved over many decades and is generally accepted for its comprehensiveness is Baddeley and Logie's (1999) contemporary definition:

> Working memory comprises multiple specialized components of cognition that allow humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals.
>
> (p. 28)

By comparing the differences in spoken language processing between young, normal-hearing listeners with high and low WMS, it is hoped that we will better understand the relative degree to which the components of working memory interact during language processing, which components share the same pool of working memory resources, and whether group differences in language processing can be explained by how alternative interpretations of the linguistic signals are generated or evaluated during processing.

Specifically, one question the present research was designed to investigate is whether group differences in spoken language comprehension by high and low WMS listeners can be partially explained by differences in auditory processing of the speech signal; i.e., is there a specialized auditory component within working memory? A second question of interest is whether high and low WMS listeners process phonological cues differently as a result of working memory constraints, thus contributing to differences in spoken language comprehension. The third and fourth questions that are addressed are whether high and low WMS listeners are influenced differentially by previously-acquired lexical and/or sentence level knowledge while processing incoming spoken language. In addition, a fifth question that is investigated is the extent to which auditory, phonological, lexical, and sentence level cues interact with each other during spoken language processing, and whether they share the same pool of working memory resources (that is, do the components interact, and does the extent to which the components interact differ for high and low WMS listeners?). A sixth question that is addressed is whether rehearsal capacity of the incoming linguistic information may help explain differences in comprehension between high and low WMS listeners.

In order to investigate the above questions, listeners were asked to listen to and identify target words presented under conditions that rendered the word ambiguous (i.e., presentation in background noise, or as incomplete segments in a gating task). These methods for investigating spoken language comprehension have been used for many decades, but seldom in the study of working memory. By implementing a gating task, a seventh question that is addressed is whether spoken language comprehension differences between high and low WMS listeners can be explained by the number of alternatives generated to an ambiguous target word and/or evaluation of the alternatives generated to the ambiguous target word.

1.3 Background

Several topics relevant to the proposed research will be discussed in this chapter. First, a review of Baddeley and Logie's (1999) model of working memory and its subcomponents will be given. It will be followed by a review of the 20-year-old debate over how spoken language is processed, including a presentation of key models of discourse processing (Kintsch, 1988; 1998; Kintsch & van Dijk, 1978), a discussion about what is known about individual differences in WMS and language comprehension (e.g., Engle, Cantor, & Carullo, 1992; Ericsson & Kintsch, 1995; Just & Carpenter, 1992; Salthouse, 1996; Stoltzfus, Hasher, & Zacks, 1996; Wingfield, 1996), a summary of the modularity vs interaction debate (e.g., Caplan & Waters, 1995; 1999), and a reflection on cognitive models of speech processing (e.g., Luce, 1986; Marslen-Wilson, 1987; McClelland & Elman, 1986). Second, a description of various types of linguistic ambiguity will be given, including perceptual-lexical ambiguity -- the type of linguistic ambiguity manipulated in these experiments. Third, the research plan will follow, including justification for choices made in the design of the studies. The justifications will appeal to literature examining the relationships between WMS and auditory, phonological, lexical, and, sentence level processing in spoken language and how differences in rehearsal capacity within the phonological loop may help to explain differences in spoken language comprehension between high and low WMS listeners. The need to determine whether individuals differ in generating or evaluating alternative responses to ambiguous words in spoken language will be framed using the models by Just and Carpenter (1992) and Stoltzfus, Hasher, and Zacks (1996). A brief description of the four experiments and the hypotheses to be tested will then be presented.

1.3.1 Working memory and its subcomponents

For almost 30 years, Alan Baddeley and his colleagues have been studying the processing and storage components of working memory (e.g., Baddeley & Hitch, 1974; Baddeley & Logie, 1999). Given that this is the most long-standing and widely used model of working memory, it will be referred to extensively in order to address the research questions posed in the present research.

The most current version of Baddeley and Logie's (1999) model illustrates the disunity in working memory in terms of the central executive, phonological loop, and visuo-spatial sketchpad. They describe key features of the model as follows:

> (1) According to our view, working memory comprises multiple specialized components of cognition that allow humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals.
>
> (2) These specialized components include both a supervisory system (the central executive) and specialized temporary memory systems, including a phonologically based store (the phonological loop) and a visuospatial store (the visuospatial sketchpad).
>
> (3) The two specialized, temporary memory systems are used to actively maintain memory traces that overlap with those involved in perception via rehearsal mechanisms involved in speech production for the phonological loop and, possibly, preparations for action or image generation for the visuospatial sketchpad.
>
> (4) The central executive is involved in the control and regulation of the working memory system. It is considered to play various executive functions, such as coordinating the two slave systems, focusing and switching attention, and activating representations within long-term memory, but it is not involved in temporary storage. The central executive in principle may not be a unitary construct, and this issue is a main focus of current research within this framework.
>
> (5) This model is derived empirically from studies of healthy adults and children and of brain-damaged individuals, using a range of experimental methodologies. The model offers a useful framework to account for a wide range of empirical findings on working memory.
>
> (Baddeley & Logie, 1999, p. 28)

Because the storage and processing components of the phonological loop, and the processing component of the central executive are most relevant to the present research, these components of the model will be described in detail.

The basic mechanisms of the phonological loop and central executive. The phonological loop is fractionated into a passive, phonological store and an active rehearsal system. The central executive provides the mechanism for control processes in working memory, including the coordination of the subsidiary memory systems, the control of encoding and retrieval strategies, the switching of attention, and the mental manipulation of material held in the slave systems (Baddeley, 1996). The central executive itself, however, is not equipped with supplementary storage capacity. Because the original model, Baddeley and colleagues have abandoned the assumption that the central executive itself stores information, proposing instead that any increase in total storage capacity beyond that of a given slave system, such as the phonological loop, is achieved by accessing either long-term memory or other subsystems. [Daneman and Carpenter's (1980) resource model, however, suggests that lexical and sentence level information are both processed and stored by a central executive in working memory rather than by separate slave systems such as those proposed by Baddeley and Logie (1999)].

The nature of working memory limitations. With respect to the nature of working memory limitations, Baddeley & Logie (1999) stated that

> Each component has constraints commensurate with the specialist function that it provides, and these constraints may arise from capacity for activation [emphasis mine] (with both the amount and duration of activation being limited) or from capacity for rehearsal [emphasis mine], or from capacity for complexity of material, or from the extent to which they are supported by acquired strategies and prior knowledge.
>
> (p. 31)

Baddeley and Logie (1999) further claim that individual differences in the phonological loop capacity reflect the amount of memory activation available, partly as a result of genetic factors or as a result of brain damage. In addition to limits on degree of activation, subjects may also differ in their rehearsal capacity of items activated within the phonological loop. Having a superior capacity to activate and rehearse incoming speech signals implies a more efficient language processing system.

The role of working memory in complex cognitive activities. Working memory plays a role in complex cognitive activities, particularly language comprehension. Specifically, the evidence indicates that processing and storage by the phonological loop is essential to the detection of errors in word order during reading although phonological loop processing is distinct from sentence level processing and reading comprehension (e.g., Baddeley, Eldridge, & Lewis, 1981; Vallar & Shallice, 1990). In contrast to the phonological loop, the central executive is believed to play an important role in comprehension at the sentence and discourse levels (Gathercole & Baddeley, 1993, as cited in Baddeley & Logie, 1999). Specifically, the central executive activates representations in long-term memory from individual words and concepts to complex schemata. Comprehension of a particular passage will depend on the existing representations in long-term memory and on the capacity of the central executive to activate representations stored by the relevant working memory components and combine them into a coherent mental model, which can then be consolidated into long-term memory. Of course, comprehension also depends on processing of external input. Studies in spoken language comprehension, however, have seldom manipulated the quality of the signal or otherwise tested the extent to which perceptual processing influences interpretation of ambiguous speech. The present research was designed to do precisely this.

The relationship of working memory to long-term memory and knowledge. Baddeley and Logie (1999) have provided evidence for the hypothesis that temporary memory is influenced by long-term knowledge as well as the operation of the components of working memory. Although some have viewed working memory as an activated portion of long-term memory (e.g., Ericsson & Kintsch, 1995), Baddeley and Logie propose that such a view is probably an unhelpful oversimplification. They believe instead that working memory and long-term memory comprise two functionally separable systems. More specifically, a major role for working memory is retrieval of stored long-term knowledge relevant to the tasks at hand, the manipulation and recombination of material allowing the interpretation of novel stimuli, and the derivation of novel interpretations or the solution to problems. They further argue that working memory also plays a major role in encoding into long-term memory the outcome of its operations. Thus, Baddeley and Logie argue that an important feature of the central executive is to activate and integrate representations from long-term memory, and that working memory and long-term memory comprise two functionally separable cognitive systems.

In terms of the phonological loop, Baddeley and Logie (1999) stated that there is considerable evidence that long-term knowledge has an influence on the performance of verbal working memory tasks. For example, memory span for familiar words is longer than for nonsense syllables, and memory span for words in sentences is longer (15-16 words) than for unrelated words (5-6 words; Baddeley, Vallar, & Wilson, 1987). In addition, nonwords that approximate English result in better immediate recall by English speakers than those that do not (e.g., Gathercole & Baddeley, 1989), and familiarizing subjects with nonwords increases serial recall capacity (Hulme, Maughan, & Brown, 1991). Furthermore, the superior memory performance that often underpins expert performance is typically based on the temporary utilization of gradually acquired schemata (Ericsson & Delaney, 1999).

Baddeley and Logie (1999) suggest that Selfridge's (1959) pandemonium model, in which acoustic information is encoded and processed serially through levels, beginning with isolated speech sounds, moving up to phonemes, then syllables, which in turn may map onto sublexical and lexical units might explain the data. Where appropriate, they argue, lexical units may activate higher level syntactic and semantic structures, which in turn may be categorized in terms of semantic schemata, and encoding will stop depending on the capacity of the system to provide useful and meaningful chunks (Miller, 1956). [Marslen-Wilson (1987) and others, however, have presented evidence that phonological and lexical information interact in parallel in utterance interpretation.]

Taking the general working memory model of Baddeley and Logie (1999) as an overall framework, it is now possible to begin to examine more closely how specific components and subcomponents function. For example, some researchers have begun exploring the specific function of different components by showing that there are correlational dissociations between complex WMS measures such as reading span and spatial span (e.g., Daneman & Tardif, 1987; Shah & Miyake, 1996). In one study (Shah & Miyake, 1996), participants' reading WMS predicted performance on reading comprehension but not spatial thinking tasks, and participants' spatial span scores predicted performance on spatial thinking but not on reading comprehension tasks. Miyake and Shah propose that the question of greatest interest is:

> Are the sources of domain-specificity due to a) domain-specific subsystems [subcomponents] or storage buffers; b) similarity-based interference; c) domain-specific knowledge, skills, and strategies; and d) domain-specific processing speed?
>
> (p. 455)

The present research is focussed on a) the extent to which subcomponents used in spoken language processing interact, b) whether they do so differently for listeners with high and low WMS, c) the extent to which previously aquired knowledge interacts directly with incoming speech information, and d) whether the interaction is different for listeners with high and low WMS.

Assuming that working memory does consist of components which process and store different aspects of spoken language for comprehension, and, assuming that some of these components can interact with each other more than others, the present research was designed to address which subcomponents (e.g., auditory, phonological, lexical and sentence level) combine to influence how spoken language is processed. In addition, the extent to which they interact depending on the listener's WMS will be explored.

Specifically, the experiments were designed to examine whether temporal auditory processing interacts with processing within the phonological loop. Temporal auditory processing is defined as

> The ability [of the auditory system] to detect changes in stimuli over time, for example, to detect a brief gap between two stimuli or to detect that a sound is modulated in some way.
>
> (Moore, 1989, p. 137).

For example, high WMS listeners may detect subtle phonetic cues better than low WMS listeners due to superior temporal auditory processing such that they are more effective in using incoming acoustic speech to disambiguate lexical items. A second focus of the research concerns the interaction of lexical and sentence level knowledge (retrieved from long-term memory perhaps via the central executive) with phonological information processed and stored within the phonological loop, and the extent to which the interaction may vary for listeners who have different WMS. A further question examined in this research is whether differences in WMS may be related to differences in rehearsal capacities within the phonological loop. Perhaps high WMS listeners, for example, are better at holding phonological information in memory while processing other aspects of the incoming linguistic information. The final question that is addressed is whether listeners with high vs low WMS differ in their generation (i.e., activation) of alternatives from stored lexical information in long-term memory, and/or evaluation of the alternatives (e.g., inhibition of incorrect responses) while processing incoming linguistic information that is unclear, or ambiguous.

1.3.2 Discourse processing

Kintsch's construction-integration model has emerged from research on discourse comprehension (Kintsch, 1988; 1998; Kintsch & van Dijk, 1978). McNamara and Kintsch (1996) reported, for example, that high-knowledge adult readers performed better than low-knowledge readers on open-ended questions after reading low-coherence history texts. To my knowledge, however, Kintsch himself did not test whether differences in WMS contributed to differences in comprehension at the discourse level, but he provided the framework for others to do so (e.g. Spilich, 1983).

Kintsch's (1998) reading comprehension model describes the integration of text with previously-learned knowledge from long-term memory. Inferential reasoning is important in linking the two sources of information. Because information from the text is compared to related items activated in long-term memory, Ericsson and Kintsch (1995) thus view working memory as an activated portion of long-term memory. For the present research purposes, however, it will be assumed that working memory and long-term memory are two functionally separable cognitive systems, as argued by Baddeley and Logie (1999).

Examining why different WMS groups process language differently by taking the argument to finer levels of language processing such as phonological, lexical and sentence level processing proves more challenging and elusive because it has been difficult to isolate the influence of each of these factors from the influence of the others, and assess the degree to which one interacts with the others. Again, the present research is unique in that the relationship between WMS and spoken rather than written language comprehension is tested.

1.3.3 Models of individual and/or age-related differences in working memory and language comprehension

Within the past decade many researchers have been specifically investigating individual and/or population differences in WMS and language comprehension. The possible underlying causes of such differences have been proposed in different models. The major types of models and their proponents have recently been reviewed and include the following:

> 1) total amount of activation resources available to the system (Engle et al., 1992; Just & Carpenter, 1992), 2) processing speed (Salthouse, 1996; Wingfield,1996), 3) efficiency of inhibitory mechanisms (Stoltzfus et al., 1996), and 4) domain-specific knowledge and skills (Ericsson & Kintsch, 1995).
>
> (Shah & Miyake, 1999, p. 11)

Although the above theorists disagree on the underlying explanation for individual differences in the relationship between WMS and language comprehension, their explanations are not necessarily mutually exclusive. One aspect of working memory that they do all agree with Baddeley and Logie (1999) about is that working memory is not a unitary process (Shah & Miyake, 1999).

The present research was not designed to specifically compare the assumptions of the above models directly. For the purposes of considering whether individual differences in processing perceptually ambiguous words are due to differences in generation or evaluation of alternatives, the model of Just and Carpenter (1992) and the model of Stoltzfus et al. (1996), respectively, provide the most useful frameworks. A more detailed description of their models and their relevance to the present research will be discussed in section 1.5.e, below.

1.3.4 Modularity vs interaction among processes in language comprehension

Given that working memory is not unitary, there continues to be a great deal of interest in how the components required in language comprehension tasks inter-relate. Carpenter stated the problem and its importance for understanding how components of spoken language comprehension interact over the course of time:

> Because language comprehension is a multi-level [domain] process (containing component processes like lexical processing, syntactic analysis, and thematic analysis), it raises the question of how processing at the various levels is coordinated, temporally and functionally. At one extreme, a particular component process or subsystem of comprehension (e.g. syntactic analysis) could operate autonomously of the other processes. This position, called modularity, was clarified by Fodor's (1983) proposal that the mind might contain a set of cognitive modules, each of which is an autonomous subsystem, and in addition, an extremely complex central cognitive system. The research emphasis has been on the modules. Within a module, once a set of initiating conditions starts it on a minimal quantum of processing, it is uninfluenced by certain classes of information that may develop elsewhere in the system. This lack of influence is called informational encapsulation, and it has been proposed as a particularly important distinguishing property of a cognitive module (Garfield, 1989)....
>
> At the other extreme, all subsystems may be completely interactive, continuously taking into account any new relevant information as it is made available by other subsystems, and also making their own outputs available to other subsystems as soon as they are developed. Although these two positions represent extremes of cognitive architecture, they do not define a useful set of research strategies by themselves. Instead, the two extremes help generate a set of second-order research questions about the nature of encapsulation or interaction in various aspects of language processing. One such question concerns the <u>time course</u> [emphasis mine] over which syntactic, semantic, and pragmatic information interact in contributing to comprehension. Two subsystems whose degree of modularity/interaction have been examined are lexical access and syntactic analysis.
>
> (Carpenter, 1995, p. 94)

Lewis emphasizes the point that interaction between two cognitive components does not preclude the possibility that they perform specialized functions; interaction does preclude, however, the notion that components operate automatically, without appeal to information from other components:

> In sentence comprehension (online parsing and comprehension), the basic modularity issue is whether there is an autonomous syntactic parser that operates automatically without appeal to semantic or contextual knowledge sources (Garfield, 1989). Such a component would be a prime example of a cognitive module (or an input analyzer in Fodor's terms). The alternative view is that comprehension is an interactive process (Marslen-Wilson, 1975) in which multiple knowledge sources (including syntax) interact to produce the meaning. The critical issue is not whether knowledge sources are applied, but how. Modularity does not claim that semantics and context are unnecessary components of the comprehension process, but simply that syntactic processing occurs independently of the other knowledge sources. Similarly, an interactive account does not deny the importance of syntactic knowledge, but does not grant it distinguished status as a separate module.
>
> (Lewis, 1996, p.76)

It has been argued that, in sentence comprehension, syntactic processing interacts with other subsystems that utilize the same working memory resources, thereby requiring the allocation of processing resources to multiple subsystems. According to one resource model of working memory, the more syntactically complex a sentence is, the more difficult and time consuming it is to understand, especially for individuals with low WMS (Just & Carpenter, 1987, 1992; Miyake, Just, & Carpenter, 1994). Caplan and Waters (1999) have argued, however, that the verbal working memory system is composed of subsystems devoted to different types of verbal tasks that do not interact, and the assignment of sentence structure and its use in interpretation comprises such a specialized subsystem. They argue that people with problems in the syntactic subsystem have specific trouble comprehending syntactically complex sentences.

Caplan and Dufault (2000) recently tested a group of older adults with dementia of the Alzheimer's type, and a group of education-matched controls on a battery of WMS tests and an on-line test of sentence processing efficiency in which the sentences presented auditorily varied in complexity and syntactic acceptability. The amount of time listeners spent processing the sentence before they pressed a button for the next phrase was recorded and compared across groups. The dementia patients scored more poorly than the controls on all working memory tasks, but, listening times were similar for the dementia patients and the controls. Scores for the grammatical acceptability judgments were also similar between the two groups on the syntactically more complex sentences. Caplan and Dufault concluded that their study provides additional support for the thesis that syntactic processing utilizes different processing resources than those measured by standard tests of verbal working memory [although see Small, Kemper, & Lyons (2000) for evidence that sentence repetition by dementia patients is influenced by sentence complexity and their performance correlates with working memory scores].

Although the research presented here does not address syntactic processing per se, the questions that are addressed pertain to the same issues as those in which Caplan and Waters (1999) and Carpenter and her colleagues (1992; 1994) are interested. These issues are: 1) Which aspects of spoken language processing interact with each other? 2) To what extent do they interact with each other? and, 3) Does the pattern of interaction differ for groups that vary in working memory limitations?

These issues are particularly relevant in addressing the question about the interaction of auditory processing with other components in working memory. One possibility is that high and low WMS listeners differ in their temporal auditory processing, resulting in differences in operation of phonological loop and in turn in language processing. On the other hand, their temporal auditory processing may be similar, and differences may lie in cognitive processing of the speech input. Before considering the possible contribution of auditory processing it is useful to consider how cognitive models of speech processing have addressed lexical access and higher level processing.

1.3.5 Cognitive models of speech processing

Word recognition and cognitive models of speech processing. Spoken language comprehension involves perceptual, lexical and message-level (syntactic/semantic) analyses (Bernstein & Auer, 1996). The notion that the language processor is composed of modules, each of which is responsible for a strictly defined function and is only allowed access to strictly defined sorts of information (Fodor, 1983) has been of great interest to psycholinguists, as described above. Informational encapsulation, the proposal that each module has a clearly defined input and output, and that any other type of information simply has no access to the internal workings of that module is the aspect of modularity in which psycholinguists have been most interested.

Activation-based models known as the "Neighborhood Activation Model" (Luce, 1986), the "Cohort" model (Marslen-Wilson, 1987; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978) and the TRACE model (McClelland & Elman, 1986) all propose that word units have various levels of activation and compete for recognition as a function of similarity with incoming perceptual information (Bernstein & Auer, 1996). The above models have been developed to further our understanding of the interactional relationship among the perceptual, lexical and message (sentence/syntax) levels of processing. A more detailed review of the literature regarding the interaction between lexical and sentence level cues with phonological cues will be presented in sections 1.5.b and 1.5.c, below.

1.4 Ambiguity in spoken language

Many classic studies aimed at furthering our understanding of how spoken language is processed have examined the question by testing how individuals process ambiguous words or sentences. Specific examples of such research will be presented throughout this section, but first the various types of ambiguity in language that have been reported in the linguistics and psycholinguistics literature will be described.

According to Tartter (1986), ambiguity at the word level, or lexical ambiguity, can be defined as "a word with a double meaning" (p. 24), and ambiguity at the sentence level, or structural ambiguity, can be defined as "a sentence with different possible structural interpretations" (p. 109). The sentence *That is a hot car* illustrates lexical ambiguity. Several interpretations of the word *hot* are possible: The car may be hot in terms of temperature, it may be recently stolen, or, it may be "excellent" or "good". The sentence *They are visiting relatives* is an example of a sentence which is structurally ambiguous. This sentence has two possible interpretations based on syntactic structure: *They* are either relatives who have come, or people who have gone to visit their relatives, depending on the deep syntactic structure (Chomsky, 1957; 1964; 1965).

Slobin (1979) describes a third type of linguistic ambiguity called pragmatic ambiguity -- one in which disambiguation occurs as a result of inferential reasoning based on acquired cultural information and world knowledge stored in long-term memory. Correct interpretation of the sentence *I wanted to read the book at the library but it wasn't open yesterday.* (p. 12) relies on cultural, or contextual knowledge that it is the library, and not the book, that wasn't open yesterday. A fourth type of linguistic ambiguity is described by Crystal (1985) as semiotic ambiguity. Semiotic information, such as vocal intonation, facial expression and bodily gestures, is called upon in order to disambiguate the meaning of sentences in which structural meaning is contradicted by "the expression on his face..." or "...the way she looked when she said it" (p. 248).

A fifth type of linguistic ambiguity pertinent to the present study is perceptual ambiguity. Sometimes the acoustic speech signal is made unclear through degradation such as background noise, or poor articulation. A classic study by Miller and Nicely (1955) showed that tokens of phonemic categories presented in noise are confused most easily in terms of place of articulation and least easily in terms of voicing. For example, /p/-/t/ (voiceless stop consonants that differ in place of articulation) would easily be confused, whereas /b/-/p/ (bilabial stop consonants that differ in voicing) would not easily be confused in the presence of noise.

A sixth, and final, type of ambiguity combines perceptual and lexical ambiguity and is referred to as perceptual-lexical ambiguity (Connine, Blasko, & Wang, 1994). This type of ambiguity occurs when an ambiguous phonetic feature in a word creates the possibility of multiple lexical interpretations. The word cat, for example, could be interpreted as cap, if the final consonant is not clearly articulated or if background noise is present.

A classic study by Warren and Warren (1970) demonstrated how listeners "fill-in" ambiguous phonetic features in words based on higher order knowledge by showing that adults were able to restore a part of a word that had been replaced with white noise given the overall meaning of the sentence (e.g., The *eel was on the orange). Further, Salasoo and Pisoni (1985) have shown that when all but the initial 150 milliseconds of a word is replaced with a "noise envelope", adults fail to identify the word when it is presented in isolation, but identify it correctly when it is presented in a normal sentence context. Wingfield (1996) has shown, further, that young and older adult listeners correctly identify target words in a gating task better when the target word is preceded than when it is followed by sentence context.

In a more recent study by Wingfield, Lindfield, and Goodglass (2000), young and older adults were presented with 1) just word onsets, 2) word onsets followed by white noise indicating the full duration of the target word, or 3) word onsets followed by a low-pass-filtered signal that indicated the number of syllables and syllabic stress of the word in the absence of segmental information. Although older adults required longer stimulus durations for word recognition under all conditions, with age differences in hearing sensitivity contributing significantly to this age difference, word recognition was facilitated by knowledge of word prosody to the same degree for young and older adults. These findings suggest that listeners can detect and utilize word stress in making perceptual judgments about ambiguous input and that this ability remains spared in normal aging. Perceptual-lexical ambiguity is the type of ambiguity which listeners will be tested on in the research described presently.

As stated earlier, there have been two decades of debate in the literature over whether and how working memory is related to language (spoken and reading) comprehension. Much of the evidence used in this debate comes from testing comprehension of syntactically ambiguous sentences (Caplan & Waters, 1995; 1999), lexical ambiguity (Just & Carpenter, 1987; 1992), and pragmatic ambiguity (Hasher & Zacks, 1988; Kintsch, 1988; 1998; Kintsch & van Dijk, 1978). The present studies are novel in that they test the interpretation of spoken perceptual-lexical ambiguities.

1.5 The research plan

As stated previously, the purpose of this research is to examine why there are differences in spoken language processing between listeners with high and low WMS. The differences may be due to several factors including: differences in auditory perception of the speech signal, differences in processing the signal at the phonological, lexical or sentence levels, interaction between processing at different levels, rehearsal capacity, and/or, generation and/or evaluation of alternative interpretations to ambiguous signals. With regard to the matter of how processing at different levels might interact, the following questions are examined: Which aspects of spoken language processing interact with each other? To what extent do they interact with each other? and, Does the pattern of interaction differ for groups that vary in working memory limitations? Each of these issues will now be discussed in turn.

1.5.1 Auditory processing

The question has not yet been directly addressed as to how the perceptual (particularly temporal auditory processing) abilities of listeners interact with other components involved in spoken language comprehension. All of the models of language comprehension discussed in the previous sections have assumed that each individual receives the same quality and quantity of perceptual input. One of the purposes of the present research is to examine whether a perceptual processing component might perhaps be included in models of spoken language comprehension, and if so, how. It is possible that high WMS individuals perform better on text-based and spoken language comprehension tasks compared to low WMS individuals due partly to their superior visual or auditory perceptual capabilities (i.e., they may be better at extracting the signal from background noise).

Accuracy in spoken language comprehension requires auditory perceptual processing prior to and/or simultaneously with processing information from long-term memory and maintenance and storage of information. Auditory processing of language involves many stages of analysis by the peripheral and central auditory nervous systems (Greenberg, 1996; Willott, 1991).

Peripheral auditory processing consists of transduction and coding or analysis of incoming sound information by the mechanisms of the middle ear and cochlea. Central auditory processing consists of the relay and further coding and analysis of auditory information by neurons in the cochlear nucleus and superior olivary complex in the lower brain stem, the inferior colliculus in the midbrain, the medial geniculate body in the thalamus and the auditory regions of the neocortex (Pickles, 1988; Willott, 1991). It is well established that peripheral processing requires temporal processing of the auditory signal (Moore, 1989), and new evidence indicates that the onset responses of cells in the central auditory system, for example, also function to separate the signal from background noise temporally (Greenberg, 1996), a necessary function for extracting the linguistic signal from background noise, thus reducing potential ambiguity. There might be individial differences in temporal processing of the speech signal due to genetic predisposition, experience, or organic pathology in the relevent neuronal pathways in the peripheral and/or central auditory systems at anytime across the lifespan.

The present research was designed to investigate the role of temporal auditory processing in spoken language comprehension. (Listeners with normal pure-tone hearing were tested to rule out basic audibility differences as a possible explanation for group differences). Toward this end, listeners were tested on their ability to extract target words, with stop consonants differing in terms of articulatory feature contrasts (specifically voicing and place of articulation), from background noise (Experiments 1 to 3). The ability of the listeners to detect voiced from voiceless contrasts will depend on their ability to recognize differences in voice-onset time (VOT), which is assumed to be a task akin to detecting a gap in a pure tone. Both tasks are assumed to rely on timing analysis of the auditory input by cells in the peripheral and central auditory systems. In Experiment 4, listeners were given a classic gating task in quiet (e.g., Grosjean, 1980), in which the target word became less ambiguous as the gates increased in duration. Listeners' temporal auditory processing as measured by a gap-detection test (Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994), was compared to their performance on lexical identification performance in Experiments 2, 3 and 4. The gap-detection test consisted of presenting two stimuli to the listener, one of which is a continuous tone and the other of which contains a gap between the beginning and end of the tone. The duration of the gap for the test used in the Experiments described here ranged from 3 ms to 20 ms. The listener was asked to discriminate the two stimuli, and the point at which discrimination was no longer reliable was considered to be the listener's gap-detection threshold.

Auditory processing and spoken language comprehension. There is good reason to believe that individual differences in language comprehension may be at least partially a result of temporal auditory processing insofar as this is an important aspect of auditory processing, and both auditory processing and speed of processing have been related to language processing. Pichora-Fuller, Schneider, and Daneman (1995) have shown that older, normal-hearing listeners with age-related declines in auditory sensitivity benefit more from a supportive sentence context than young adult listeners when identifying spoken words in noise, and memory for words heard in noise is poorer for older than younger listeners (although memory for read words is similar for both groups). One possible explanation for these results is that due to poorer temporal auditory processing, the ability to extract a clear speech signal from the background noise is impaired, necessitating heavier reliance on sentence context, which is more cognitively demanding. Performance on both perception and memory tasks declines for both groups as background noise increases. Pichora-Fuller (1996) has also shown that bimodal perceptual cues (visual and auditory) prevent the decline in word-recall that occurs when unimodal auditory cues are presented in noise. Pichora-Fuller, Schneider and Daneman (1995) and Pichora-Fuller (1996) argue that their results support a resource model in which resources in working memory are reallocated from storage to perceptual processing during word identification when listening becomes effortful and word identification has priority.

Studies that directly compare the relationship between temporal psychoacoustic and speech perception performance are more equivocal. Tyler, Summerfield, Wood, and Fernandes (1982) found that psychoacoustic abilities (temporal difference limen and gap-detection scores) are poorer for hard-of-hearing listeners than for normal listeners, and that they correlate with ability to identify words in noise for both groups, but not with discrimination of VOT boundaries tested in quiet. On the other hand, Strouse, Ashmead, Ohde, & Grantham (1999) presented evidence that suggests that psychoacoustic (gap-detection thresholds and interaural time difference thresholds) and speech perception abilities (syllable identification/discrimination of phonemes varying in VOT and speech masking level difference) are poorer for older than young, normal-hearing listeners, but that the correlation between psychoacoustic and speech perception abilities is not significant for either group.

Two more recent studies have found a correlation between gap-detection scores and performance on word identification tasks in which the availability of temporal information in the speech signal was varied in speech rate for both young and older adult listeners. The discrimination of word pairs contrasting in terms of presence or absence of an inserted silent interval (e.g., *cash-catch*, *slit-split*) presented in quiet as well as three noise conditions, was more highly correlated with gap-detection scores when the words were presented at a fast compared to a slow speaking rate (Haubert & Pichora-Fuller, submitted). Lam (in preparation) found that sentence repetition performance correlated significantly with gap-detection scores for speech presented at a speeded (time-compressed) rate, but not for speech presented at a normal rate. These studies suggest that temporal auditory processing may influence the comprehension of incoming speech at the lexical and sentence levels, even for young listeners with normal audiograms, at least for fast speech.

Recall that Baddeley and Logie's (1999) model, as well as the others discussed above, all assume that perceptual input is processed similarly by all individuals, regardless of modality. The present research is novel in that it addresses this assumption directly. If differences in temporal auditory processing are found between high and low WMS listeners, and if the temporal auditory processing scores correlate with comprehension performance, it will suggest that perceptual processing may interact with other components of working memory, thus influencing group differences in language comprehension. If non-speech temporal auditory processing differences are not found to differ between high and low WMS listeners, or if they are found but do not correlate with spoken language comprehension, it could be argued that auditory-level perceptual processing operates as a separate component from other aspects of working memory during spoken language comprehension, at least for this population.

Thus, if high WMS listeners have better temporal auditory processing, it is predicted that they will generate fewer alternative responses to an ambiguous target word by honing in on the correct response more quickly than low WMS listeners. If, on the other hand, temporal auditory processing is not significantly different between high and low WMS listeners, any differences between low and high WMS groups in generating alternatives could be attributed to other causes, such as cognitive processing during lexical access.

According to a cognitive model of processing, it is predicted that high WMS listeners may generate more alternative responses to an ambiguous target word, thereby providing them with more alternatives from which to select their final response. They may, therefore, respond more slowly because they consider more alternatives.

1.5.2 Interaction between phonological and lexical level cues

The present research was designed, in part, to examine whether phonological, lexical and sentence level components interact during comprehension in the same manner for high and low WMS listeners. It has been well documented that word-frequency (lexical level information) interacts strongly with phonological level processing, whereas sentence level context interacts less (Altmann, 1990). The next two sections are devoted to reviewing the literature on the influence of lexical and sentence level knowledge on the interpretation of perceptual-lexical ambiguities.

The effects of lexical status on phonemic perception are well documented (e.g., Connine & Clifton, 1987; Ganong 1980; Samuel, 1981, 1990). Ganong (1980) has shown, for example, that when presented with a word-nonword speech continuum such as *dash-tash*, listeners identified the stimulus-initial ambiguous phoneme such that a word was formed, suggesting that ambiguous information at word onset is admitted access to the lexicon.

Andruski, Blumstein, & Burton (1994) have also shown that there is interaction between the phonological and lexical levels of analysis. They provided evidence that words which are phonologically similar to the intended word candidate are also activated to some extent during lexical access, whether the input provides a good phonetic representation of the intended word or a poorer one. Based on these results, Andruski et al. argue that the presence or absence of a real-word counterpart contributes to the density or size of the set of activated lexical candidates.

According to Altman (1990, p.4), early versions of the cohort model of spoken-word recognition (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978) proposed that the acoustic input available at a particular point in time activates the set of all lexical candidates compatible with that input, with subsequent disconfirming evidence causing members of the cohort to drop out. One reason that support for the original model declined was the assumption that the matching process between acoustic input and each member of the cohort was an all-or-none process --- if the input matched, the cohort member was activated, but if there was a mismatch, it was inactivated. The revised cohort model (Marslen-Wilson, 1987) overcame this problem by proposing that the cohort consists of elements whose activation levels are determined by goodness of fit to the acoustic input with higher-level processes evaluating the most highly activated elements, in parallel, for integration into the utterance interpretation, rather than operating on an all-or-none basis. Evidence counter to the original model was also provided by increasing evidence that word frequency can affect the activation levels of lexical candidates in the initial stages of the access process (Marslen-Wilson, 1990; Zwitserlood, 1989). In addition, Luce's Neighborhood Activation model (1986) proposes that the number of cohort members that are activated has an effect on word recognition, that is, the more cohort members that are activated, the longer the decision process takes. Marslen-Wilson's (1987) data, however, suggest otherwise, supporting a parallel processing model.

Whether lexical level knowledge influences the perception of prelexical information in a top-down fashion is a matter of some debate (Norris, McQueen, & Cutler, 1998, as cited in Cutler & Clifton, Jr., 1999). Elman and McClelland (1986) provided evidence that listeners who heard an ambiguous token between /s/ and /ʃ/ at the end of the words *Christma** and *fooli** shifted their responses in the same direction as the actual phonemes at the end of *Christmas* and *foolish*. The results were also simulated in TRACE by Elman and McClelland, who attributed the top-down influence of the lexical knowledge to the feedback connections in their connectionist model of word-recognition. According to the model, which is based on the principles of interactive activation, information processing takes place through excitatory and inhibitory interactions of a large number of simple processing units. Each unit constantly updates its own activation based on the activations units to which it is connected. The model is called TRACE because a dynamic processing structure called "the Trace," is formed by the network of units. "The Trace" serves simultaneously as the perceptual processing mechanism and as the system's working memory. The model simulates a large number of empirical findings concerning the interactions of phoneme and word perception. Norris (1993), however, also simulated the same findings with another connectionist model -- but one with no feedback connections, and argued that top-down connections are not necessary to achieve the same results.

Norris et al. (1998) further argue that if lexical knowledge did have an influence on phonological processing, such an interaction would result in poorer performance, i.e., more errors. (Of course, this would depend on how extensive the listeners' lexical knowledge-base is, and whether the lexical items activated influence the ambiguous pre-lexical information in the direction toward the "correct" response.) The relevance of this issue to the present research is that when presented with perceptual-lexical ambiguities, the low WMS listeners may show greater interaction between lexical and/or sentence level knowledge and phonological processing, resulting in more errors when the higher level sources do not support the target word, whereas the high WMS listeners may not be so influenced by higher levels of information. The high WMS listeners might, on the other hand, be more influenced by higher order sources insofar as they generate more alternatives to the ambiguous speech signal, but process them more efficiently, resulting in fewer errors, whether or not the higher order information supports the target.

1.5.3 Interaction between phonological and sentence level cues

Some evidence has been presented to suggest that sentence-level processing does not interact strongly with early perceptual encoding (Connine, 1990; Samuel, 1981) and that sentence-level processing is post-perceptual and operates similarly to payoff bias (i.e., decision-making that is expressed in terms of the costs and/or benefits associated with taking a particular course of action; Connine & Clifton, 1987). As stated earlier, there appears to be agreement that lexical information is capable of altering the percept of the individual sound segments that make up a word, but sentential information usually has only a much more indirect effect on lower-level processing (Altmann, 1990, p. 18). The present research was designed to test the influence of sentence level cues on phoneme perception when the speech signal is degraded, making it ambiguous.

In the proposed experiments it is assumed that cohort effects will occur, i.e., multiple interpretations of perceptual-lexical ambiguities will be activated to varying degrees (i.e., the word *pat*, when presented in noise would perhaps activate competitors such as *bat*, *mat*, *hat*, *sat*, etc.), and that the alternatives generated will increase with an increase in background noise. Word frequency in Experiments 1 and 4, and sentence context in Experiments 2 and 3 were manipulated to determine the extent to which they are employed by the listener in resolving the perceptual-lexical ambiguity. Whether lexical or sentence level knowledge interacts with phonological processing of incoming speech information differentially for high and low WMS listeners in challenging perceptual conditions is of primary interest.

Influences of sentence context on perceptual-lexical ambiguities

Sentence level knowledge has been shown to influence high WMS listeners in their interpretation of words in which ambiguous phonetic features make more than one interpretation possible (albeit less strongly than lexical level knowledge). In an experiment conducted by Connine, Blasko, and Wang (1994), high and low WMS listeners were presented with sentences in which the contextual information at the sentence level was followed by a perceptual-lexical ambiguity or an unambiguous control word in counterbalanced order (e.g., *Let's climb to the* *dip/tip* vs *Let's climb to the roof*). The ambiguous words (e.g. *dip/tip*) were created by removing portions of the voiced segment of the voiced consonants, and replacing them with equally long voiceless segments of the voiceless consonants. In this manner, a continuum of voiced-ambiguous-voiceless intitial stop consonants for each pair of words was created. After each sentence was presented, it was followed by a target word presented visually (e.g., *swim*) either immediately or after an 850-ms delay, at which point the listener was required to indicate as quickly and accurately as possible whether the visual target was related to the sentence. It was found that, although the accuracy of the responses was equally high for the two groups, the high WMS listeners showed more effective multiple activation of the perceptual-lexical ambiguities than did low WMS listeners. This was indicated by greater differences in response latencies to the ambiguous vs unambiguous controls for the high WMS group. Connine et al. suggested that the high WMS listeners took longer to compare alternative hypotheses because more alternatives were activated, whereas low WMS listeners did not require as much time because fewer alternative hypotheses were activated. Moreover, when an 850-ms delay was introduced after the sentences were heard, more effective use of context to resolve the ambiguity was shown for high WMS listeners (the differences in response latencies to the ambiguous vs unambiguous controls decreased significantly). Those low WMS listeners who <u>could</u> effectively activate and maintain multiple lexical hypotheses, were less effective in using context to resolve the ambiguity than the high WMS listeners.

Results from a study by Gernsbacher and Faust (1991) in which low comprehenders showed similar perseveration of inappropriate homonym meanings are consistent with Connine et al.'s (1994) results. Connine et al. argue that this convergent evidence demonstrates a parallel between processing homonyms and perceptual-lexical ambiguities and implicates similar processes. Connine and her colleagues (1994) did not test <u>directly</u>, however, whether the differences they found were due to <u>generation</u> of alternatives to the ambiguous word and/or <u>evaluative</u> processing of the alternatives generated, as Experiment 4 in the present research was designed to do.

More evidence that sentence context influences language processing under conditions of uncertainty come from the literature on reading. Just and Carpenter (1987) cite studies which have shown that sentence context might contribute most to comprehension under three conditions: 1) when the text is of poor physical quality, 2) when the reading conditions (e.g., illumination) are poor, or 3) when the reader's encoding skill is poor.

Furthermore, good readers are found to be better than poor readers at word recognition, when words are presented both in and out of context (Stanovich, 1980; Stanovich & West, 1981). Moreover, context helps poor readers recognize a word more than it helps good readers. Good and poor readers from fourth grade were tested on their word-encoding and lexical access processes, with words that were printed in: 1) normal print, 2) moderately degraded print, and 3) highly degraded print (Perfetti & Roth, 1981; Roth, Perfetti, & Lesgold, 1979). The time readers took to pronounce words was then measured and compared. Good readers were faster at pronouncing words than were poor readers, and good readers and poor readers were faster when the print was normal than when it was degraded. To determine how much context helped word recognition, the experimenters then measured how long readers took to pronounce the words in the presence of context. It was found that the poor readers were helped more by the context, but only because they typically take longer than good readers to encode and access words. The authors concluded that the contextual influences on the speed of word recognition are not a primary source of individual differences, but that context effects are due to differences in the speed of word encoding and lexical access.

1.5.4 Rehearsal capacity

As discussed in section 1.3.a above, Baddeley and Logie (1999) argue that individual differences in phonological loop capacity may reflect the amount of memory activation and/or rehearsal capacity available to the listener. In Experiments 2 and 3, described below, one of the hypotheses tested concerns whether high and low WMS listeners differ in their ability to generate and hold multiple interpretations of degraded speech signals and to rehearse them while considering the sentence context.

1.5.5 Evaluation/generation

The present research is designed, in part, to also address whether the explanations put forth by Carpenter et al. (1995), or Stoltzfus et al. (1996) can be upheld. That is: Are individual differences in spoken language comprehension due to differences in generation of alternatives or to differences in inhibition of incorrect alternatives?

In Just and Carpenter's (1987) new implementation of their Reader model of written language comprehension, working memory is described as consisting of two components: processing and storage. The processing component is responsible for analyzing and interpreting linguistic input, and the storage component is responsible for temporarily holding information in memory, to be later integrated with further incoming information. Carpenter and colleagues, in their Capacity-constrained (CC) Reader model (for a review see Carpenter et al., 1995), implemented working memory as capacity-limited (Ebbinghaus, 1885/1913; Miller, 1956), such that as the allocation of resources to one component increases, fewer resources remain for the other. [Recall that Baddeley and Logie (1999) also argued for a distinction between processing and storage in working memory; however, this distinction occurs in the components rather than the central executive with capacity-limitations occurring in both the components and the central executive.]

Carpenter et al.'s (1995) CC Reader model of language comprehension is a hybrid model. In addition to a connectionist component, CC Reader employs a production-based, directed-activation component which takes previously processed information into account in order to make decisions about how to process further incoming linguistic information. This component operates independently in a self-activated, feed-forward manner within the connectionist framework.

According to the CC Reader model, upon encountering lexical or syntactic ambiguity, the processor first generates multiple representations based on lexical, syntactic and/or pragmatic contextual information. Each source of information activates the processor to varying degrees, depending on the strength of the source. Working memory supports the following computational or processing components: syntactic parsing, thematic role assignment, comparison, integration of information, inference, referential assignment and logical operations. During activation, each element (word, phrase, grammatic or thematic structure) has an associated activation threshold level. Productions can fire reiteratively over successive cycles so that activation levels of the output elements become gradually incremented until they reach their activation threshold. Reintegration may sometimes be necessary to resolve ambiguity, as shown experimentally by longer response times. Because language is represented temporally, it is important that the storage component of working memory maintains the intermediate and final results of these computations.

Carpenter and her colleagues (1995) further propose that errors in language comprehension occur as a result of working memory constraints that are imposed on the system. In the case where working memory capacity is reached, reallocation of resources occurs. As stated earlier, Carpenter et al. have proposed that working memory involves two components: storage of information and processing. When the system is stressed, as may occur when a person with a low WMS encounters an ambiguous or complex sentence, either storage of previously processed information, or processing of incoming information is compromised. Deallocation of resources from the storage function of working memory results in forgetting; deallocation of resources from the processing function of working memory results in longer processing time or errors. The model predicts that comprehension speed and error rates will increase as a result of intrinsic memory load (e.g., syntactic complexity or lexical ambiguity), text distance (e.g., distance between topical agent and patient in continuous discourse), extrinsic memory load (e.g., holding related items in memory) and temporal restrictions induced by rapid presentation rates.

Just and Carpenter (1992) have provided data from human readers that illustrate deallocation from storage to the processing component of working memory. In an on-line reading task, human readers with either high or low reading WMS were asked to read sentences either low or high in complexity. It was found that the high and low WMS readers performed equivalently when presented with a sentence such as the subject-relative sentence *The reporter that attacked the senator admitted the error*. Readers with higher WMS, however, took significantly less time to read sentences such as the more complex object-relative sentence *The reporter that the senator attacked admitted the error* than did readers with low WMS. Other researchers, however, interpret similar difficulties aphasics encounter in interpreting syntactically complex sentences to be due to a deficiency in the ability to access certain syntactic components (e.g., Berndt & Caramazza, 1980; Caplan, Baker, & Dehaut, 1985; Caplan & Waters, 1995). Similarly, an alternative hypothesis for Carpenter et al.'s results may be that low WMS listeners may simply be poorer sentence processors.

Just and Carpenter (1992) concluded that complex sentences require more working memory resources than do simple sentences. Further evidence in support of their conclusion was provided by the computer simulation, CC Reader. The working memory component (activation limit) of the CC Reader simulator can be increased or decreased to simulate individuals with high or low WMS. When highly complex sentences were presented to the CC Reader simulator, a higher activation limit was required to simulate the human data for a high WMS reader and a lower activation limit was requiredto simulate the data for a low WMS reader.

Further evidence has been provided by the CC Reader model in another on-line task which demonstrates that high WMS readers are better able than low WMS readers to hold multiple meanings of an ambiguous word in memory until the word is disambiguated by further sentence processing (Miyake et al., 1994). Miyake et al. presented readers with sentences in which the <u>subordinate</u> interpretation of a homograph was correct. An example is: *Since Ken really liked the <u>boxer</u>, he took a bus to the nearest pet store to buy the animal.* (The matching sentence in which the <u>dominant</u> interpretation of the homograph was correct is *Since Ken really liked the <u>boxer</u>, he took a bus to the nearest sports arena to see the match.*) The effect of an increase in reading time by the low WMS readers due to the lexical ambiguity in the former sentence was first observed soon after the occurrence of the first disambiguating word, e.g., *pet*. The low WMS readers showed the largest increase in reading time, though, on the last word of the sentence, but only when they read sentences in which the subordinate rather than dominant meaning for the target word was indicated. The authors suggest that these readers may have been mentally back-tracking to find an alternative (subordinate) meaning for the ambiguous word, whereas the high WMS readers only had to chose from the alternatives that were still activated above threshold. Relevant to the present research is the assumption that the high WMS readers may have been better at activating, or generating multiple alternatives to the ambiguous homograph, and that they have slower decay.

Carpenter et al.'s (1995) CC Reader model is somewhat consistent with Hasher and Zack's (1988) General Capacity Theory of Working Memory (subsequently updated by Hasher, Stoltzfus, Zacks, & Rypma, 1991; and, Stoltzfus et al., 1996) which states that cognitive performance depends upon the extent to which the contents of working memory (that is, the currently activated information) reflect the subject's current task goals. Rather than activation limitations, according to their model, individual differences in language comprehension are due to deficient inhibitory control over the contents of working memory. These limitations lead to the system being preoccupied by task-irrelevant information, with consequent difficulties in accessing and retrieving task-relevant information (Stoltzfus et al., 1996). Stoltzfus et al. have argued that this account appears to explain the pattern of cognitive deficits in older adults, that they have less inhibition and slower decay of irrelevant items, and it may well be relevant to the study of individual differences in performance more generally.

1.5.6 The experiments and hypotheses

Four experiments were designed to examine why there are individual differences in spoken language processing between listeners with high and low WMS. The differences may be due to several factors including: differences in temporal auditory processing of the incoming acoustic signal, influences at the phonological, lexical and/or sentence levels, differences in rehearsal capacity, and/or, differences in generating and/or evaluating alternative interpretations of ambiguous words.

A reading working memory span test[1] developed by Daneman and Carpenter (1980) was administered once to listeners in Experiment 1 and twice (for reliability) to listeners in Experiments 2 to 4 to determine their reading working memory spans (WMS). Based on the results, each subject was then assigned to the high or low WMS group.

In the first three experiments, target words which varied in a minimal pair contrast were presented in a variety of spoken sentences to young, normal-hearing adult listeners, who had either high or low WMS. The sentences were presented in varying degrees of background noise, (and in quiet for Experiments 2 and 3). As the background noise increased, the acoustic signal became increasingly ambiguous. The listeners were asked, in each experiment, to identify the ambiguous target word. The sentences varied in degree and type of supportive information. Word frequency was varied in the first experiment and the degree to which sentence context supported one interpretation or another of the target word was varied in second and third experiments. The target words, whose frequency was controlled in Experiments 2 and 3, were embedded in three levels of sentence context: Context-Consistent, Context-Inconsistent, and Context-Neutral (see Appendix A). All sentences were of the moderately complex co-ordinated syntactic type; in this sentence type the target word precedes the supportive context of the sentence, requiring the listener to hold the ambiguous target word in working memory while processing the meaning of the sentence (see Wingfield, 1996).

The listeners heard the sentences through headphones, and, in Experiment 1 they were asked to indicate what they heard by drawing an arrow from the subject(s) to the object(s) of the sentence and writing the verb(s) above the arrow(s) on a response form (see Appendix B). In Experiments 2 and 3, the listeners were then asked to indicate their interpretation of the sentence by

[1] Reading and listening WM tests are significantly correlated (for a review see Daneman & Merikle, 1996), at least when sensory systems are normal and presentation conditions are ideal. To the extent that language processing is modality-independent, then either reading or listening span tests should tap the same individual differences in language processing ability related to WM. In cases where sensory impairment or non-ideal conditions of presentation are thought to alter language processing in a more modality-dependent fashion, and the experimenter is interested in understanding how those conditions impact WM during language processing, then it could be argued that WMS should be tested in the modality and conditions of interest in order to appraise altered processing. Alternatively, it could be argued that "true" language processing, uncontaminated by perceptual stress, is more legitimately measured in the unaffected modality or in ideal conditions.

clicking on the corresponding box on a computer screen (see Appendix C). In Experiment 2, in order to ensure that listeners were processing the sentence in a deep, meaningful fashion, thereby taxing working memory, the responses on the computer screen consisted of the target word and a synonym for the final verb in each sentence [see Craik & Lockhart (1972) for the distinction between deep and shallow processing]. Experiment 3 was identical to Experiment 2, except the responses on the computer screen consisted of the target word and the actual final verb of each sentence (rather than a synonym for the final verb; see Appendix D). Thus the task required only shallow processing. The results for the high and low WMS listeners in each experiment were then compared to assess the degree to which each group made use of the phonological, lexical and sentence level information in identifying the target word.

In the fourth experiment, a classic gating task (e.g., Grosjean, 1980; Marslen-Wilson, 1990), was implemented in which high-frequency words with low-frequency competitors and low-frequency words with high-frequency competitors were each presented in 50 ms increments to high and low WMS listeners. (Background noise was not required in the fourth experiment, as ambiguity was introduced by word-incompleteness.) The fourth experiment was designed to assess whether listeners <u>generate</u> multiple alternatives to ambiguous words differentially or <u>evaluate</u> alternatives differentially (e.g., through inhibitory mechanisms).

In Experiments 2 to 4, temporal auditory processing (via gap-detection scores) was compared to comprehension performance in order to evaluate whether temporal auditory processing interacts with components of working memory in the processing of ambiguous spoken language.

<u>Experiment 1</u>

Experiment 1 was designed to examine whether phonological or word-frequency cues are used differently by young, normal-hearing adult listeners with low vs high WMS.

Hypothesis 1 (Experiments 1-4):

H_r: A significant WMS x Phonological Cue interaction will be found.

Prediction: It is predicted that high WMS listeners will discriminate phonological cues significantly better than low WMS listeners, especially voicing cues.

Rationale: Low WMS readers have more difficulty than high WMS readers comprehending language. Because comprehension relies on word identification, which, in turn, is influenced partly by phonological cues, it may be that individuals differing in WMS may differ at the level of phonological encoding.

H_o: A significant WMS x Phonological Cue interaction will not be found. This will indicate that phonological processing does not interact with components of working memory that influence comprehension.

Hypothesis 2 (Experiment 1):

H_r: A significant WMS x Word-frequency interaction will be found.

Prediction: It is predicted that low WMS span listeners will make more Target-Word identification errors based on word-frequency cues than high WMS listeners.

Rationale: Low WMS readers have more difficulty than high WMS readers comprehending language. Because comprehension relies on word identification, which, in turn, is influenced partly by word frequency, it may be that individuals differing in WMS may also show differences in the degree to which they are influenced by word frequency during word identification.

H_o: A significant WMS x Word-frequency interaction will be not found. This will suggest that word frequency, an aspect of lexical knowledge, influences word-recognition equivalently for high and low WMS listeners.

Experiments 2 and 3

Experiments 2 and 3 were designed to test whether high and low WMS listeners process ambiguous target words differently based on sentence level context. A second question these two experiments were designed to examine was the extent to which perceptual vs cognitive factors explains individual differences in spoken language processing. A third question of interest was whether the pattern of results provides evidence for differences between high and low WMS listeners in generation or evaluation of alternatives. A final question was whether the deep processing task (Experiment 2) requires more general working memory resources than the shallow task (Experiment 3).

Hypothesis 3 (Experiments 2 and 3):

H_r: If differences occur in the interaction of sentence level knowledge with phonological processing between high and low WMS listeners, then a Sentence Context x WMS interaction for Correct Responses will be found.

Prediction: It is predicted that low WMS listeners will show more errors than high WMS listeners on Context-Inconsistent than Context-Consistent sentences in high noise.

Rationale: Stanovich and West (1979) showed that young adults relied more heavily on top-down information (context) under conditions of stimulus degradation compared to non-degraded conditions. If low WMS listeners rely more on sentence context, due to poorer perceptual encoding, by reallocating working memory resources from storage to processing, as Carpenter et al. (1995) would suggest, this group should show more errors on Context-Inconsistent than Context-Consistent sentences in high noise.

H_0: If differences in the interaction of sentence level knowledge with phonological processing between high and low WMS listeners do not occur, then a Sentence Context x WMS interaction for Correct Responses will not be found.

Hypothesis 4A (Experiments 2 and 3) -- Perceptual Argument:

H_r: A positive correlation between Correct Responses and gap-detection performance and a Correct Responses x WMS interaction will be found.

Prediction: It is predicted that high WMS listeners will demonstrate better gap-detection scores and significantly more Correct Responses than low WMS listeners, and a positive correlation between Number of Correct Responses and gap-detection scores will be found for both WMS groups.

Rationale: Reallocation of resources from storage to processing may be taking place for low WMS listeners due to poorer perceptual encoding (Pichora-Fuller et al., 1995). High WMS listeners may deploy superior temporal auditory processing in honing in on the correct interpretation of perceptual-lexical ambiguities. Such results would support the argument that temporal auditory processing is an interactive component in working memory for language.

H_o: A positive correlation between temporal auditory processing and comprehension performance will not be found, but a Correct Responses x WMS interaction will be found.

Hypothesis 4B (Experiments 2 and 3) -- Cognitive Argument:

H_r: A positive correlation between temporal auditory processing and comprehension performance will not be found, but a Correct Responses x WMS interaction will be found.

Prediction: It is predicted that high and low WMS listeners will demonstrate similar gap-detection scores. Significantly more Correct Responses are predicted for high than low WMS listeners, but a positive correlation between Number of Correct Responses and gap-detection scores will not be found for both WMS groups.

Rationale: Baddeley and Logie (1999) argue that individual differences in spoken language processing may be dependent on activation and rehearsal capacity within the phonological loop. This would indicate that comprehension errors may be due to cognitive differences. Therefore, any differences between low and high WMS groups in language processing could be attributed to the high WMS group's superior ability to generate and rehearse multiple interpretations of the signal while considering the sentence context. Such results would support a cognitive rather than perceptual explanation for group differences.

$\underline{H_o}$: A positive correlation between Correct Responses and gap-detection performance and a Correct Responses x WMS interaction will be found (perceptual argument).

Subsidiary hypotheses for Experiments 2 and 3 regarding time course of activation in noise

Hypothesis 4C (Experiments 2 and 3) — Perceptual Argument:

$\underline{H_r}$: Significant Correct Response x WMS, Noise x WMS, and Response Time x WMS interactions will be found. A significant positive correlation between Correct Response and gap-detection scores for all listeners and between gap-detection scores and Correct Response for high and low WMS listeners analyzed separately will be found

Prediction: It is predicted that high WMS span listeners will demonstrate better gap-detection scores than low WMS listeners, make fewer comprehension errors than low WMS listeners, especially in high noise, have faster response times than low WMS listeners, expecially in high noise, and that comprehension scores will positively correlate with gap-detection scores.

Rationale: Reallocation of resources from storage to processing may be taking place for low WMS listeners due to poorer perceptual encoding (Pichora-Fuller et al., 1995). High WMS listeners may deploy superior temporal auditory processing in honing in on the correct interpretation of perceptual-lexical ambiguities. Furthermore, if high WMS listeners generate fewer alternatives in high noise compared to low WMS listeners as a result of their more finely-tuned temporal auditory processing, they would be able to respond faster. A perceptual explanation for differences in comprehension between WMS groups would be supported.

$\underline{H_o}$: Significant Noise x WMS, Correct Response x WMS, and Response Time x WMS interactions will be found, but a significant correlation between Correct Response and gap-detection scores for all listeners, and between gap-detection scores and Correct Response for high and low WMS listeners analyzed separately will not be found.

Hypothesis 4D (Experiments 2 and 3) -- Cognitive Argument:

H_r: Correct Response x WMS, Noise x WMS, and Response Time x WMS interactions will be found, but a significant correlation between Correct Response and gap-detection scores for all listeners will not be found.

Prediction: It is predicted that high WMS span listeners will make more comprehension errors than low WMS listeners in high noise compared to quiet, have slower response times than low WMS listeners in high noise compared to quiet, and their comprehension scores will not correlate with their gap-detection scores.

Rationale: Baddeley and Logie (1999) argue that individual differences in spoken language processing may be dependent on activation and rehearsal capacity within the phonological loop. Furthermore, high WMS listeners may perceive an equally clear signal as the low WMS listeners, but generate and rehearse multiple interpretations of the signal, thereby taking more time to consider the alternatives. Thus, superior activation and/or rehearsal capacities within the phonological loop may underlie differences in comprehension, supporting a cognitive rather than perceptual explanation for group differences.

H_o: Significant Noise x WMS, Correct Response x WMS, and Response Time x WMS interactions will be found, and a significant positive correlation between Correct Response and gap-detection scores for all listeners will be found (perceptual argument).

Hypothesis 5 (Experiments 2 and 3):

Recall that Experiment 2 was designed to tax general working memory resources with a deep processing task, whereas the shallow processing task used in Experiment 3 was designed not to tax general memory resources as much.

H_r: Interactions between WMS and Sentence Context for Correct Responses will be greater for Experiment 2 than 3.

Prediction: High WMS listeners are expected to have significantly more Correct Responses than low WMS listeners in Experiment 2, but less difference is expected for Experiment 3.

Rationale: Carpenter et al.'s (1995) resource model states that processing and storage share a common pool of resources. If the deep processing task requires more general working memory resources than the shallow task, it should be more difficult a task for the low compared to high WMS listeners.

H_o: Interactions between WMS and Sentence Context for Correct Responses will be equal for Experiments 2 and 3.

Experiment 4

Experiment 4 was designed to investigate more directly the influence of lexical level information on whether high WMS listeners generate more alternatives to perceptual-lexical ambiguities or evaluate the alternatives (through inhibition of the incorrect response) more successfully.

Hypothesis 6 (Experiment 4):

H_r: High and low WMS listeners will differ in the number of low-frequency competitors they produce to high-frequency target words and in the number of high-frequency competitors they produce to low-frequency target words.

Prediction: High WMS listeners will produce more low-frequency competitors to high-frequency target words and more high-frequency competitors to low-frequency target words than the low WMS listeners.

Rationale: Carpenter et al.'s (1995) CC Reader model states that poorer language processing of ambiguous sentences by individuals with low WMS is explained by lower activation levels, and, similarly, Baddeley and Logie's (1999) working memory model states that differences in WMS may be accounted for by differences in activation within the phonological loop.

Therefore, differences in comprehension between high and low WMS groups may depend on the number of competitors activated to ambiguous target words.

$\underline{H_o}$: High and low WMS listeners will not differ in the number of low-frequency competitors they produce to high-frequency target words and the number of high-frequency competitors to low-frequency target words..

Hypothesis 7 (Experiment 4):

$\underline{H_r}$: A WMS x Response Time interaction will be found.

Prediction: High WMS listeners will generate the same number of competitors to high and low target words as the low WMS listeners, but identify the target words sooner in the gating task.

Rationale: Stoltzfus et al. (1996) argue that high and low WMS listeners generate the same number of alternatives to ambiguous target words, but low WMS listeners may lack inhibitory mechanisms during evaluation of the alternatives. Thus, if high and low WMS listeners generate the same number of alternatives to target words, but inhibit incorrect target words differently, then differences will be shown in RTs rather than Correct Responses in Experiment 4.

$\underline{H_o}$: A WMS x Response Time interaction will not be found.

II EXPERIMENT 1

2.1 Interaction of phonological and lexical processing

Experiment 1 was designed to examine the extent to which phonological and word-frequency cues interact with WMS in young, normal-hearing adult listeners. A second question was whether reallocation of working memory resources from storage to processing occurs more for low WMS listeners as Carpenter et al.'s (1995) model would predict.

In order to test the above questions, 12 high and 12 low WMS listeners between 18 and 35 years of age, and with normal pure-tone hearing, were tested on their ability to recognize target words in spoken sentences. Three sets of target words which varied in word frequency (high, medium, or low) and phonological cues [their final consonants varied in terms of VOT (e.g., voiced or voiceless)], and place of articulation (e.g., bilabial, alveolar, or velar) were embedded in sentences which were of low, medium or high syntactic complexity (Caplan, Baker, & Dehaut 1985).

Because the sentences were presented in three levels of background noise, the target words were rendered perceptual-lexically ambiguous. (The ambiguity of the signal increases as background noise increases.) Perceptual-lexical ambiguity occurs when an ambiguous phonetic feature of a word creates the possibility of multiple lexical interpretations. (The word <u>cat</u>, for example, could be interpreted as <u>cap</u>, if the final consonant is not clearly articulated or background noise is present.)

The listeners' task was to listen to each sentence through headphones and indicate what they heard by drawing an arrow from the agent(s) of the sentence to the patient(s) and writing the verb(s) above the arrow(s) on a response form [hereafter referred to as the "Pen and Paper" task; Lloyd & Pichora-Fuller, 1996; see Appendix B]. An on-line memory task was given during testing to provide an additional measure of WMS and to ensure that the storage component of working memory was stressed.

2.2 Method

2.2.1 Participants

Twenty-four undergraduate students between the ages of 19.5 and 35.5 years (mean age 25.4 years; standard deviation 4.5 years), who gave their informed consent, participated in this study. Participants were recruited through newspaper and poster advertisements distributed in and around the University of British Columbia. All participants were native English speakers, and gender was represented approximately equally in each group. Participants received payment for their participation.

Participants were assigned to one of two groups depending on their reading WMS score (Daneman & Carpenter, 1980) which was measured at the beginning of the testing session.

During the reading working memory test, participants were asked to read aloud unrelated sentences varying in syntactic complexity and content that were presented on a computer screen. Sentences such as: *When at last his eyes opened, there was no gleam of triumph, no shade of anger.* and *The taxi turned up Michigan Avenue where they had a clear view of the lake.* were presented. After reading two such sentences, the participant was asked to recall the final word of each sentence, e.g., *anger* and *lake.* The sentences were initially presented in 5 sets of 2 sentences. If the participant successfully recalled the last words from a minimum of three of the sets, they were presented with 5 sets of 3 sentences, and so on, until a maximum set-size of 6-sentences was reached.

The highest set-size for which the participant successfully recalled the last words from a minimum of 3 sets of sentences was considered to be their reading WMS. If the participant was only successful on 1 or 2 sets at a given level, a score between the two levels was given. For example, if 2 out of 5 sets at the 3-sentence set-size were successfully recalled, a score of 3.67 was given. Possible WMS scores ranged from 1.33 to 6, and the average score for young adults has been reported to be 2.67 (ibid.). In an attempt to clearly distinguish low from high WMS participants in this study, it was decided that those who scored at or below 2.67 would be assigned to the low WMS group and those who scored at or above 3.33 would be assigned to the high WMS group. Potential participants who scored exactly 3 were not included in the study. The mean and standard deviations for ages and working memory scores for the low and high WMS groups were 24.98 years (SD = 4.166) and 25.88 years (SD = 4.879), and 2.41 (SD = .29) and 3.63 (SD = .36), respectively.

Listeners from each WMS group performed equivalently on pure-tome threshold testing, word discrimination and speech recognition thresholds, hearing history, years of education, and vocabulary.

To ensure that peripheral auditory perceptual abilities for pure tones between participants in the low and high WMS groups were similar, pure-tone thresholds for all participants were tested and only those with normal audiograms were included in the study. The pure-tone air-conduction thresholds in both ears for all participants was between -10 and +25 dBHL for tones from .25 to 8 kHz (tested in octave steps). In addition, all participants had normal hearing histories.

A word discrimination score (N.U. Auditory Test, No. 6, Form D) and a speech recognition threshold (Spondaic word lists A-1, A-2 and A-3) were measured in each ear. All participants scored higher than 92% on the standard word discrimination test, and their speech recognition thresholds were between -10 and +5 dB, well within the normal range.

Because the speech stimuli were presented 50 dB above each participant's threshold for the eight-talker speech, a threshold measure for the babble was measured in the right ear only. The babble threshold ranged from -5 to +15 dB, consistent with the pure-tone thresholds and speech recognition thresholds.

The number of years of education at the time of testing ranged from 14 to 24 for all participants, with an average of 17.1 for the low WMS group and 17.5 for the high WMS group. The Mill Hill vocabulary test (Raven, 1938) revealed similar vocabulary scores across groups. They ranged from 12/20 to 17/20, with an average score of 14.6/20 for the low WMS group and 15.3/20 for the high WMS group.

2.2.2 Design

The questions this experiment was designed to address were whether lexical knowledge from long-term memory (varied in terms of word-frequency) and phonological cues (varied in terms of voicing and place of articulation) differentially influence the detection of target words in a range of sentence types in high and low WMS groups as background noise increases. Toward this end, a mixed-factor design with one between-groups factor and three repeated-measures factors was implemented in this experiment. The between-groups factor is reading WMS (high and low), and the repeated-measures factors are background noise level (high, medium and low), sentence complexity (high, medium, and low), and word frequency (high, medium, and low). In addition, target words varied dichotomously in voicing (voiced or voiceless), and place of articulation (bilabial, alveolar, or velar).

Each participant listened to three lists of digitized sentences. The order of sentence lists was counterbalanced across participants, however, the lowest level of background noise (0 dB S:N) was always presented with whichever list the listener heard first, followed by the next level of background noise (-3 dB S:N) and finally the highest level of background noise (-6 dB S:N). This ensured that the listener gradually became accustomed to progressively higher levels of background noise as the experiment proceeded.

2.2.3 Stimuli

Nine consonant-vowel-consonant (CVC) target words were embedded in sentences spoken by a young female whose speech was recorded with a Sennheiser KU3 unidirectional microphone placed 6 inches in front of her lips. The speech signal was transmitted through a Proport model 656 amplifier at a sampling rate of 32 kHz which, once digitized, was converted to a 22.05 kHz sampling rate with SoundRecorder audio software on a NeXT computer.

<u>Voicing</u> In order to assess the degree to which information at the phonological level influences word recognition, the final consonants of the target words in each sentence differed phonetically with respect to voicing and place of articulation cues (Ladefoged, 1975). The final consonants were: voiced velar-stop /g/, voiceless velar-stop /k/, voiceless alveolar-stop /t/, voiced alveolar-stop /d/, voiceless bilabial stop /p/, and voiced bilabial stop /b/. The final consonants in each block of target words included at least one final consonant from a voiced and one from a voiceless phonemic category. Regarding place of articulation cues, in one block one final consonant was alveolar and two were velar, in another block one final consonant was alveolar and two were bilabial, and in another block, two were velar and one was bilabial. Thus, the CVC target words in each block were: (1) *dog* /dag/, *dot* /dat/, *dock* /dak/, (2) *cat* /kæt/, *cap* /kæp/, *cab* /kæb/, and (3) *pig* /pIg/, *pick* /pIk/, *pip* /pIp/.

<u>Word-frequency</u> A second question of interest is the degree to which lexical knowledge stored in long-term memory (varied in terms of word-frequency) influences word recognition. Because words that are high in frequency are identified more quickly than low-frequency words

during brief presentations (Howes & Solomon, 1951), and high-frequency words are perceived more easily than low-frequency words in noisy conditions (Broadbent, 1967), it was predicted that high-frequency words presented as target lexical items in the sentences would be more easily identified than low-frequency words. Hence, each of the three target word blocks included a range of high-, medium- and low-frequency words (Francis & Kucera, 1982). The frequency count (based on number of times the word occurred per 1 million in written text[2]) for the target words is as follows: dog (147), dot (22), dock (7), cat (42), cap (22), cab (15), pig (14), pick (5), and pip (1). (A confound that was discovered after the data were collected is that the high-frequency words were also animate, and the medium and low-frequency words were inanimate. This is not a serious confound, however, because animacy is still a lexical knowledge-based cue).

Sentence Complexity In order to make the comprehension task vary in difficulty, each target word, which is always the agent of action in the sentence, was embedded in three different types of sentences which varied according to complexity. Sentences with two propositions are significantly more difficult to comprehend than sentences with one proposition when presented in noise to young listeners with normal hearing (Dillon, 1995), and to aphasics (Caplan, 1987). Therefore, the sentences used in this study corresponded to simple one-proposition (cleft subject), moderately complex two-proposition (co-ordinated), and highly complex two-proposition (subject-object) sentence types identified by Caplan (1987) and modified by Dillon (1995). Each of the sentences below was presented twice per set in quasi-random order. Each set of sentences was presented in one of each of the three levels of background noise for each participant. Thus three lists of 54 sentences were presented to each participant:

[2] It is reasonable to expect that word-frequency for written text is not the same as for spoken language. It is speculated that low frequency words may be more frequent in written text because it is generally a more formal expression of language. Nonetheless, while corpora of spoken speech exist, published frequency counts comparable to counts for written text (Francis and Kucera, 1982) were not available.

It was the **dog** that chased the duck.	Cleft Subject (low complexity)
The **dog** chased the duck and bumped the mouse.	Co-ordinated (moderate complexity)
The duck that the **dog** chased bumped the mouse.	Subject-object (high complexity)

It was the **dot** that chased the duck.
The **dot** chased the duck and bumped the mouse.
The duck that the **dot** chased bumped the mouse.

It was the **dock** that chased the duck.
The **dock** chased the duck and bumped the mouse.
The duck that the **dock** chased bumped the mouse.

It was the **cat** that chased the duck.
The **cat** chased the duck and bumped the mouse.
The duck that the **cat** chased bumped the mouse.

It was the **cap** that chased the duck.
The **cap** chased the duck and bumped the mouse.
The duck that the **cap** chased bumped the mouse.

It was the **cab** that chased the duck.
The **cab** chased the duck and bumped the mouse.
The duck that the **cab** chased bumped the mouse.

It was the **pig** that chased the duck.
The **pig** chased the duck and bumped the mouse.
The duck that the **pig** chased bumped the mouse.

It was the **pick** that chased the duck.
The **pick** chased the duck and bumped the mouse.
The duck that the **pick** chased bumped the mouse.

It was the **pip** that chased the duck.
The **pip** chased the duck and bumped the mouse.
The duck that the **pip** chased bumped the mouse.

In order to ensure that the target words became increasingly ambiguous, the sentences were presented in three levels of background noise. The question of interest was the degree to which background noise affects word recognition. Therefore, each digitized sentence was mixed with a segment of eight-talker babble in a digitized stereo file using SoundRecorder on the NeXT computer. The babble began 500 ms before each sentence and ended 500 ms after each sentence. Although the mixed sound files were composed of two separate sound tracks so that the signal-to-noise ratios of the sentence and babble could be adjusted for each sentence list during testing, the sentence and babble tracks were presented monaurally (to the participant's right ear only) during testing.

In order to ensure that the listener's working memory was, indeed, stressed during testing, an on-line memory score was obtained from each participant during testing by having the participants listen to sets of two, three, four or five words, and repeat them back to the experimenter after every third sentence they heard. On-line memory test words were chosen semi-randomly from the Daneman and Carpenter (1980) reading WMS test and were spoken live by the experimenter during testing.

2.2.4 Procedure

Screening The participants were tested in one 60 to 90 minute session, depending on how quickly each one completed the comprehension test. Before testing began, the participant's reading WMS measure was measured, and the hearing and vocabulary questionnaires were completed. The audiogram, speech-threshold and word-recognition tests were then conducted. Next, the experimenter trained the participant to complete the response form, and finally the sentence comprehension test was administered.

Training on the "Pen and Paper" Task During the training session for the "Pen and Paper" task, the listener sat at a desk in a sound-proof booth and was familiarized with the target words as the experimenter read each one aloud. The participant was also familiarized with the verbs *chased* and *bumped* which were to be written on the response form above the arrows to be drawn from the target word representing the agent of action to the word representing the patient (see Appendix B).

The experimenter then familiarized the participant with each sentence by reading each one aloud. The experimenter demonstrated how to complete the practice response form for the first three practice sentences, and read out the remaining 24 sentences while the participant finished completing the response form. Feed-back was given, and only when the participant felt comfortable with the task did the experimenter proceed with testing. The experimenter then placed the headphones on the participant and the experimenter moved to the adjacent sound booth control room to begin testing.

Set-up During testing, the participant sat at the same desk in the sound-proof booth that had been used for training. The experimenter was seated in the adjoining control booth facing the participant through the common window. Each booth was well lit. The experimenter presented the audio stimuli from a NeXT computer. The sentence stimuli and the background noise were transmitted from the computer, via the audiometer, to the participant's right ear only. By utilizing the "Pen and Paper" task (Lloyd & Pichora-Fuller, 1996), after listening to each sentence, the participants indicated their interpretation of the sentence by drawing arrows from the agent to the patient(s) on the response form.

Reliability of "Pen and Paper" Task Lloyd and Pichora-Fuller (1996) tested the reliability between the "Pen and Paper" and "object manipulation" task (OM task; Caplan, 1985; Dillon, 1995) used to assess comprehension of spoken sentences varying in complexity. The (OM) task test was adapted by Dillon (1995) from Caplan et al. (1985). Although the OM task is effective in assessing lexical, semantic and thematic errors in the comprehension of spoken language, it is very labor intensive to score, and, therefore, a procedure requiring less time to score was developed and the outcome compared to results from the OM task.

The "Pen and Paper" task was adapted by Lloyd and Pichora-Fuller (1996) from Cook (1975) and Kemper and Anagnopoulos (1993). The "Pen and Paper" and OM tasks were used to test normal-hearing young adults in a within-subject design. The subjects were tested with the OM task identical to that implemented by Dillon (1995) and on the newly adapted "Pen and Paper" task, in counterbalanced order. The results of the two tasks were then compared.

It was found that the number of sentences which participants interpreted erroneously was not significantly different between the "Pen and Paper" and OM tasks, and that the pattern of within-subject errors was consistent across tasks. In addition to compatibility with the OM task, the "Pen and Paper" task provides greater opportunity for testing a wider range of sentence types.

Familiarization with sentences in noise Three sample sentences were presented at the beginning of each new list of sentences in order to familiarize the participant with the higher level of background noise. The listener was asked to repeat each sample sentence back to the experimenter. The duration of the testing phase was participant-controlled; after completing the response form for each sentence, the participant would say "Go", indicating to the experimenter that she was ready for the next sentence. The experimenter would then depress a button on the computer, and the next sentence would play through the headphones. The participant could talk to the experimenter through a microphone in the sound-booth whose signal was transmitted through the audiometer and into an earphone worn by the experimenter. The experimenter could also talk to the participant by using the "talk over" feature on the audiometer.

On-line memory task During testing, to ensure that the working memory of the participant was, indeed, stressed, the experimenter verbally presented words for the participant to remember while comprehending the sentences. To ensure optimum encoding, the participant was asked to read the experimenter's lips (via the window that separated the sound-booths) while the words to be remembered were spoken at 0 dB HL over a microphone with equal stress on each syllable. The experimenter started by presenting 2 words to be remembered, then the participant heard and completed the response form for three digitized sentences. The participant was then asked to repeat the 2 memory words back to the experimenter. If 3 blocks of 2 words were accurately recalled, the experimenter would present a block of 3 memory words, and so on. The maximum number of words presented at once in a set was 5. If the participant did not accurately remember 3 blocks of words at a given set-size, the experimenter would continue to present the number of words at that set-size for the remainder of that particular list of sentences. The words to be remembered, such as *advance*, *sensitivity*, and *elders* varied randomly in number of syllables, word class and the degree to which they were concrete or abstract in meaning. They were chosen randomly from the Daneman and Carpenter (1980) reading WMS test. A sample of a three-sentence trial follows:

1. Presentation of words to be remembered: *advance, sensitivity, elders.*

2. Participant listens to Sentence 1:

   The cab chased the duck and bumped the mouse.

3. Participant completes form:

   cab cap cat dock dog dot duck mouse pick pig pip

4. Participant listens to Sentence 2:

   The duck that the dock chased bumped the mouse.

5. Participant completes form:

   cab cap cat dock dog dot duck mouse pick pig pip

6. Participant listens to Sentence 3:

   It was the pig that chased the duck.

7. Participant completes form:

   cab cap cat dock dog dot duck mouse pick pig pip

8. Instruction to recall words: *advance, sensitivity, elders.*

2.2.5 Apparatus

Two double-walled IAC sound-attenuating booths with a common window were required for testing (the experimenter sat in one, and the listener in the other). A NeXT work station with a motorola 68040 25 MHz cpu, a GSI 16 Audiometer, and a set of TDH-50P Telephonics 296D200-2 headphones was also required to transmit the digitized sentences and background noise to the participant. In addition, response forms (Lloyd & Pichora-Fuller, 1996) were required for the participants to indicate their interpretation of the sentence they heard.

2.3 Results

Phonological Cues It was predicted that, if the high WMS listeners were more able to resolve ambiguity than low WMS listeners (e.g. Just & Carpenter, 1987, 1992; Connine, Blasko, & Wang, 1994) because of superior temporal auditory processing, which would result in easier extraction of the signal from background noise, then they would correctly detect more target words than low WMS listeners in noisy conditions, as phonological cues become more ambiguous. In other words, as listening becomes more stressful, the low WMS listeners were expected to reallocate WMS resources from storage to processing, which, due to capacity limitations, would result in more errors in detection of phonological cues (see Just & Carpenter, 1987).

A mixed 2 (WMS) between-subject x 3 (Noise) x 2 (Phonological Feature) within-subject analysis of variance (ANOVA) with Percent of Phonological Cues Correct as the dependent variable was conducted to investigate the relative roles of voicing vs place of articulation cues in word identification. Main effects for WMS, Noise, and Phonological Feature (Voicing vs Place of Articulation) were found, revealing that the high WMS listeners detected both types of phonological cue better than the low WMS listeners ($p < .01$), and that voicing cues were confused less than place of articulation cues by both groups across all noise levels ($p < .01$).

High WMS listeners identified significantly more phonological cues (voicing and place features) correctly compared to Low WMS listeners [(F (1, 22) = 8.248, $p < .01$; see Figure 1)]. Listeners correctly identified significantly fewer voicing and place features as background noise increased (F (2, 44) = 13.468, $p < .01$). Post-hoc Newman-Keuls tests revealed that significantly fewer phonological features were correctly identified in the high vs low and high vs medium noise conditions ($p < .01$), but not the medium vs low conditions ($p > .05$; see Figure 2). Listeners' correct identification of voicing features was significantly better than their correct identification of place features [(F (1, 22) = 104.302, $p < .01$; see Figure 3)].

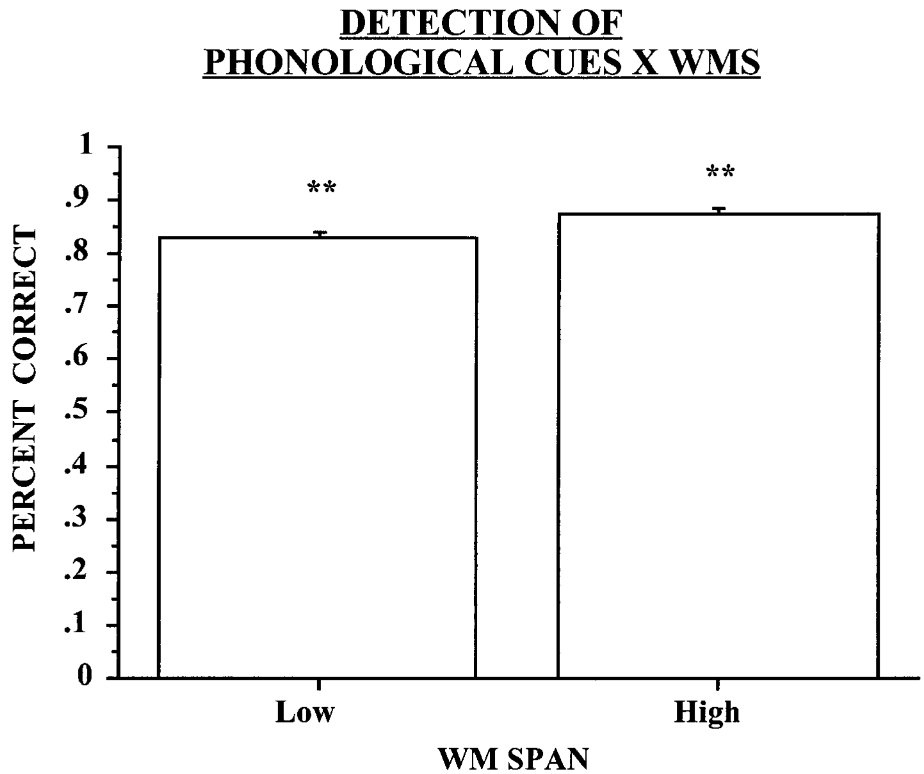


Figure 1: Percent correct detection of phonological cues (voicing and place of articulation features collapsed) by WMS in Experiment 1.

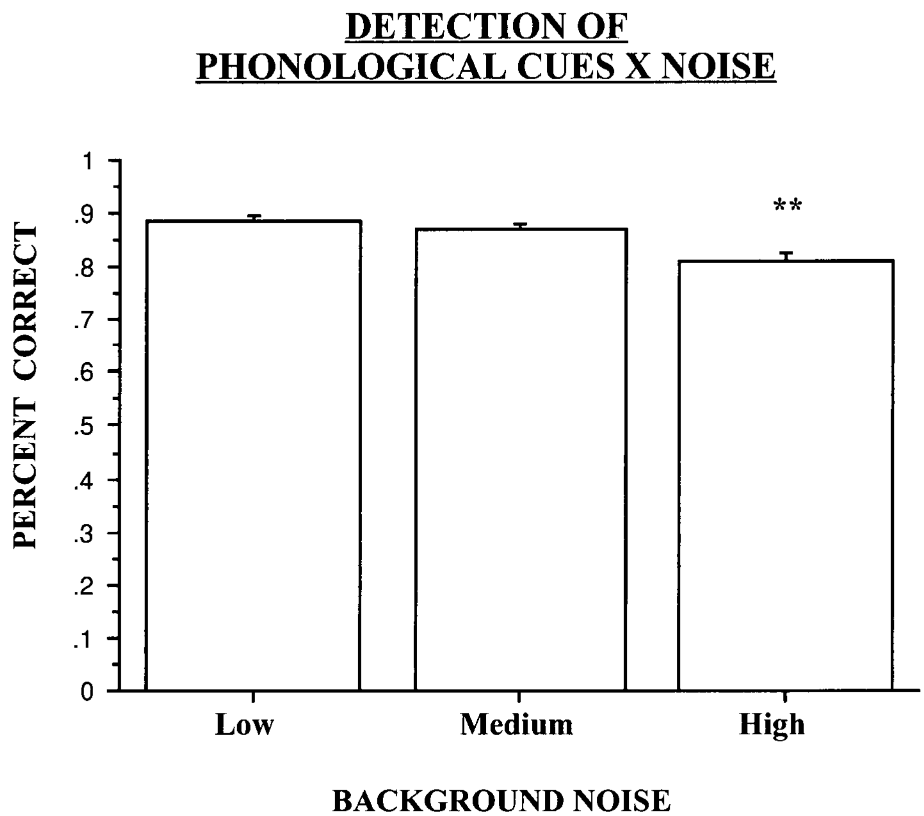


Figure 2: Percent correct detection of phonological cues (voicing and place of articulation features collapsed) analyzed by noise condition in Experiment 1.

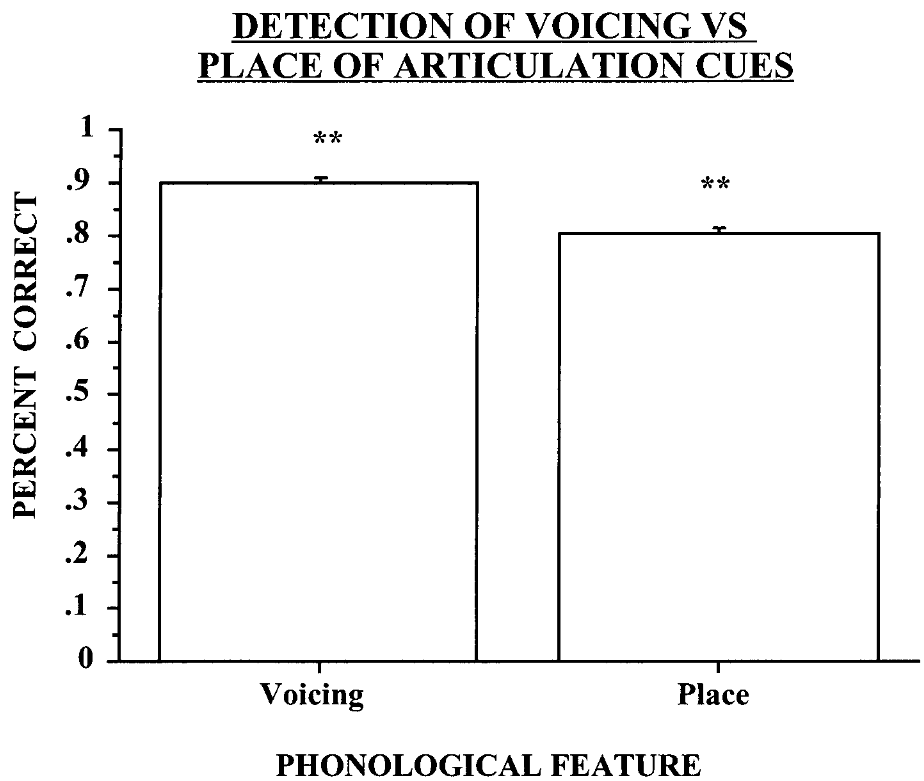


Figure 3: Percent correct detection of voicing vs place of articulation cues in Experiment 1.

Lexical Cues It was predicted that, if high WMS listeners' superior perceptual abilities enable them to extract the phonological signal from background noise more easily, then they would rely less than low WMS listeners on word-frequency in determining their responses. Such results would provide converging evidence for the hypothesis that word-frequency interacts with phonetic level processing (Altmann, 1990) and provide new evidence that this occurs to a greater extent for low than high WMS listeners.

A mixed 2 (WMS) between-subject x 3 (Noise) x 3 (Sentence type) x 3 (Word-frequency) within-subject analysis of variance (ANOVA) was conducted with Number of Correct Target Words as the dependent variable. The results showed that high WMS listeners identified significantly more target words than low WMS listeners [(F (1, 176) = 9.285, $p < .01$; see Figure 4)].

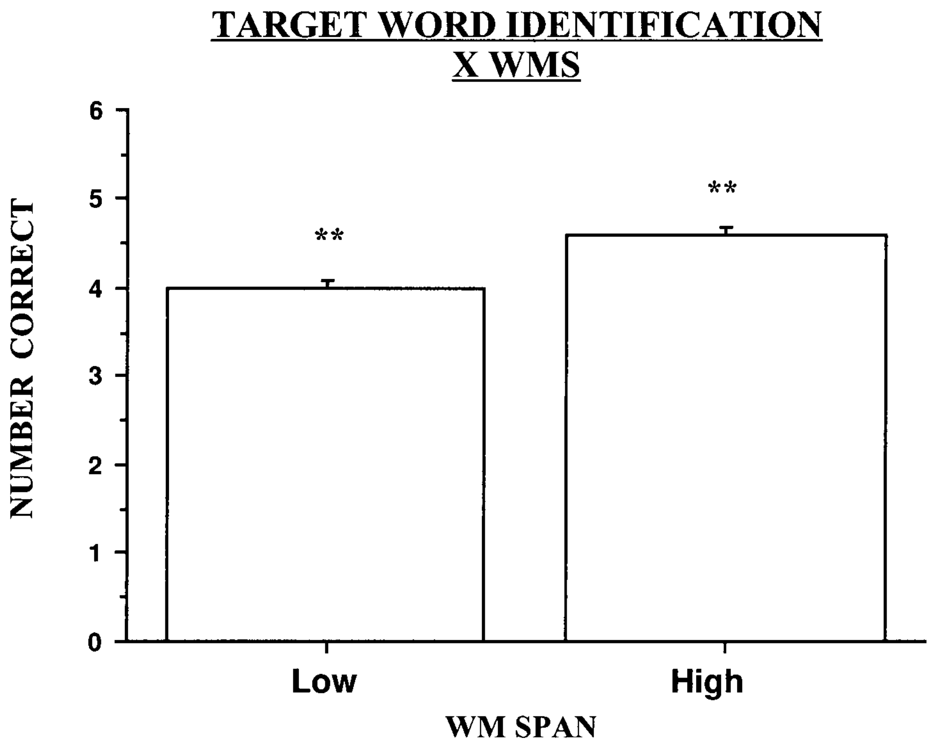


Figure 4: Number of Target Words correctly identified by WMS in Experiment 1.

Main effects for Noise, Sentence Complexity and Word Frequency also showed that each factor plays an important role in spoken language comprehension. Listeners identified significantly fewer target words as background noise increased (F (2, 44) = 16.51, $p < .0001$). Newman-Keuls post-hoc tests revealed that significantly fewer target words were identified correctly in the high noise condition compared to the medium and low noise conditions ($p < .01$), but no significant difference was found between the medium and low conditions ($p < .05$; see Figure 5).

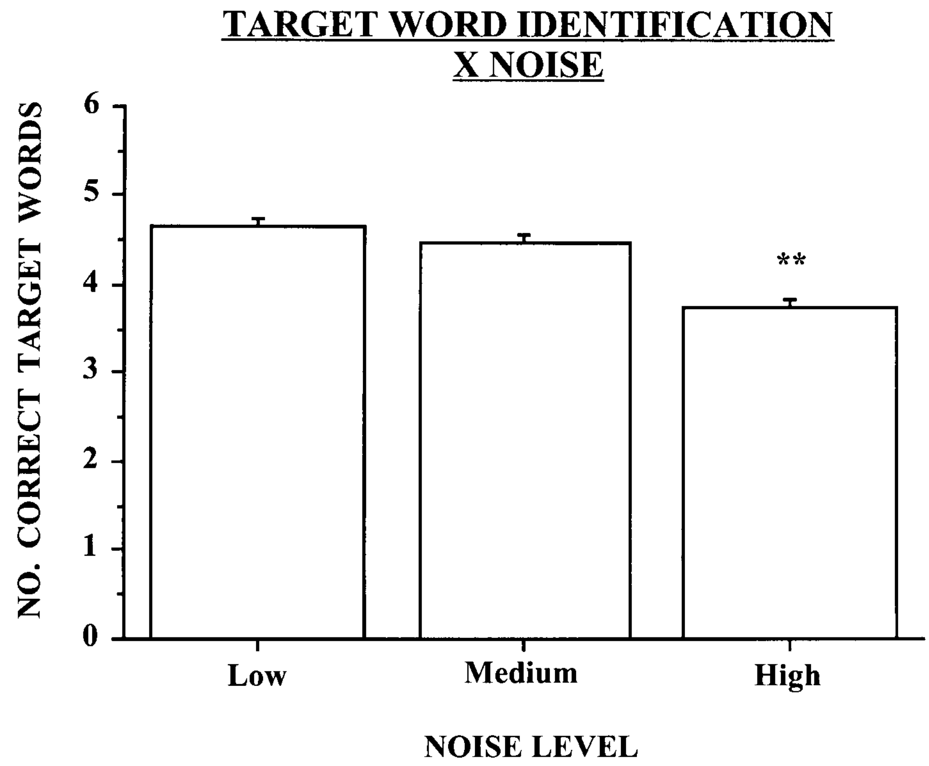


Figure 5: Number of Target Words correctly identified by noise condition in Experiment 1.

Listeners identified significantly fewer target words as sentence complexity increased (F $(2, 44) = 9.68$, $p < .001$). Newman-Keuls post-hoc tests revealed that significantly fewer target words were identified correctly in the high vs low and medium complexity conditions ($p < .01$), but no significant differences were found between the medium and low complexity conditions ($p > .05$; see Figure 6).

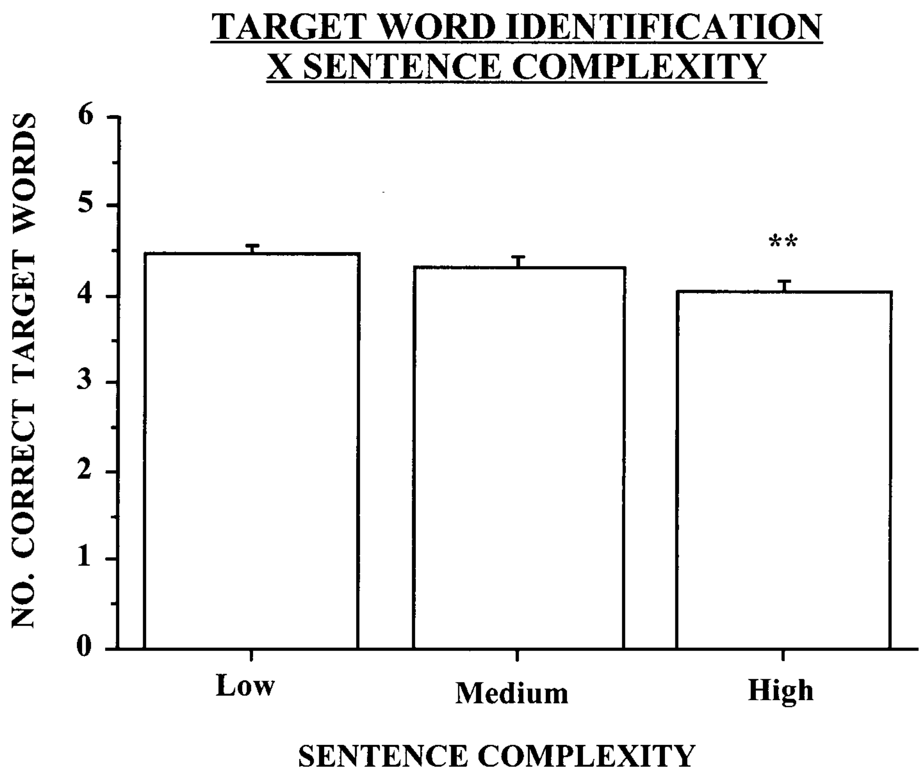


Figure 6: Number of Target Words correctly identified by sentence complexity in Experiment 1.

Listeners identified significantly more target words as the words increased in frequency count (F (2, 44) = 48.579, $p < .0001$). Newman-Keuls post-hoc tests revealed that significantly more high-frequency vs medium- or low-frequency target words were identified ($p < .01$), but no significant difference was found between the medium and low word frequency target words ($p > .05$; see Figure 7).

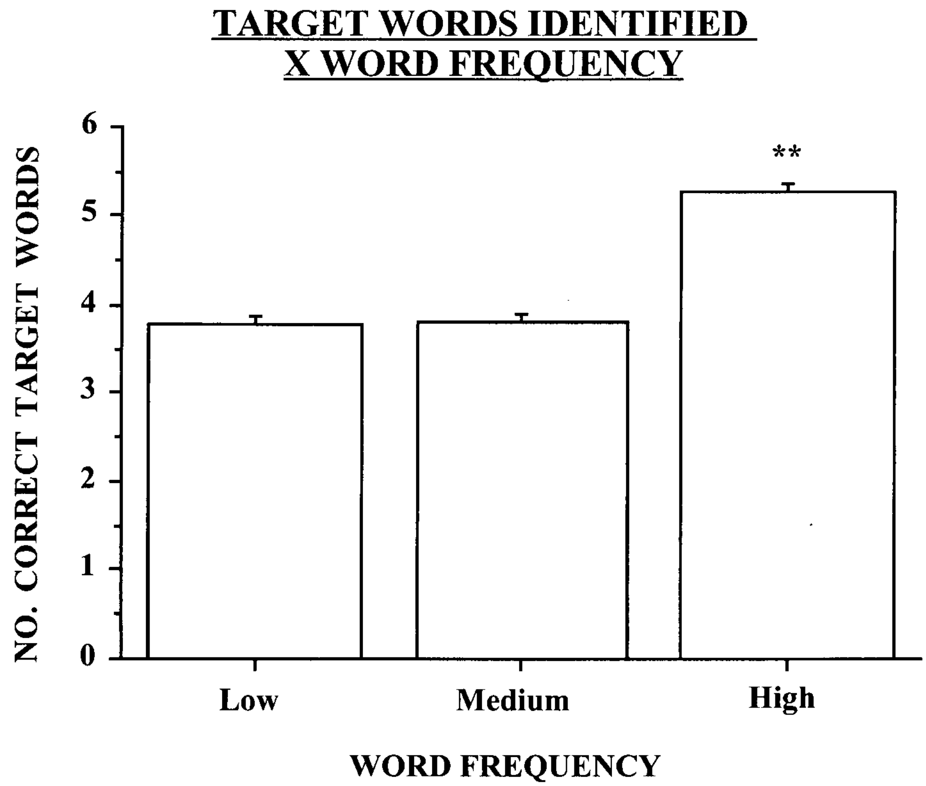


Figure 7: Number of Target Words correctly identified by word frequency in Experiment 1.

Although the interaction between WMS, Noise and Word-Frequency was only marginal (F $(4,88) = 1.93$, $p = .1124$), planned protected t-tests (Fisher's LSD test, hereafter referred to as planned protected t-tests) indicated that high WMS listeners detected low-frequency words significantly better than low WMS listeners in both the low and high noise conditions ($p < .01$), and high WMS listeners also detected high-frequency words significantly better than low WMS listeners in the high noise ($p < .01$), but not the low noise condition ($p > .05$; see Figure 8).

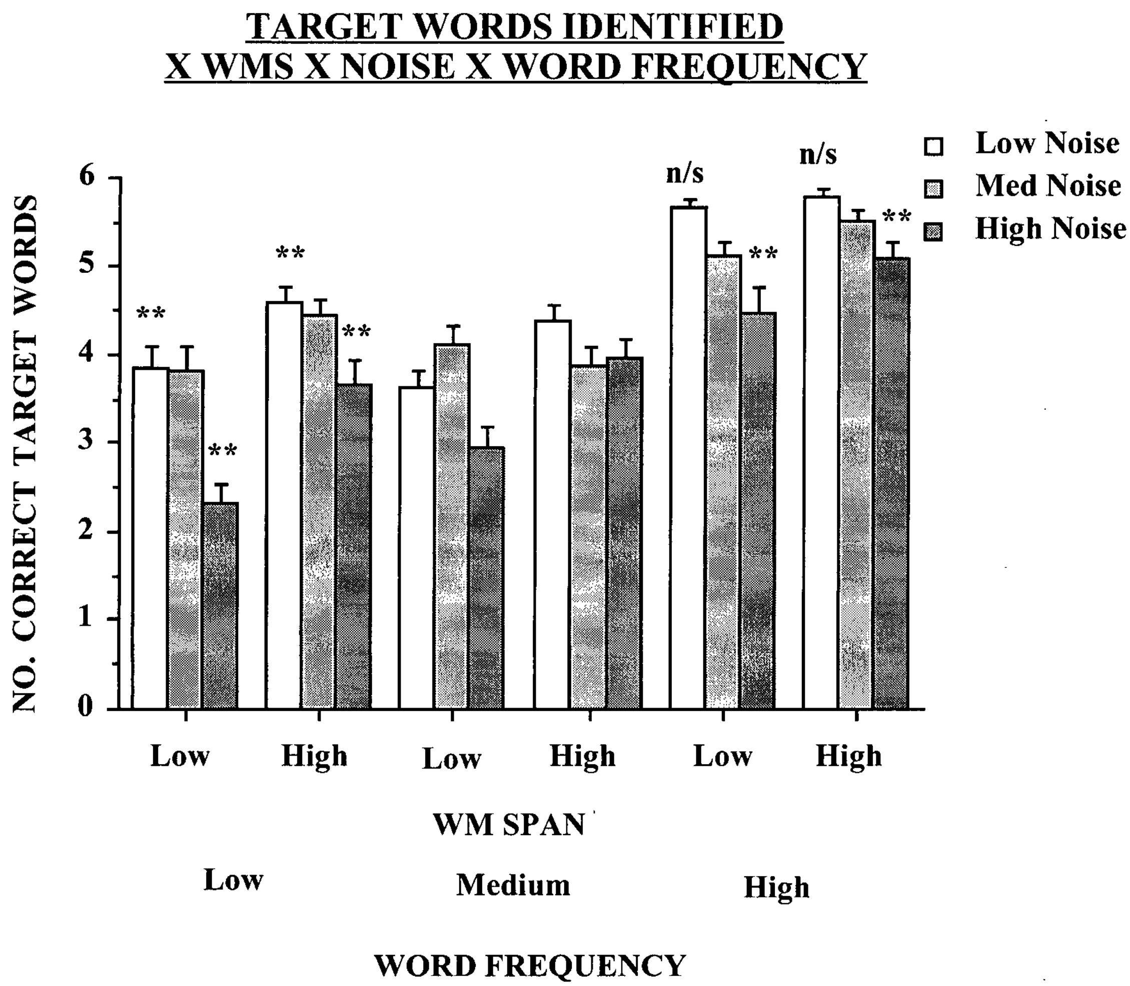


Figure 8: Number of Target Words correctly identified by WMS, noise and word frequency in Experiment 1.

A mixed 2 (WMS) between-subject x 3 (Noise) x 3 (Sentence type) x 3 (Word-frequency) within-subject analysis of variance (ANOVA) was also conducted with Number of Correct Sentences as the dependent variable (all nouns, verbs, and direction of arrows were required to be correct in order for the sentence to be scored as correct). A trend for WMS was found, and main effects for Noise, Sentence Complexity and Word-Frequency also showed that each factor plays an important role in correct sentence interpretation. Although the interaction between WMS and Noise was only marginal ($\underline{F}$ (2,44) = .496, $\underline{p} > .05$), planned protected t-tests indicated that high

WMS listeners correctly interpreted significantly more sentences than the low WMS listeners in the high noise condition ($p < .01$; see Figure 9).

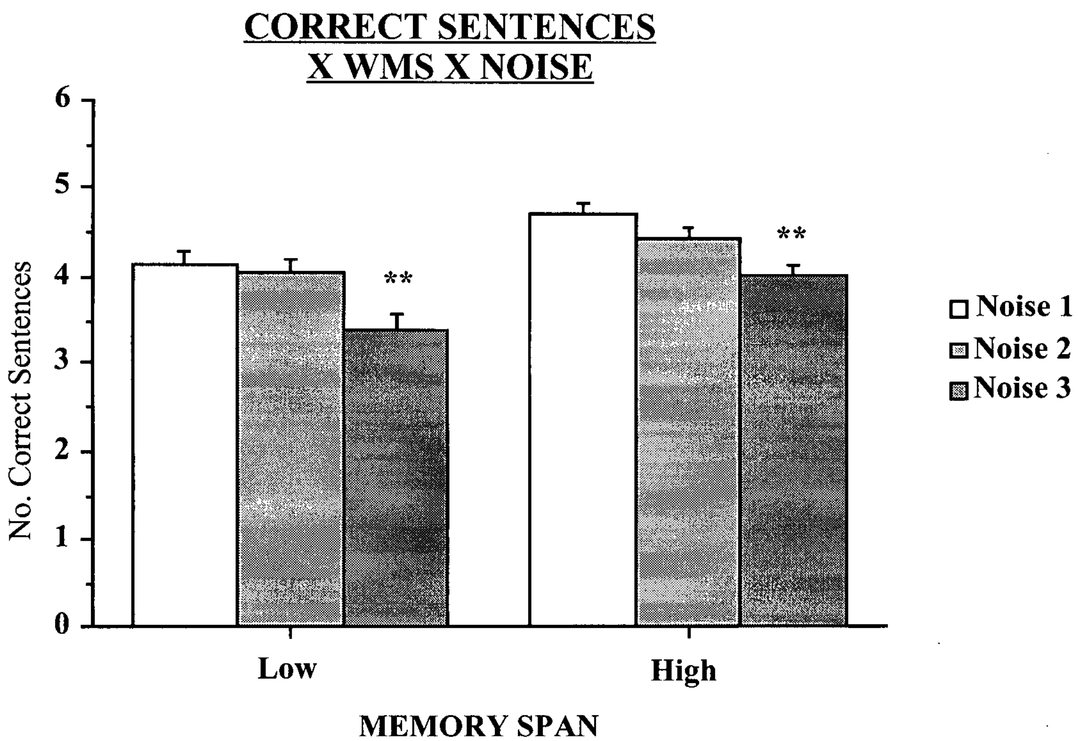


Figure 9: Number of Sentences correctly completed by WMS and noise condition in Experiment 1.

Although the interaction between WMS, Noise, and Word-Frequency was only marginal (F (4,88) = 1.698, $p = .1575$), planned protected t-tests indicated that high WMS listeners interpreted significantly more sentences correctly than the low WMS listeners in the high noise condition if the sentences contained low-frequency target words ($p < .01$), but not if they contained high-frequency target words ($p > .05$; Figure 10).

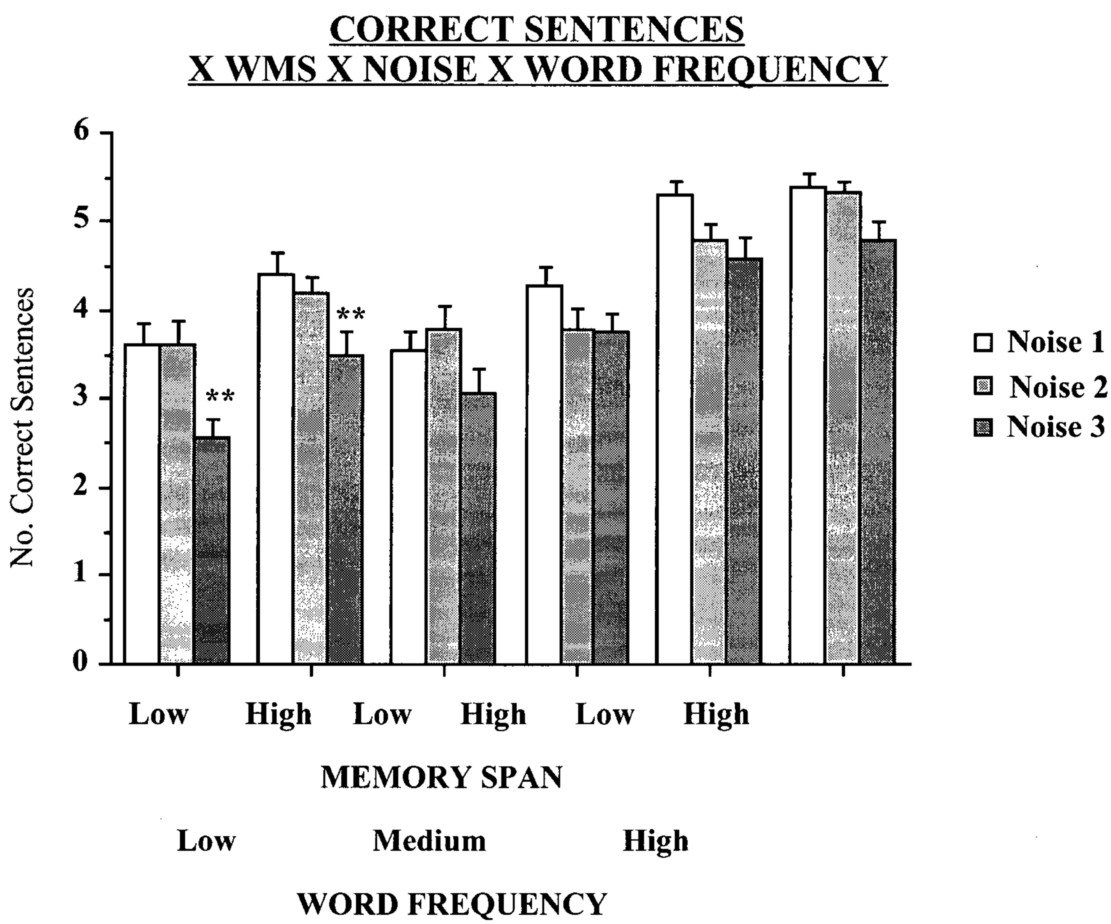


Figure 10: Number of Sentences correctly completed by WMS, noise condition and word frequency in Experiment 1.

Word-frequency vs animacy Because word frequency and animacy were confounded in the experimental design, a mixed 2 (WMS) between-subject x 3 (Sentence Complexity) x 3 (Word Category) within-subject analysis of variance (ANOVA) was conducted with Percent Correct for Animacy as the dependent variable.

A significant main effect for Word Category was found (F (2,44) = 50.623, $p < .0001$. Newman-Keuls post-hoc tests revealed that Percent Correct for Animacy was significantly lower in the *cat/cap/cab* category than the *dog/dot/dock* or *pig/pick/pip* categories, indicating that, overall, animacy was not as strong a factor as word frequency in listeners' correct identification of the target word (Figure 11).

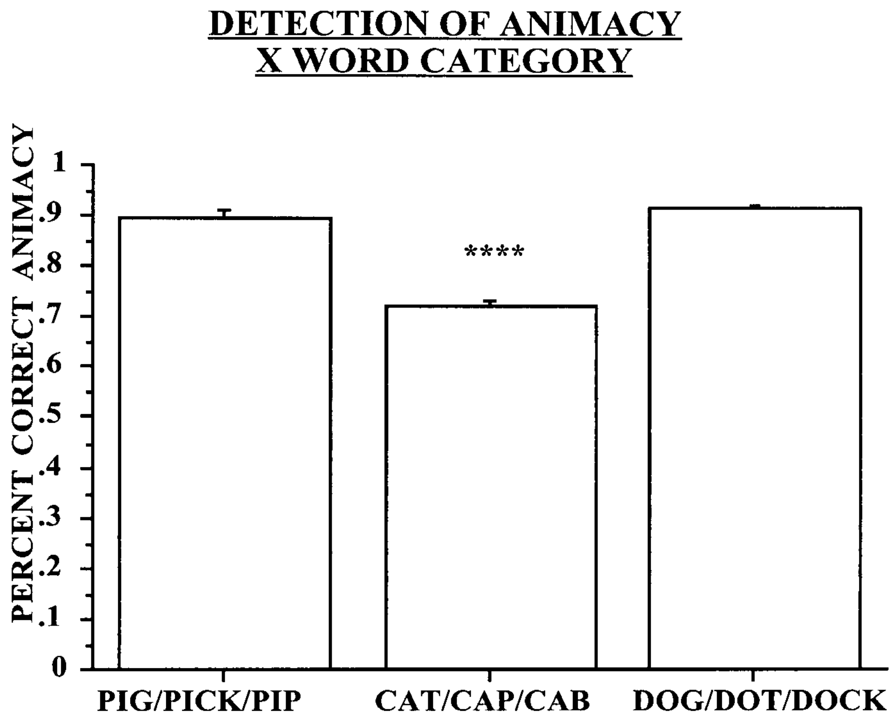


Figure 11: Percent correct detection of animacy by word category in Experiment 1.

Thus, for the categories in which the high-frequency/animate word was also voiced (dog/dot/dock and pig/pick/pip), the high-frequency/animate words were highly distinguishable from those that were medium- and low-frequency/inanimate, but also voiceless. For the category in which the high-frequency/animate word was not also voiced (cat/cap/cab), the distinguishability of the high-frequency word from the medium- and low-frequency/inanimate words was significantly lower than for the high-frequency words in the other two word categories. Thus, it appears that the voicing cue can be an important factor above and beyond animacy as a factor.

The WMS x Word-frequency interaction may also be partly based on Animacy. Either way, the effect demonstrates that the influence of lexical level information, be it frequency or animacy, on phonological processing is stronger for low than high WMS listeners. In addition, because the high-frequency/animate target words were the agents of action in all sentence types, it is plausible that, for high WMS listeners, syntactic information may not interact with phonological processing as much as it does for low WMS listeners.

On-line memory scores It is important to note that a significant correlation was found between online memory score and WMS score (Pearson product moment correlation = +.541, t (22) = 3.02, p = .01 (two-tailed), because it provides converging evidence, from an on-line memory test, for the measure of working memory between groups.

2.4 Discussion

Hypotheses 1 and 2:

The results of Experiment 1 showed that phonological and word-frequency cues both interact with WMS to some degree, with the dominance of the cue depending on WMS. High WMS listeners demonstrated superior abilities in discriminating differences in phonological cues and were less influenced by lexical level cues than low WMS listeners in interpreting ambiguous target words. Moreover, it was speculated that this ability may be attributed to superior perceptual abilities possessed by high WMS listeners.

The results from Experiment 1 set the stage for further investigating why high WMS listeners identified ambiguous target words better across all conditions, why they interpreted more sentences correctly in high noise, and why they interpreted sentences with low-frequency words more accurately than low WMS listeners. The results could be due to the generation of fewer alternative responses to the ambiguous target word by high WMS listeners as a result of superior abilities in extracting the signal from background noise based on perceptual processing, or better evaluation of the alternatives by the high WMS listeners based on superior abilities to compare the input to possible interpretations stored in long-term memory. Experiments 2 and 3 were designed to test the above questions, and also to test the degree to which context at the lexical vs sentence levels interacts with phonological cues when listeners disambiguate speech signals in noise.

III EXPERIMENTS 2 and 3

3.1 Interaction of phonological and sentence level processing

Experiments 2 and 3 were designed to test whether high WMS listeners generate more alternatives to or evaluate ambiguous targets more successfully than low WMS listeners, and whether contextual information at the sentence level interacts with WMS as a partial explanation for why high and low WMS listeners process ambiguous target words differently. A second question these two experiments were designed to examine is the extent to which temporal auditory processing abilities may underpin individual differences in spoken language processing.

Phonological cues were varied by contrasting VOT of the initial consonant of the target words, and contextual information that would link to knowledge stored in long-term memory was varied at the sentence level. The word-frequency of target words was controlled. High and low WMS listeners were presented with normed[3], tightly controlled sentences in which equally frequent target words (which varied in voicing of the initial consonants, e.g., *bat*, *pat*) were embedded in three levels of sentence context: Context-Consistent, Context-Inconsistent, and Context-Neutral. (An example of the Context-Consistent type is *The kiddies pat the doggies and chase toy rockets*, Context-Inconsistent is *The kiddies bat the doggies and chase toy rockets*, and Context-Neutral is *The families pat the things and discuss the action.*) All sentences were of the moderately complex co-ordinated type; in this sentence type the target word precedes the informative part of the sentence, requiring the listener to hold the ambiguous target word in working memory while processing the meaning of the sentence (see Appendix A).

After hearing each sentence, the listeners were asked to indicate their interpretation of the sentence by clicking the appropriate box on a computer screen which showed the target words (bat or pat, which will be referred to as the Target Words). Each target word was paired with a synonym for the last verb in each sentence in Experiment 2, and, in Experiment 3, the actual last

---

[3] The sentences were normed by having a group of 32 undergraduate students indicate whether the word *bat* or *pat* best fit each sentence. Only sentences for which the target word was chosen significantly greater than chance ($p < .05$) were used in the Context-Consistent condition; significantly less than chance ($p < .05$) were used in the Context-Inconsistent condition; and, at chance ($p > .05$) were used in the Context-Neutral condition.

verb in each sentence. The synonym for, or the actual last verb of each sentence (for Experiments 2 and 3, respectively), will be referred to as the Context Word. The boxes on the computer screen displayed such choices as: Pat (Follow), Bat (Follow), Pat (Converse) in Experiment 2 (see Appendix C), and: Pat (Chase), Bat (Chase), Pat (Discuss) in Experiment 3 (see Appendix D), and so on. For the sentence *The kiddies pat the doggies and chase toy rockets*, for example, the correct response in Experiment 2 would be *Pat (Follow)*, and in Experiment 3 it would be *Pat (Chase)*. The sentences were presented in quiet and in three levels of background noise.

The target words in Experiments 2 and 3 differed only in voicing contrasts. It was predicted that high WMS listeners would identify the target words more accurately, as they had in Experiment 1, because it has been shown that voicing features are confused less than place of articulation features as background noise increases (Miller & Nicely, 1955), that both high and low WMS listeners identify voicing contrasts significantly better than place cues, and that voicing and place are confused less by high than low WMS listeners as background noise increases (Lloyd, 1998a; 1998b, Experiment 1 above).

It was predicted that the identification of target words in sentence contexts which are consistent with the dominant interpretation would be identified more accurately by both groups of listeners than those whose context is inconsistent or neutral, because it has been shown that target words are identified more quickly by adults and children when they are presented in semantically meaningful sentences than in sentences in which semantic context is anomalous (Marslen-Wilson & Tyler, 1980; Tyler & Marslen-Wilson, 1981). For example, in the Context-Consistent sentence *The athletes bat the average and secure the lead.*, the target word *bat* was expected to be interpreted as *bat* more frequently than in the Context-Inconsistent sentence *The kiddies bat the doggies and chase toy rockets*, in which it was expected sometimes to be erroneously interpreted as *pat* which would be a more expected word given the sentence context. The question of interest was whether the low WMS group would make significantly more errors, indicating that their interpretation of the ambiguous target words is influenced more than the high WMS listeners' by the sentence context, as it was by word-frequency in Experiment 1.

In Experiment 2, in order to ensure that the listeners were processing the sentences in a deep, meaningful fashion, and hence, that their working memory was taxed, the responses on the computer screen consisted of the target word and a synonym for the final verb in each sentence, as stated above. The frequency of the target words was held constant and contextual information from long-term memory was varied by manipulating sentence context. In order to control for the confound between word-frequency, animacy and phonological cues present in Experiment 1, the target words were matched for frequency, verbs rather than nouns were used, and the initial consonants of the target words differed in voicing cues only. A second reason to control the frequency of the target words when testing the influence of sentence level context on their interpretation is based on the findings of Marslen-Wilson (1990) that when multiple interpretations of an ambiguous word are activated, those higher in frequency are activated more quickly.

Experiment 3 was identical to Experiment 2, except the responses on the computer screen consisted of the target word and the final verb of each sentence (rather than a synonym for the final verb). Thus the task required only shallow processing. For Experiments 2 and 3, a psychoacoustic test (gap detection) was administered to each listener so that a comparison between temporal auditory processing performance and comprehension performance could be made.

3.2 Experiment 2: Deep-processing task

In the second experiment, sixteen high and sixteen low WMS listeners were tested on their ability to correctly identify naturally-produced CVC target words with either voiced or voiceless initial consonants when the target words were embedded in sentences which varied in context. The sentence context which followed the target word either supported the correct identification of the target word (Context-Consistent), supported the incorrect identification of the target word (Context-Inconsistent), or neither supported the correct nor the incorrect identification (Context-Neutral). Immediately after each sentence was played through the headphones, the listener selected the response (target word and context word) on the computer screen which corresponded with their interpretation of the sentence. The responses included the target word and a synonym for the final verb in the sentence to ensure that the listener was processing the sentence in a deep, semantic manner.

3.3 Method

3.3.1 Participants

Thirty-two (16 high and 16 low WMS) university students between the ages of 19 and 33 years who gave their informed consent, participated in this study. The mean age was 24.10 yrs., with a standard deviation of 4.11. Participants were recruited through newspaper and poster advertisements distributed in and around the University of British Columbia. All participants were native English speakers, and participants received a token payment for their participation.

As in the previous experiment, participants were assigned to a high or low WMS group depending on their WMS scores. In this experiment the listeners' WMS scores were tested twice for reliability (within a 48 hour period), and only if the listener scored below 3 both times, or above 3 both times, were they included in the study. The mean scores from both tests were used as the final WMS score. If a listener scored above 3 on one test and below 3 on the other, a third test was given. If a score of 3 was achieved on the third test, the listener was not included in the study. Otherwise, they were assigned to the group into which 2 of their scores fell.

The means and standard deviations for the ages of the low and high WMS groups were 24.1 years. (SD = 3.84 years) and 24.1 years. (SD = 4.5 years). The means and standard deviations for the WMS scores were 2.4 (SD = .25) and 3.6 (SD = .41), for the low and high WMS groups, respectively. Listeners from each WMS group were matched on peripheral auditory processing abilities, word discrimination and speech recognition thresholds, hearing history, years of education, and vocabulary.

To ensure that peripheral auditory processing abilities between participants in the low and high WMS groups were similar, hearing sensitivity of all participants was tested and only those with normal audiograms were included in the study. All participants had pure-tone air-conduction thresholds below 25 dBHL in both ears for tones from .25 to 8 kHz tested in octave intervals.

All participants scored higher than 92% on the word discrimination test, and the speech recognition threshold were between -10 and +15 dBHL. In addition, all participants had normal hearing histories based on a hearing history questionnaire.

The number of years of education at time of testing was also matched for the high and low WMS groups and ranged from 13 to 24 years for all participants, with an average of 15.5 for the low WMS group and 17.5 for the high WMS group. The groups were also matched for their scores on the Mill Hill vocabulary test, with scores ranging from 10/20 to 16/20. The average scores for the low and high WMS groups were 13.4/20 and 14.8/20, respectively. Finally, the temporal auditory processing of each individual was tested using the gap-detection test (see below for results).

3.3.2 Design

A mixed 2 (WMS) between-subject x 3 (Noise) x 3 (Sentence Context) x 2 (Voicing Cue) within-subject design was implemented. As stated earlier, each list of sentences was presented in counterbalanced order in 4 conditions (quiet and three levels of background noise). Two exemplars of two different sentences for each of two different target words (voiced and voiceless) for each of three different sentence types -- Context-Consistent, Context-Inconsistent, and Context-Neutral -- were recorded separately and included in each block (2 x 2 x 2 x 3 = 24 sentences per block). Each list included 5 blocks of sentences (2 x 2 x 2 x 3 x 5 = 120 per list). Within each list, the sentences were presented in semi-random order in each of 5 blocks, with the order of sentences within the blocks varying, and rotating through the lists to result in four different orders of list presentation (see Table 1, below).

Table 1

Sentence Lists for Experiment 2

ORDER 1		ORDER 2		ORDER 3		ORDER 4	
List 1	Block 1	List 2	Block 6	List 3	Block 11	List 4	Block 16
Quiet	Block 2	Quiet	Block 7	Quiet	Block 12	Quiet	Block 17
	Block 3		Block 8		Block 13		Block 18
	Block 4		Block 9		Block 14		Block 19
	Block 5		Block10		Block 15		Block 20
List 2	Block 6	List 3	Block 11	List 4	Block 16	List 1	Block 1
Low Noise	Block 7	Low Noise	Block 12	Low Noise	Block 17	Low Noise	Block 2
	Block 8		Block 13		Block 18		Block 3
	Block 9		Block 14		Block 19		Block 4
	Block10		Block 15		Block 20		Block 5
List 3	Block 11	List 4	Block 16	List 1	Block 1	List 2	Block 6
Med Noise	Block 12	Med Noise	Block 17	Med Noise	Block 2	Med Noise	Block 7
	Block 13		Block 18		Block 3		Block 8
	Block 14		Block 19		Block 4		Block 9
	Block 15		Block 20		Block 5		Block10
List 4	Block 16	List 1	Block 1	List 2	Block 6	List 3	Block 11
High Noise	Block 17	High Noise	Block 2	High Noise	Block 7	High Noise	Block 12
	Block 18		Block 3		Block 8		Block 13
	Block 19		Block 4		Block 9		Block 14
	Block 20		Block 5		Block10		Block 15

3.3.3 Stimuli

In order to create target words that were clearly distinguishable in quiet, but perceptual-lexically ambiguous in background noise, two naturally-produced CVC speech tokens were used. Because voicing features are confused less than place of articulation features as background noise increases (Miller & Nicely, 1955), and voicing cues are confused less by high than low WMS listeners as background noise increases (Lloyd, 1998a; 1998b, Experiment 1 above), tokens which varied in voicing contrast of the initial consonant were presented. The voiced token used was *bat* /bæt/ and the voiceless token was *pat* /pæt/. Word frequency for the verb form is relatively low for both /bæt/ (1) and for /pæt/ (5) (Francis & Kucera, 1982).

In Experiment 2, only one type of syntactic structure (co-ordinated) was used in creating the stimulus sentences because the results of Experiment 1 demonstrated that this sentence type elicits similar responses by high vs low WMS listeners (although the particular analysis is not included in this thesis). A second reason to use co-ordinated sentences is that they require listeners to hold the ambiguous target word (which in this case is the first verb) in memory while processing the context which follows, thereby stressing working memory during the listening task (e.g. Wingfield, 1996). Finally, in Experiment 2 the ambiguous cue was the initial rather than the final consonant of the target word, based on phonetic analyses which show that voiced final stop consonants are not always articulated distinctly (Ladefoged, 1975). The CVC tokens were embedded in three types of sentences: Context-Consistent, Context-Neutral, and Context-Inconsistent, as shown in Appendix A.

The sentences were designed with the following controls: 1) Vowels before and after the target consonant are held constant, 2) /b/ and /p/ were not included in sentences except in the target words, 3) 9 words were included per sentence, and 4) 12 syllables were included per sentence. In addition, the sentences were normalized by having a group of 32 undergraduate students indicate whether the word *bat* or *pat* best fit each sentence. Only sentences for which the target word was chosen significantly greater than chance ($p < .05$) were used in the Context-Consistent condition, those chosen significantly less than chance ($p < .05$) were used in the Context-Inconsistent condition, and words chosen at chance ($p > .05$) were used in the Context-Neutral condition.

Each of four lists of sentences were presented at 70 dB SPL in quiet and in three levels of background noise that increased in relation to the speech signal [+5 dB S:N, 0 dB S:N, and -5 dB S:N]. Each listener heard all four lists, in counterbalanced order, with 2 repeats of 5 samples of each of 12 sentences presented per list (for a total of 120 sentences per list). Although the lists were presented in counterbalanced order, the listener always heard the first list in quiet and subsequent lists in increasing levels of background noise.

Furthermore, it is worthwhile to consider some general characteristics of the /b/ vs /p/ contrast:

1) Although voiceless consonants are generally more intense than voiced consonants, (Miller & Nicely, 1955, p. 346), the intensities for /p/ and /b/ are relatively similar (7 and 8 dB,

respectively, greater in intensity relative to the intensity of /θ/ as in *thin*, which is the English sound with the lowest intensity; Fry, 1979, p. 127). Thus, amplitude of the initial consonant for *bat* /bæt/ and*pat* /pæt/ is not a strong cue, so differences in discrimination are unlikely to be attributed to amplitude.

2) Vowel duration and intensity for /bat/ and /pat/ are identical for words spoken in carrier phrases (Ladefoged, 1975, p. 164), but may vary for words spoken in running speech depending on the position of the word in the sentence, the degree of stress on each word, and the personal characteristics of the speaker (p. 165).

3) VOT is the time between the onset of burst and onset of periodic activity (Lieberman & Blumstein, 1988, p. 216). In running speech, the average VOT is 7 ms for /b/ in initial position, and 28 ms for /p/ in initial position, but it varies considerably between talkers (Lisker & Abramson, 1964, p. 411).

Thus, there is no reason to expect that vowel duration or intensity will vary systematically between the words for *bat* /bæt/ and*pat* /pæt/. Rather, differences will be due to VOT and burst intensity.

4) The VOT burst may last no more than about 10 or 15 ms where the sound has little or no aspiration (e.g., /b/); where there is marked aspiration (e.g., /p/), it may be on the order of 50 ms. During the burst, the noise energy is spread rather widely over the spectrum but peaks of energy tend to occur at different frequency regions according to the place of articulation of the consonant. In the bi-labial sounds, [p] and [b], the maximum energy is generally in the low frequencies, at about 600-800 Hz. Although the intensity of the noise burst is much less in the voiced sounds, (Fry, 1979, p. 122-125), VOT is much longer for voiceless sounds (Ladefoged, 1975, p. 163).

An analysis of the specific stimuli used in Experiments 2 and 3 yielded the following results:

1) the words *bat* /bæt/ and*pat* /pæt/ varied in duration and average intensity (but remained between 391.25 and 533.20 ms, and .5201 and .6736 RMS volts, respectively),

2) the vowel duration of for *bat* /bæt/ and*pat* /pæt/ varied in duration and intensity (but remained between 221.05 and 266.95 ms, and .7014 and .9034 RMS volts, respectively),

3) the VOT varied between 6.25 and 9.95 ms for /bat/, and 40.45 and 64.4 ms for /pat/, but the energy of the overall burst for both words remained between .3499 and .5847 RMS volts. (Thus, any difference in perception between the two words would most likely have been based on VOT rather than differences in energy that occurred during the burst; see Appendix E for time waveforms of two sample sentences) and,

4) **only one exemplar** of the two words that preceded the target word (e.g., the kiddies, the athletes, the families) in each sentence were spliced onto the beginning of each sentence in its respective group so that any differences in target word identification between the sentences could not be attributed to differences in the first two words (Connine, Blasko, & Hall, 1991, p. 237).

3.3.4 Procedure

Using the Computerized Speech Research Environment (CSRE 4.5) developed by Avaaz Innovations Inc. (1994), interfaced with the Tucker Davis Technologies system, digitized sentences were presented binaurally through headphones to the participant who sat in front of a computer screen in a sound-attenuating booth. The participant was instructed to listen to each sentence and then as quickly and accurately as possible, indicate her interpretation of the sentence by clicking on the appropriate box on the computer screen with the mouse. The forced-choice responses which appeared on the computer screen for Experiment 2 are shown in Appendix C). A gap-detection task (Schneider, 1994) was also conducted with each listener completing 3 trials per ear.

3.3.5 Scoring

The Number of Correct Target Word and Context Word Responses, the Response Times for Correct Sentences, and the Gap-detection Thresholds were obtained.

3.4 Results

A mixed 2 (WMS) between-subject x 4 (Noise) x 3 (Sentence context) x 2 (Voicing contrast) within-subject analysis of variance (ANOVA) with Number of Correct Responses (both Target Word and Context Word were required to be correctly identified) as the dependent variable was conducted[4]. Main effects were found for Noise (F (3,44) = 497.703, $p < .0001$) and Voicing (F (1, 30) = 4.375, $p < .05$), and significant interactions were found for WMS x Noise (F (3, 90) = 4.127, $p < .01$) and WMS x Noise x Voicing (F (3, 90) = 3.289, $p < .05$). Planned protected t-tests showed that both high and low WMS listeners correctly interpreted fewer sentences in the High noise than Quiet condition [(F (1, 15) = 3536.56, $p < .001$) and (F (1, 180) = 2350.34, $p < .001$), respectively], but (as shown below) counter to the results in Experiment 1, the low WMS listeners accurately identified more sentences (both target word and context word correctly) than the high WMS listeners in the high noise condition [(F (1, 180) = 98.693, $p < .001$; see Figure 12)].

[4] Only responses with less than 5 second latencies were included in the data analyses, because it was decided a priori that any responses more than 3 standard deviations above the mean would be excluded. According to this criterion, the cut-off points were 5.3 seconds and 4.1 seconds for the high and low WMS listeners, respectively. In addition, the data from the first block in each list were not included in any of the data analyses for Experiment 2, because a A mixed 2 (WMS) Between-subject X 4 (Noise) X 5 (Block) Within-subject analysis of variance (ANOVA) with RT as the dependent variable showed a significant interaction between Noise and Block (p < .001). Follow-up Newman-keuls tests showed that RTs were significantly longer for Block 1 than Blocks 2-5 in the quiet and high noise conditions for all listeners (p < .01). Thus, the total number of correct responses for each sentence type per list was 16 (4 sentences of each type x 4 blocks), rather than 20 (4 sentences of each type x 5 blocks).

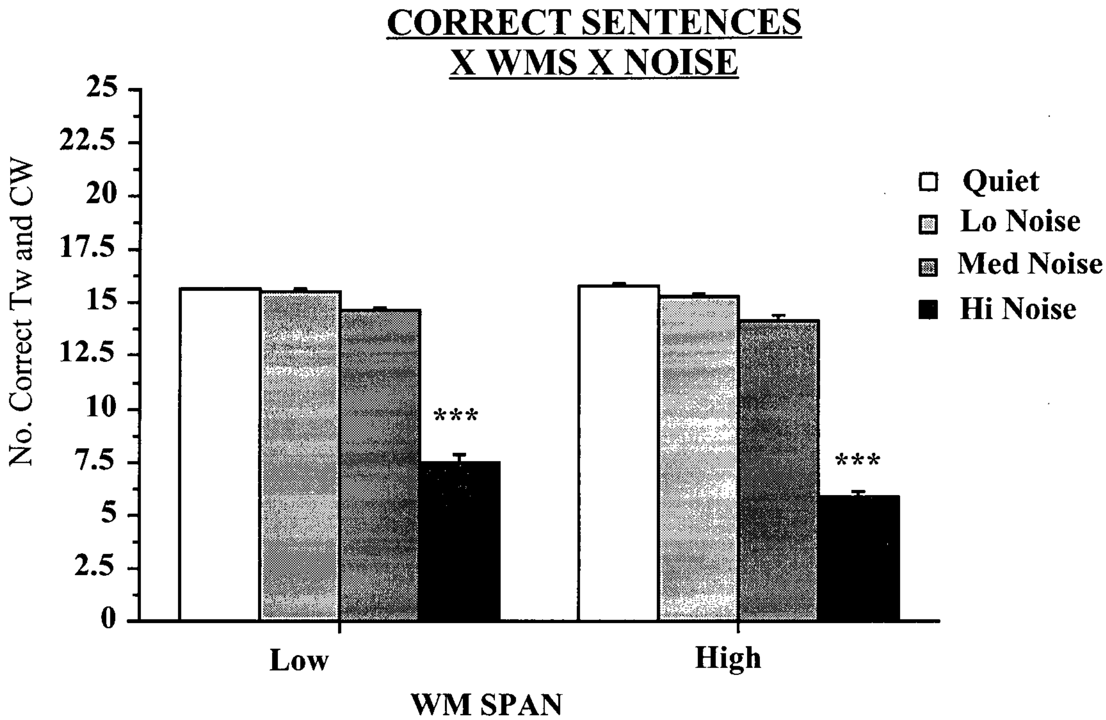


Figure 12: Number of Correct Target Words and Context Words (combined) by WMS and noise condition in Experiment 2.

Planned protected t-tests showed there was no significant difference for world-knowledge cues (sentence context) in Experiment 2 ($p > .05$) -- both low and high WMS groups responded correctly to the Context-Inconsistent and Context-Consistent sentences equally in high noise (see Figure 13).

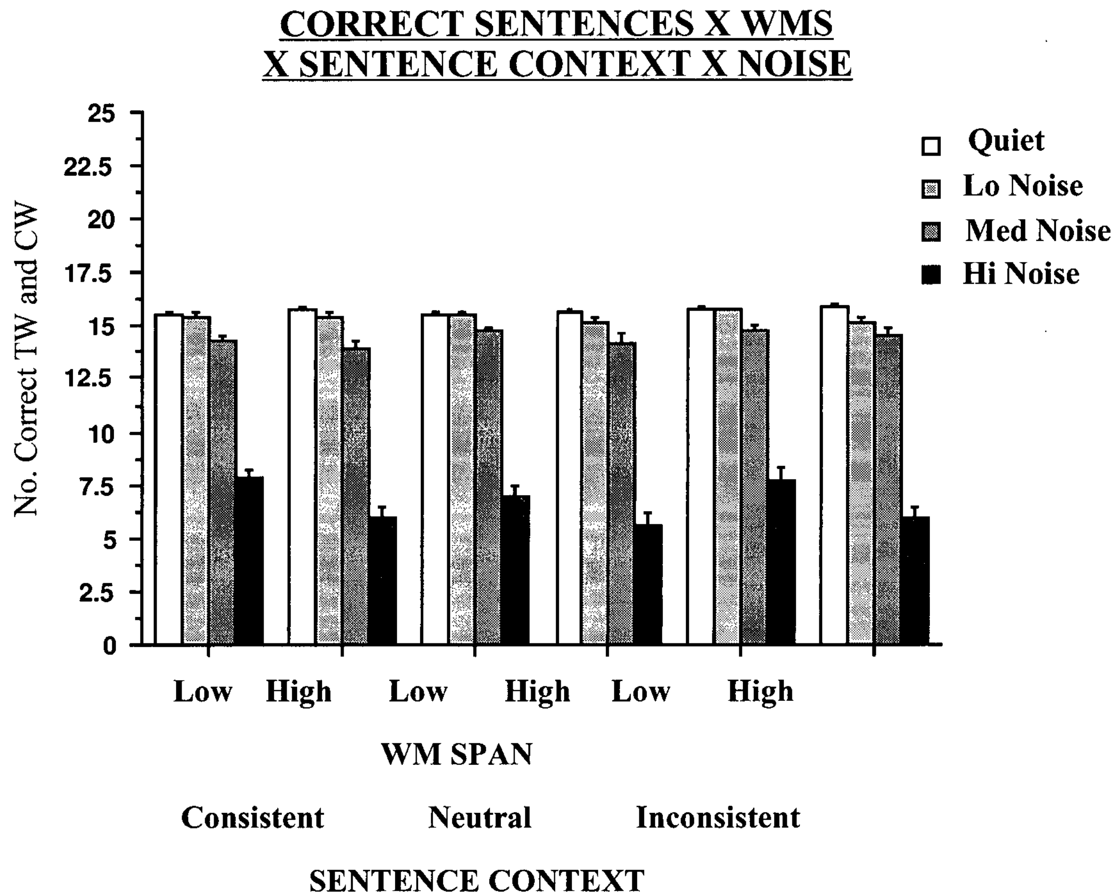


Figure 13: Number of Correct Target Words and Context Words (combined) by WMS, sentence context and noise condition in Experiment 2.

Thus, it appears that a WMS x Noise interaction is clearly present; however, the low rather than the high WMS listeners gave more correct responses in the high noise condition in which the speech signal was most ambiguous! Whereas Experiment 1 provided converging evidence for the body of literature that shows that lexical information appears to be capable of altering the percept of the individual sound segments that make up a word (Altmann, 1990, p. 18) and that this may be particularly relevant for low WMS listeners, the results of Experiment 2 provide converging evidence for the finding that sentential information only has a much more indirect effect on lower-level processing (ibid.). Perhaps by analyzing the correct responses to target word and context word separately, more insight may be gained regarding the generation or evaluation of alternatives, and differences in rehearsal capacity between the two groups of listeners.

Recall that the results of Experiment 1 showed the high WMS listeners correctly interpreted more sentences with low-frequency words in high noise and than did the low WMS listeners. They also identified target words better than the low WMS listeners. Thus, the strategy for the high WMS listeners may have been to attend to the target word better than or equally as well as the low WMS listeners, and complete the task more accurately because they were able to quickly extract the target word from background noise and hold the word in working memory while processing the remaining sentence input, whereas the low WMS listeners may have been more influenced by word-frequency due to inadequate processing time and poorer rehearsal capacity, which, for them, resulted in more errors. Unfortunately, because response latencies were not collected as a dependent variable in Experiment 1, it was not possible to examine this explanation for the different error rates.

To examine whether the low WMS listeners may have responded more accurately in Experiment 2 because they were able to extract the target word from background noise more easily than the high WMS listeners, an ANOVA for target word alone as the dependent variable was conducted. A main effect for WMS was not found (F (1, 30) = 1.096, $p > .05$), suggesting that the high and low WMS listeners attended equally to the target word. A further analysis for context word correct showed the low WMS listeners correctly identified significantly more context words than high WMS listeners (F (1, 30) = 4.453, $p < .05$). The low WMS listener's strategy may have been to listen for the target word and context word and respond as quickly as possible without considering alternative interpretations, indicating poorer rehearsal capacities (e.g., Baddeley & Logie, 1999). The high WMS listeners, on the other hand, may have considered alternative interpretations of the target word before settling on their identification of the target word, then proceeded to consider the context word, ultimately running out of rehearsal time to consider the context word adequately.

If, indeed, the high WMS listeners generated multiple activations for the ambiguous target word, it would be expected that they would show longer response latencies while considering the alternative responses than low WMS listeners in Experiment 2 (see also Connine et al., 1994). To test this hypothesis, an ANOVA using median response time (Med RT) for Correct Responses (target word and context word) as the dependent variable was conducted. It showed main effects for Noise (F (3, 90) = 62.019, $p < .0001$) and Sentence Context (F (2, 60) = 9.515, $p < .001$). The critical question was whether the high WMS listeners took longer to respond in the high noise condition (when the signal was most ambiguous) compared to the low WMS listeners.

Planned protected t-tests showed that both high and low WMS listeners did take significantly longer to respond in the high noise (ambiguous) than quiet (unambiguous) condition [(F (1,180) = 239.679, $p < .001$) and (F (1,180) = 99.958, $p < .001$), respectively], but the high WMS group took significantly longer to respond in high noise than the low WMS group (F (1, 180) = 25.39, $p < .01$). There was no significant difference in Med RT between the WMS groups in the quiet condition ($p > .05$; see Figure 14).

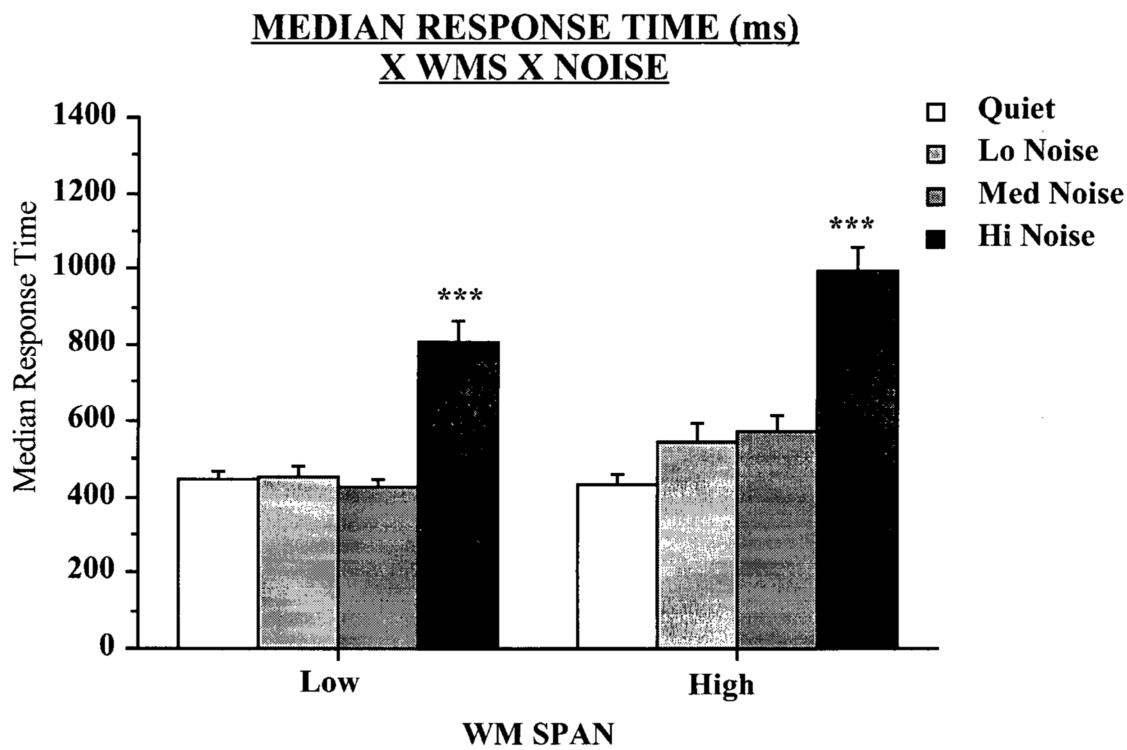


Figure 14: Median response time (ms) by WMS and noise condition in Experiment 2.

Thus, converging evidence is provided for Connine, Blasko and Wang's (1994) thesis that high WMS listeners take longer to respond to ambiguous vs unambiguous speech tokens than low WMS listeners as a result of generating multiple alternatives.

In order to contrast this finding with the influence of sentence context on listeners' interpretation of the ambiguous target word, Med RT to Context-Inconsistent vs Context-Consistent sentence contexts was examined in the high noise condition. If sentence context (and thus evaluation of multiple interpretations) was responsible for the high WMS groups' longer response latencies, they should perform faster on Context-Consistent than Context-Inconsistent sentences. Planned protected t-tests showed, in fact, that the high WMS listeners responded more slowly to Context-Consistent than Context-Inconsistent sentence contexts (F (1,180) = 9.12, p < .01), counter to the predicted direction of responses. In addition, the low WMS group showed no difference in Med RT to the Context-Consistent vs Context-Inconsistent contexts (p > .05; see Figure 15).

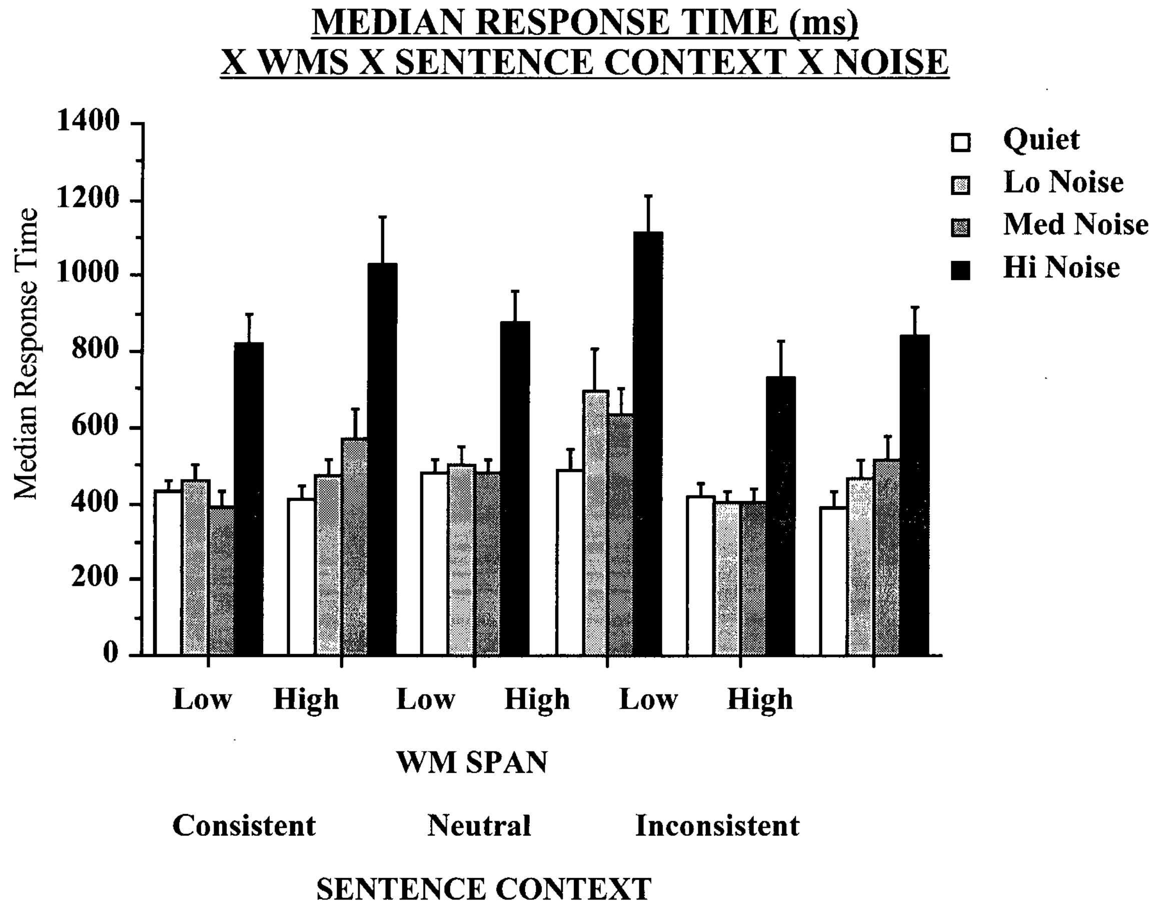


Figure 15: Median response time (ms) by WMS, sentence context, and noise condition in Experiment 2.

It appears that the High WMS listeners did not respond faster to the sentences with supportive context, but rather considered multiple interpretations of the ambiguous target word in a modular fashion, independent of sentence context.

The results of Experiment 2 support the hypothesis that high WMS listeners differed from low WMS listeners in their responses to ambiguous speech tokens as a result of generating more interpretations than low WMS listeners rather than as a result of more accurately evaluating the ambiguous tokens by using sentence context.

Recall that gap-detection thresholds were measured for each listener. The listeners were trained on the gap-detection task until they were able to detect gaps smaller than 10 ms. Then scores below 10 ms from the next three trials were included in the data analysis. In order to investigate the relationship between auditory perceptual abilities and WMS, a mixed 2 (WMS) between-subject x 3 (Trials) x 2 (Ear tested) within-subject analysis of variance (ANOVA) with Gap detection score as the dependent variable was conducted. A significant main effect for Ear tested, with the right ear showing larger gap scores was found. This is explained by the fact that the right ear was tested first in all listeners. A main effect for Trial was also shown, indicating a practice effect.

No main effect for WMS or interactions involving WMS were significant, indicating no significant differences in perceptual abilities across WMS groups. A trend for Ear x WMS ($p=.065$) was shown, however, indicating the high WMS listeners showed larger gap scores for the right ear (first ear tested). Thus, high WMS listeners may have had more difficulty warming up to the gap detection test. Whether this was due to perceptual processing, strategy in approaching the task (i.e., they may have used alternative strategies for performing the task, or interpreting the pure-tone signals) is left to speculation.

Several Pearson product-moment correlational tests were conducted to examine the relationship between gap thresholds (the gap score from the third trial of the left ear was used in the analysis to ensure the score reflected the listener's best performance), WMS and Number of Correct Responses and Median RTs on Correct Responses in the high noise condition.

The correlation between WMS and Gap thresholds was not significant ($p > .05$). Irrespective of WMS, perhaps a correlation would be found between Gap threshold and comprehension performance. The correlation between Gap threshold and the Number of Correct Responses in high noise in Experiment 2 was not significant ($p > .05$) for the high nor low WMS listeners. The correlation between Gap score and Med RT on Correct Responses in high noise was also not significant ($p > .05$) for either WMS group, suggesting that temporal auditory processing is not an underlying factor in the comprehension differences between high and low WMS listeners.

It appears that the results from Experiment 1 show that, as argued by Altmann (1990), word-frequency may interact directly with phonetic feature processing, whereas the results from Experiment 2 show the influence of sentence context on making lexical decisions in conditions of uncertainty is less direct.

3.5 Discussion

Hypothesis 3:

A Sentence Context x WMS interaction was not found, indicating that both groups relied equivalently on sentence context to resolve perceptual-lexical ambiguity. The low WMS listeners did not rely more than high WMS listeners on sentence context, and therefore, support is not evident for the hypothesis that low WMS listeners use sentence context to evaluate the alternatives any more than high WMS listeners. In addition, support is not evident for the argument that errors by low WMS listeners can be explained by reallocation of resources from storage to processing when contextual information is varied at the sentence level, as Carpenter et al. (1995) would predict.

Hypotheses 4A and 4B:

The results of Experiment 2 show that, compared to low WMS listeners, high WMS listeners made equivalent target word identification errors, more context word comprehension errors, and that their comprehension scores were not positively correlated with their gap-detection thresholds. These results suggest that they perceived an equally clear signal as the low WMS listeners, but generated more alternatives to the ambiguous signal, and were able to take more time to consider the alternatives in addition to the sentence context, perhaps due to superior rehearsal capacities.

Hypotheses 4C and 4D:

Clear, positive correlations between temporal auditory processing and comprehension performance were not found, high and low WMS listeners showed similar performance in most respects except response time, and longer response times were found for the high WMS listeners. Therefore, Baddeley and Logie's (1999) argument that individual differences in spoken language processing may be dependent on activation and rehearsal capacity within the phonological loop is supported. Thus the high WMS listeners may be more able to generate and hold multiple interpretations of the signal, rehearsing them while considering the sentence context better than low WMS listeners (although the high WMS listeners appeared also to ultimately run out of processing time, because they made more errors on context words; or the target and context words may have been less well coupled for high WMS listeners).

3.6 Experiment 3: Shallow-processing task

Experiment 3 was designed to investigate whether reducing the stress on working memory might result in an increase in the Number of Correct Responses (target word and context word) and reduce the Med RT for the high WMS group in relation to the low WMS group. The same young adults who participated in Experiment 2 were tested in Experiment 3. The stimuli and procedure for Experiment 3 were identical to Experiment 2, except this time, rather than choosing a response consisting of the target word and Synonym for the context word, the listeners were asked to choose the target word and the actual context word that appeared in the sentence. For example, for the sentence *The kiddies bat the doggies and chase toy rockets*, *Bat* (*Chase*) would be shown on the screen rather than *Bat* (*Follow*) as a response choice (see Appendix D). This change in response choice reduced demands placed on the listeners that required them to compare the input to knowledge-based information from long-term memory during processing.

3.7 Method

3.7.1 Participants

In the next experiment the same 32 young adults (16 High and 16 Low WMS listeners) were tested.

3.7.2 Design

The design was identical to Experiment 2.

3.7.3 Stimuli

The stimuli were identical to Experiment 2.

3.7.4 Procedure

The procedure was identical to Experiment 2 except the responses included the target word and the actual final verb of each sentence, rather than a synonym, as shown in Appendix D.

3.7.5 Scoring

The Number of Correct target word and context word Responses, the Median Response Times for Correct Sentences, and Gap-detection thresholds were analyzed.

3.8 Results

The results from an ANOVA with Number of Correct Responses (target word and context word) as the dependent variable in Experiment 3 showed a main effect for Noise (F (3, 90) = 360.953, $p < .0001$), Sentence Context (F (2, 60) = 6.311, $p < .01$), Voicing (F (1, 30) = 6.101, $p < .05$), and significant interactions for Noise x Context (F (6, 180) = 3.496, $p < .01$), Context x Voicing (F (2, 60) = 11.015, $p < .0001$), and Noise x Context x Voicing (F (6, 180) = 14.703, $p < .0001$)[5]. Although the WMS x Noise interaction was not significant in Experiment 3 as it was in

[5] Only responses with less than 5 second latencies were included in the data analyses, because it was decided a priori that any responses more than 3 standard deviations above the mean would be excluded. According to this

Experiment 2, planned protected t-tests were conducted to evaluate group differences in high noise.

As in Experiment 2, planned protected t-tests showed that both high and low WMS listeners showed significantly more Correct Responses in quiet than in high noise [$\underline{F}$ (1, 180) = 4092.57, $\underline{p} < .001$), and $\underline{F}$ (1, 180) = 3552.34, $\underline{p} < .001$], and again, the low WMS listeners showed more Correct Responses than the high WMS listeners in the high noise condition [($\underline{F}$ (1, 180) = 19.59, $\underline{p} < .01$; see Figure 16)].

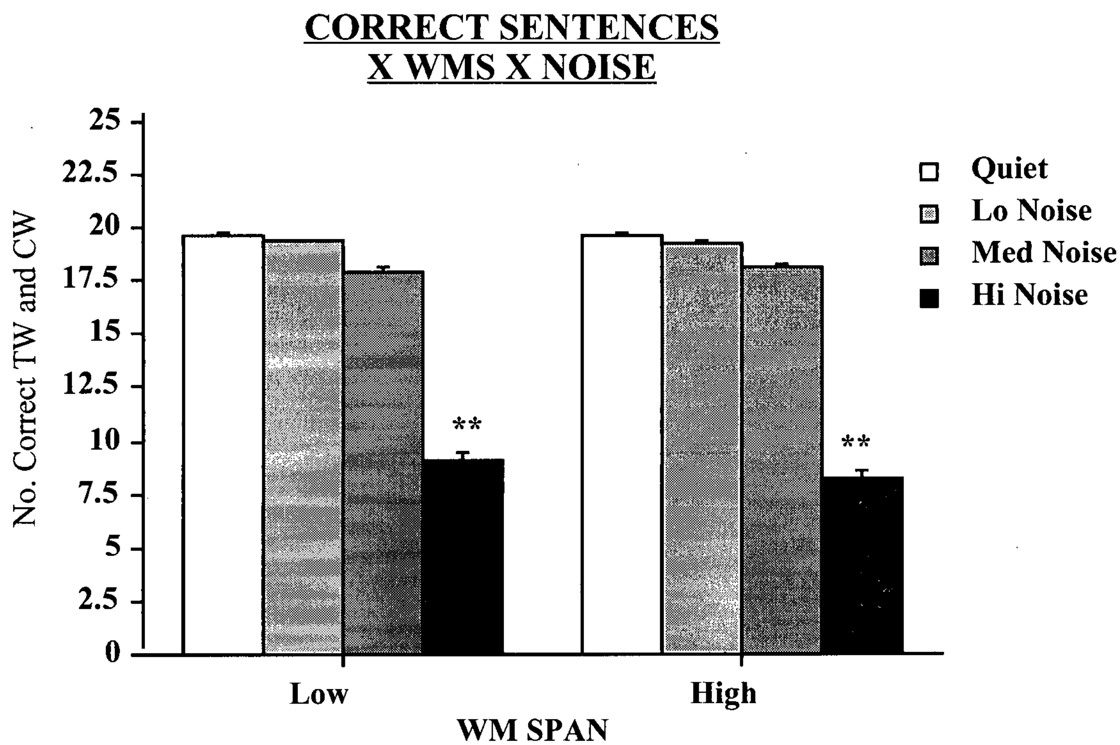


Figure 16: Number of Correct target words and context words (combined) by WMS and noise condition in Experiment 3.

---

criterion, the cut-off points were 5.1 seconds and 3.5 seconds for the high and low WMS listeners, respectively. The data from the first block in each list were included in all of the data analyses for Experiment 3, since a A mixed 2 (WMS) Between-subject X 4 (Noise) X 5 (Block) Within-subject analysis of variance (ANOVA) with RT as the dependent variable showed a significant interaction between Noise and Block (p < .001), however, follow-up Newman-keuls tests did not show that RTs were significantly longer for Block 1 than Blocks 2-5 in the quiet condition for all listeners (p > .05). Thus, the total number of correct responses for each sentence type per list was 20 (4 sentences of each type x 5 blocks).

Planned protected t-tests were conducted to evaluate the influence of sentence context on Number of Correct Responses in high noise with no difference found between Context-Consistent and Context-Inconsistent for high WMS listeners ($p > .05$). Low WMS listeners responded correctly to Context-Inconsistent significantly more than Context-Consistent sentences, (F (1, 180) = 6.399, $p < .05$) again, contrary to the predicted direction (see Figure 17).

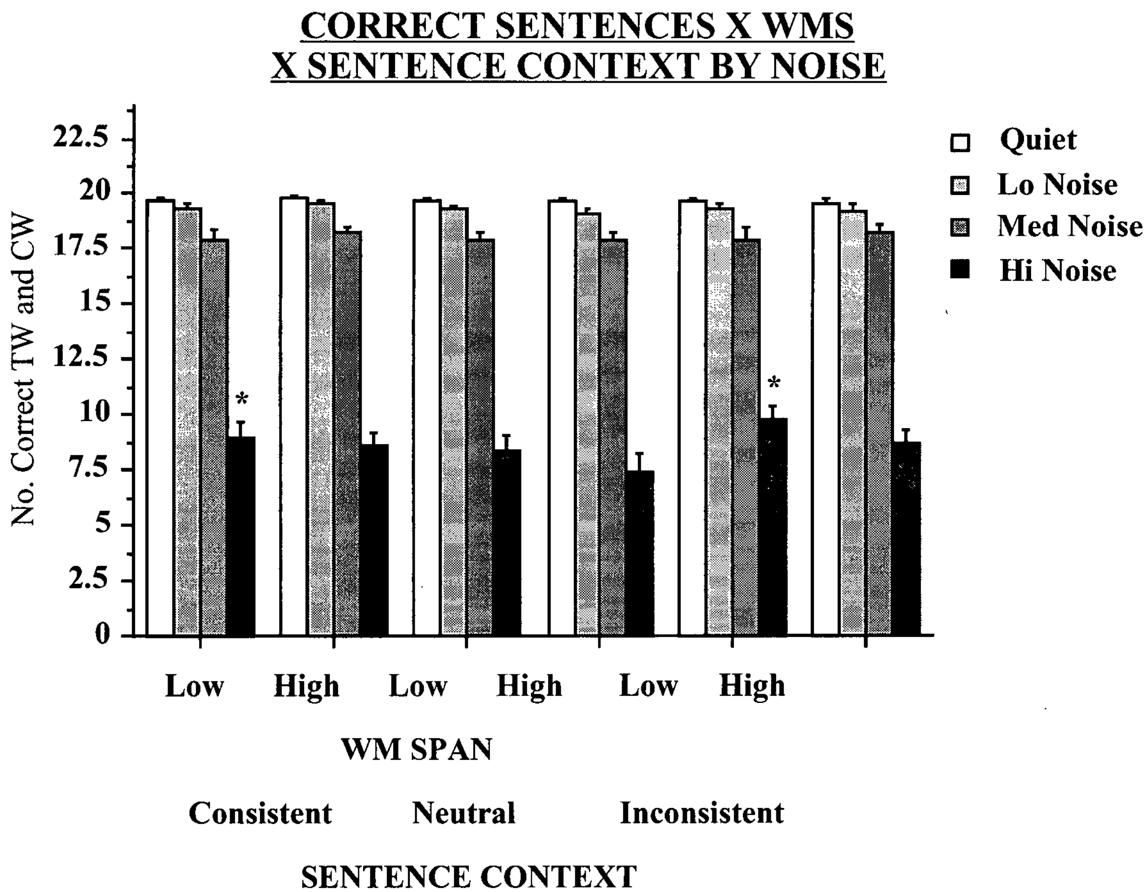


Figure 17: Number of Correct target words and context words (combined) by WMS, sentence context and noise condition in Experiment 3.

The next question tested is whether identical results to Experiment 2 will also be found in an ANOVA with Med RT as the dependent variable. As in Experiment 2, main effects for Noise (F (3, 90) = 38.011, $p < .0001$) and Sentence Context (F (2, 60) = 10.529, $p < .0001$) were found. Unlike Experiment 2, a significant interaction for Noise x Context was found (F (6, 180) = 2.194, $p < .05$).

The critical question, again, was whether WMS differences in Med RT were shown in high noise. As in Experiment 2, planned protected t-tests showed both high and low WMS listeners took longer to respond to sentences in high noise (ambiguous) than quiet (unambiguous) [F (1, 180) = 308.19, $p < .001$ and F (1, 180) = 142.25, $p < .001$, respectively]. Also, as in Experiment 2, the high WMS group showed longer response latencies in the high noise condition than the low WMS group [(F (1, 180) = 57.35, $p < .01$; see Figure 18)].

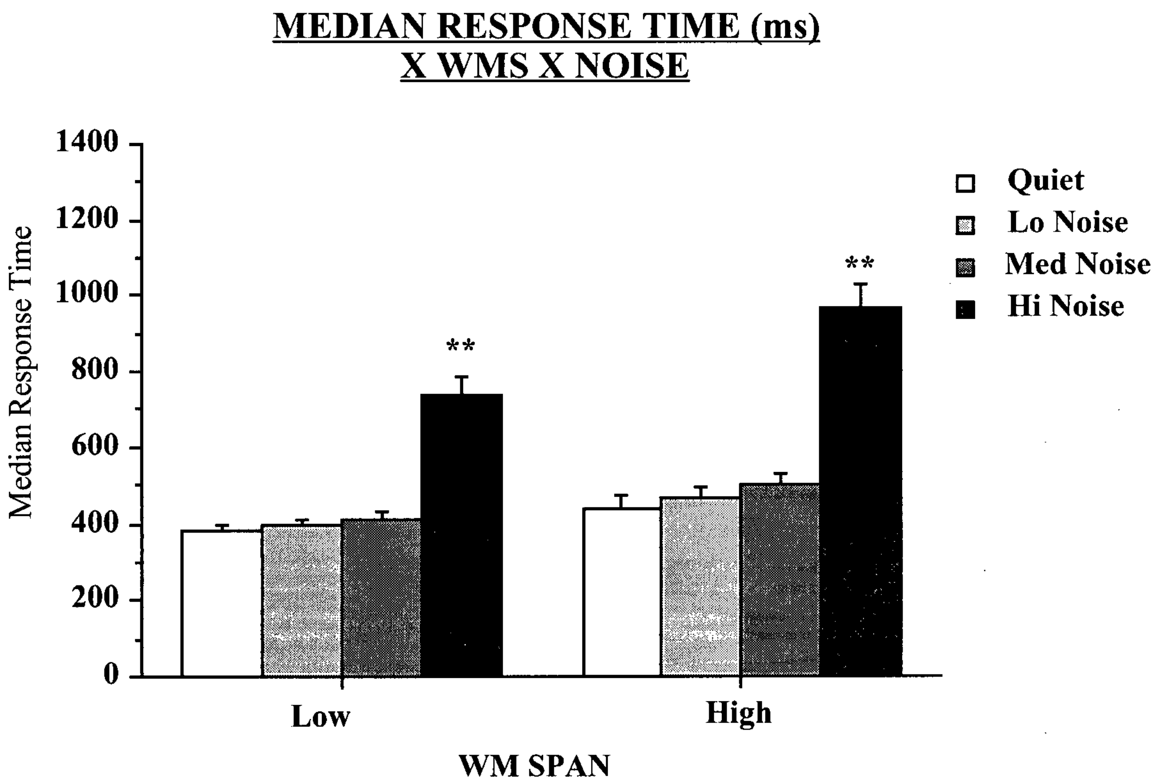


Figure 18: Median response time (ms) by WMS and noise condition in Experiment 3.

When planned protected t-tests were conducted to examine the influence of sentence context on response latencies, the low WMS (F (1, 180) = 7.248, $p < .01$), but not high WMS listeners ($p > .05$) showed longer latencies to Context-Inconsistent than Context-Consistent context sentences (see Figure 19).

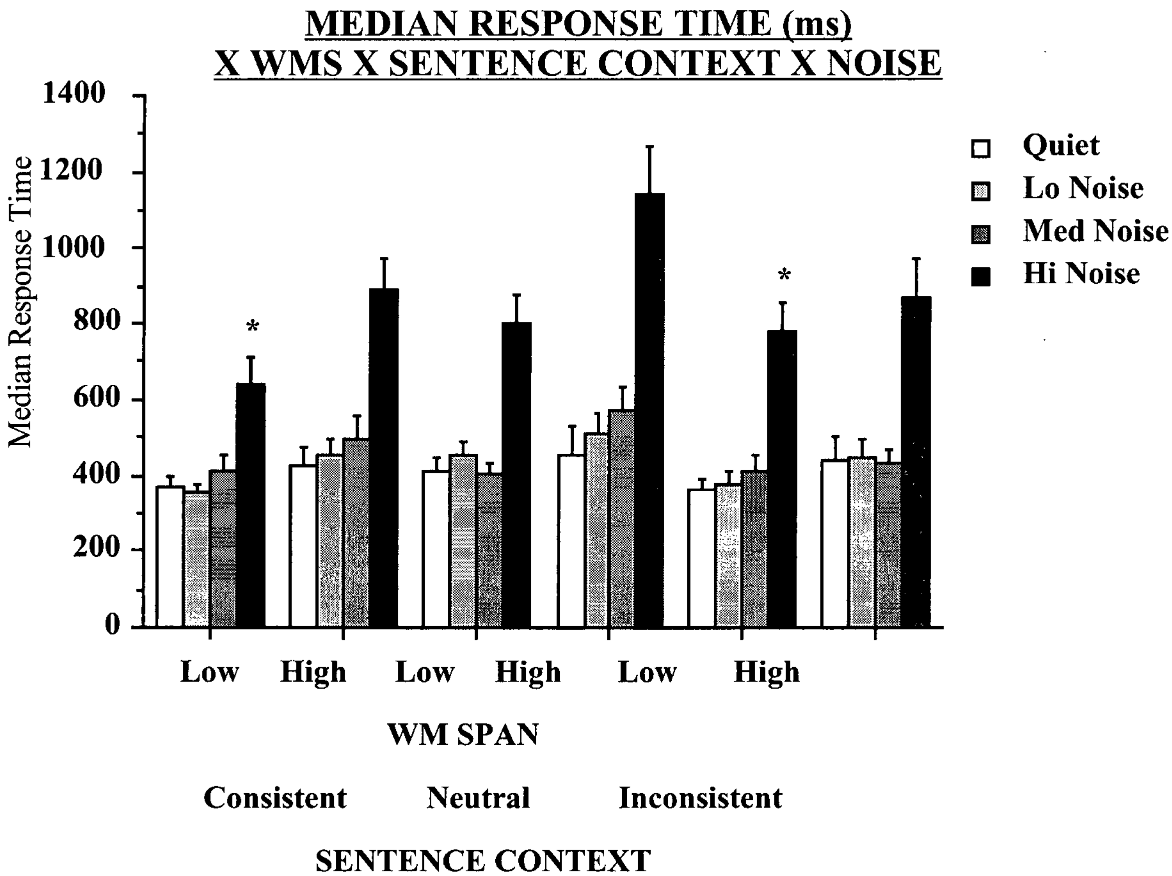


Figure 19: Median response time (ms) by WMS, sentence context, and noise condition in Experiment 3.

When Number of Correct Responses to target word and context word were analyzed separately, no WMS differences were shown [F (1, 30) = .369, p > .05; and F (1, 30) = .080, p > .05, respectively]. This indicates that even when high and low WMS listeners attended to the target word and context word equivalently, longer Med RTs by the high WMS listeners to correct sentences were still found.

Thus, the results of Experiment 3 replicate those of Experiment 2 almost perfectly. The hypothesis that high WMS listeners differ in their Med RTs due to the generation of multiple alternatives is reinforced. The finding that sentence context influenced the low WMS group in high noise indicates that sentence context may interact with response choices to ambiguous signals at times, and, as with word-frequency, low WMS listeners were more easily influenced by contextual cues.

As in Experiment 2, in order to test whether Gap thresholds correlated with comprehension performance, a Pearson Product-moment correlational analysis was conducted between Gap threshold and Med RTs to Correct Responses in high noise, and Number of Correct Responses in high noise in Experiment 3.

The correlation between Gap detection threshold (from third trial from left ear) and Med RTs on Correct Responses in high noise in Experiment 3 was not significant ($p > .05$), nor was the correlation between Gap threshold and Number of Correct Responses in high noise in Experiment 3 ($p > .05$).

The above findings indicate that, insofar as the gap detection test measures temporal auditory processing relevant to the comprehension of spoken sentences, temporal auditory processing does not appear to be related to differences in spoken language comprehension between high and low WMS listeners in the conditions tested.

To further test the possible effects of Gap detection ability on performance, the same analyses of variance conducted in Experiments 2 and 3 were performed, except Gap detection threshold (High or Low), rather than WMS, was entered as a between-group factor.

The results did not show a clear pattern such as those found in the WMS analyses. For example, the Low Gap group had significantly more Correct Responses (target word and context word) than the High Gap group in Experiment 2, but the High Gap group had significantly more Correct Responses in Experiment 3. There was no difference between Gap groups for Med RT to Correct Responses in Experiment 2, but the High Gap group took significantly longer to respond in high noise in Experiment 3. The lack of a consistent, coherent pattern in the analyses comparing High Gap to Low Gap groups suggests that either a) gap-detection thresholds are not a relevant measure for perception of complex signals, b) the groups were too homogeneous in temporal auditory processing (gap detection thresholds) to show a difference, c) temporal auditory processing is independent from working memory processing, or d) the speech presented was slow enough that temporal auditory processing was not an issue.

With respect to b), an independent t-test that was conducted between the High and Low Gap groups (on the third trial of the left ear) showed a significant difference between groups (t (15) = 11.17, $p < .001$), ruling out the possibility that gap scores were too homogeneous among the listeners to show a difference. According to c), irrespective of temporal auditory processing, high WMS individuals appear, nonetheless, to generate more alternatives to an ambiguous signal than the low WMS listeners. Insofar as gap detection thresholds are an appropriate measure of this aspect of auditory processing, the results of Experiment 2 and 3 tend to bear out this third explanation.

Thus, high WMS listeners may generate more alternatives in ambiguous conditions. Although WMS differences do appear when the auditory perceptual system is stressed, it is still unclear whether these differential results can be explained by differences in temporal auditory processing. An alternative explanation for the results is that a greater number of target word competitors were activated in the mental lexicon for high WMS listeners, and longer processing time was required for them to compare the input to the word units in the mental lexicon. They may also have been considering multiple context words for each sentence.

In Experiment 1, however, when the ambiguous target words held world knowledge at the lexical level, lexical knowledge appears to have influenced the low WMS more than the high WMS listeners' interpretation of the target word. In Experiments 2 and 3, when the ambiguous target words were equal in frequency, and world-knowledge was varied at the sentence level, the low WMS listeners appeared to be released from the direct influence of world-knowledge and interpreted the sentences more quickly (and in some cases even more accurately) than the high WMS listeners. This pattern of results may also have resulted from the nature of the task in Experiments 2 and 3, which was a two-alternative forced-choice task for the target words, whereas multiple target word responses were available in Experiment 1. Thus, the low WMS listeners may have generated fewer alternatives to the ambiguous target words in Experiments 2 and 3 than

Experiment 1, resulting in more success identifying the correct target words and context words (although they were still more influenced by Context-Consistent vs Context-Inconsistent in high noise than the high WMS listeners in Experiment 3).

3.9 Discussion

Overall, the results of Experiment 3 replicate those of Experiment 2, except that the low WMS listeners took longer to respond to Context-Inconsistent than Context-Consistent sentences in Experiment 3.

Hypothesis 5:

Recall that general working memory resources were taxed in Experiment 2 with a deep processing task, whereas the shallow processing task used in Experiment 3 did not tax general working memory resources as much. As stated previously, an interaction between WMS and sentence context found in Experiment 2 but not Experiment 3 would support Carpenter et al.'s (1995) resource model that states processing and storage share a common pool of resources. If an interaction between WMS and Sentence Context was not found, and the pattern of results was similar between Experiments 2 and 3, Baddeley and Logie's (1999) model, that states processing and storage use separate resources, would be supported. Because a WMS x Sentence Context interaction was found in the shallow (Experiment 3), but not the deep (Experiment 2) processing task, the results support the latter better than the former explanation for individual differences in spoken language processing because they indicate that processing of knowledge in long-term memory at the sentence level does not directly interact with phonological processing.

When gap-detection thresholds of the two groups were compared to accuracy and response time measures in Experiments 2 and 3, the only significant correlation was found for response time to Correct Responses in Experiment 3 (i.e., the better the gap threshold, the longer the response time). This pattern of results does not support the hypothesis that temporal auditory processing ability underlies WMS group differences in young, normal-hearing adult listeners.

An alternative explanation, that high WMS listeners generate a greater number of alternative responses to ambiguous signals, thereby requiring more processing time and an expanded capacity for matching the input to word units in the mental lexicon may be a more plausible explanation for the results. (Such a strategy, however, may result in more context word errors for the high WMS listeners, as found in Experiment 2).

Nevertheless, the results from the three experiments do indicate that the level of analysis of the linguistic input does influence spoken language comprehension differently for high and low WMS listeners in adverse conditions. The low WMS listeners appear more influenced by world-knowledge cues, especially those at the lexical rather than the sentence level.

IV EXPERIMENT 4

4.1 Evaluation vs generation

The fourth experiment was designed to gather more direct evidence that low WMS listeners generate fewer alternatives to ambiguous words than do high WMS listeners. By revisiting whether the influence of word-frequency on the processing of ambiguous words is a result of the high WMS listeners generating more alternatives or evaluating the choices more successfully than the low WMS listeners, a better understanding will be realized regarding whether the Word-frequency x WMS interaction found in Experiment 1 can be explained by low WMS listeners' relative lack of inhibitory mechanisms (Hasher & Zacks, 1988) or lower activation thresholds (Carpenter, Miyake, & Just 1995; and Just & Carpenter, 1987, 1992).

Thirty-two young adults (16 High and 16 Low WMS listeners) were tested on their ability to correctly identify naturally-produced monosyllabic words which varied in word frequency in a classic gating task (e.g., Grosjean, 1980; Marslen-Wilson, 1990). Gating tasks have been useful in illuminating lexical access processes including:

1. The amount of acoustic-phonetic information needed to identify a stimulus, such as a syllable, a word, a group of words, etc.
2. The role played by phonetic, lexical and contextual variables during identification.
3. The underlying processes leading to identification.
4. The nature of lexical representations.

(Grosjean, 1996, p. 597)

Each of 24 words [12 high-frequency words with low-frequency competitors [e.g., dog, dock (High-Low condition)] and 12 low-frequency words with high-frequency competitors [e.g., robe, road (Low-High condition)] was presented diotically through headphones to the listener starting with the first 50 ms of the target word and increasing in 50 ms increments until the word was completely heard (see Table 2). After each word segment was played, the listeners verbally guessed the identity of the word and gave their confidence rating on a 0 to 10 scale (0 representing low confidence, or a wild guess, 5 representing 50% confidence, and 10 representing high confidence, or absolute certainty that they had heard the word correctly). The increment at which the listeners correctly identified the target word (and continued to state the correct word until the final gate) is referred to as the Isolation Point. The average Isolation Points for each group of listeners were compared to the Alignment Points, which were set at vowel offset for each of the monosyllabic word pairings.

4.2 Method

4.2.1 Participants

Thirty-two (16 high and 16 low WMS) university students between the ages of 18 and 35 years (with a mean age of 25.5 years and a standard deviation of 4.9 years) participated in this study. Participants were recruited through newspaper and poster advertisements distributed in and around the University of British Columbia. All participants were native English speakers, and participants received a token payment for their participation.

As in the previous experiments, participants were assigned to a high or low WMS group depending on their WMS scores. The mean ages (with standard deviations in parentheses) for the low and high WMS groups were 23.7 years (4.3) and 25.4 years (5.5), respectively. The mean WMS scores (with standard deviations in parentheses) for the low and high WMS groups were 2.5 (.24) and 3.5 (.34), respectively.

Listeners from each WMS group were scored within the same criteria on peripheral auditory processing abilities, word discrimination and speech recognition thresholds, hearing history, years of education, and vocabulary.

To ensure that peripheral auditory processing abilities between participants in the low and high WMS groups were similar, hearing sensitivity of all participants was tested and only those with normal audiograms for their age-group were included in the study. All participants had pure-tone air-conduction thresholds below 25 dBHL in both ears for tones from .25 to 8 kHz.

All participants scored higher than 92% on the word discrimination test, and the speech recognition threshold was between -10 and +10 dBHL. In addition, all participants had normal hearing histories based on responses to a hearing history questionnaire.

The number of years of education at time of testing was also equivalent for the high and low WMS groups, and ranged from 13 to 24 years for all participants, with an average of 17 for the low WMS group and 18 for the high WMS group. The groups were also equivalent in their scores on the Mill Hill vocabulary test, with scores ranging from 12/20 to 18/20. The average scores for the low and high WMS groups were 14/20 and 15/20, respectively. Finally, the temporal auditory processing of each individual was tested using the gap-detection test.

4.2.2 Design

A 2 (WMS) x 2 (Word Frequency) Between-Within design was implemented. Each of the 24 High-Low and Low-High target words was presented in quiet to the listener in "gates" that increased by 50 ms. A list (List 1; see Table 3) of target words was constructed such that (1) not more than two High-Low or Low-High frequency words were presented in a row, and (2) each target word was followed by a word with a different consonant sound at the beginning and end. To control for potential order effects, a second list (List 2; see Table 3) was constructed by simply reversing the order of the words in the first list. Lists 1 and 2 were presented to the high and low WMS listeners in counter-balanced order.

4.2.3 Stimuli

Two word sets were constructed that balanced the factors of Frequency of the Target word and the Competitor in the manner illustrated in Table 2, below.

Table 2

Organization of the stimulus set

	Frequency	
Pair Type	Target Word	Competitor
High-Low	High	Low
Low-High	Low	High (and Low)

Specifically, two sets of 12 target word pairs each were constructed. Half of the listeners heard List 1 and the other half heard List 2. The target words were spoken naturally by an adult female talker and digitally recorded (see Table 3, below, for the target words).

Table 3

Word lists

List 1		List 2	
Guide	(H)*	Clash	(L)
Space	(H)	Dog	(H)
Gaze	(L)	Brush	(H)
Mouse	(L)	Robe	(L)
Great	(H)	Jazz	(H)
Fig	(L)	Club	(H)
Streak	(L)	Bead	(L)
Cloud	(H)	Fret	(L)
Debt	(L)	Spite	(H)
Quick	(H)	Fad	(L)
Give	(H)	Book	(H)
Brisk	(L)	Strive	(L)
Strive	(L)	Brisk	(L)
Book	(H)	Give	(H)
Fad	(L)	Quick	(H)
Spite	(H)	Debt	(L)
Fret	(L)	Cloud	(H)
Bead	(L)	Streak	(L)
Club	(H)	Fig	(L)
Jazz	(H)	Great	(H)
Robe	(L)	Mouse	(L)
Brush	(H)	Gaze	(L)
Dog	(H)	Space	(H)
Clash	(L)	Guide	(H)

* H = High-Low; L = Low-High Target Words

The words were chosen from the Kucera and Francis' (1982) corpus and were controlled in the following fashion: All of the words were monosyllabic, and started with a consonant or consonant cluster which was followed by a vowel and ended with a consonant or consonant cluster, all of the words were root words including verbs, nouns, or adjectives and proper nouns were not included as target words nor were target words which had competitors that were proper nouns.

The pool of potential competitors was restricted to other monosyllabic root words beginning with the same initial consonants or cluster and vowel as the target word, but ending in a different consonant sound. In the rare case in which a competitor was not listed in the corpus in its root form, the listed word classification was included as a competitor (e.g., clamps, clasped were included as competitors for *clash*). Words in the Low-High set had high and low-frequency competitors, whereas words in the High-Low set had only low-frequency competitors. To control for the phonetic relatedness of competitors, words were chosen with at least two featural differences in manner and place of articulation between the last consonant of the target word and that of the nearest competitor, and, where possible, the two features differed in voicing contrast.

Target words in the two sets were matched for uniqueness points and for alignment points (see Marslen-Wilson, 1990, p. 153), which are two aspects of spoken words that it is essential to control for. The point in the word at which it becomes uniquely identifiable, in the sense of separating from the other members of its word-initial cohort is referred to as the uniqueness point (Marslen-Wilson & Welsh, 1978; Grosjean, 1980). Variations in uniqueness points may be confounded with the experimental variable, unless this point is held constant across sets. The alignment point is critical for tasks involving timed responses. Response times (RTs) to spoken words need to be measured relative to equivalent points in the different words. The timing of the responses will be affected by possibly irrelevant differences in word-length if RTs are measured from word onset. Therefore, RTs need to be measured from a point that is informationally equivalent across different words. The challenge is to determine the equivalent points in different words. For example, where are the equivalent measurement points in words like <u>steep</u> and <u>stone</u>, which differ quite markedly in the temporal distribution of information about their final consonants (Warren & Marslen-Wilson 1987, 1988)?

The problem of uniqueness points for these sets was solved by using only monosyllabic words, all of which terminated in a consonant, and became discriminable only when this final consonant was heard. The problem of an alignment point was solved by matching the phonetic properties of the last segment across stimulus sets. Each set contained 4 words terminating in voiced plosives (/b/, /d/, /g/), 4 terminating in unvoiced plosives (/p/, /t/, /k/), 2 in voiced fricatives (/z/, /v/), and 2 in unvoiced fricatives (/s/, /sh/). This meant that the alignment point (relative to which response time was measured) could be controlled across sets. For example, for any High-

Low stimulus terminating in an unvoiced stop, there was a corresponding stimulus from the Low-High set also terminating in an unvoiced stop. The alignment point was set at the vowel offset for each word, which assured informational equivalence across sets (see Appendix F). Vowel offset was set at the final period in the vowel in all words except in those which ended with voiced or voiceless fricative sounds, in which case the vowel offset was set at the beginning of the frication. Two coders measured the vowel offset for all words, with an average difference of .433 ms (SD = .506 ms, ranging from .05 to 1.4 ms) between their measurements. The alignment point was set at 50 ms before the vowel offset (Marslen-Wilson, 1990, p. 162).

The average number of 50 ms segments per word was 18 (SD = 2.9) for the High-Low set, and 18 (SD = 2.2) for the Low-High set. The average duration of the full word was 862.5 (SD = 145.1) for the High-Low set, and 876.6 (SD = 112.2) for the Low-High set. The average RMS volts was .57 (SD = .15) for the High-Low set and .53 (SD = .11) for the Low-High set. The average alignment point was 466.8 ms (SD = 103.0) for the High-Low set and 466.5 ms (SD = 103.1) for the Low-High set.

Additional characteristics of the target words and their competitors were:

High-Low target words:

1) The mean number of all competitors for the High-Low target words was 3.8 (SD = 2.0) with a range of 2 competitors (*brush, cloud*) to 8 competitors (*great*).

2) High-Low target words had frequencies that ranged from 28 (*cloud*) to 665 (*great*) with a M of 172.75 (SD = 186.09; see Figure 20, below),

3) The difference in frequency between the High-Low target words and their nearest competitors was required to be at least 30 (the closest being *guide* with a frequency of 36 and *guise* with a frequency of 6). One exception to this restriction was the target word *cloud*, with a frequency of 28 and whose nearest competitor, *clown*, had a frequency of 3 (difference is 25),

4) Only words with a frequency ≤ 13 (for both target words and competitors) were considered to be "low-frequency",

5) Although the nearest competitor for High-Low target words was required to be less than 35, to ensure that only truly low-frequency competitors were included in the coding and analyses, only words with a frequency ≤ 13 were considered to be "low-frequency competitors",

6) High-Low target words had an average of 3 low-frequency competitors (SD = 1.13) that ranged from 2 competitors (*brush*) to 6 competitors (*quick*)). The M frequency of the highest of the low-frequency competitors was 6.25 (SD = 4.14; see Figure 20, below) with a range of 1 (*brush*) to 13 (*book*),

Low-High target words:

7) The mean number of all competitors for the Low-High target words was 5.8 (SD = 2.66) with a range of 2 competitors (*fret*) to 10 competitors (*robe*).

8) Low-High target words had frequencies that ranged from 1 (*fret*) to 13 (*debt*) with a M of 6.67 (SD = 3.96; see Figure 20, below),

9) Low-High target words were required to have at least one high-frequency competitor above a frequency of 50. Only the two highest competitors were considered to be "high-frequency competitors" for scoring purposes. The average of all of the two "high-frequency competitors" to the Low-High target words was 137.05 (SD = 105.35; see Figure 20, below). Their frequencies ranged from 16 to 447.

10) The Low-High target words had an average of 2.59 low-frequency competitors (SD = 1.68) which ranged from 0 competitors (fret) to 6 competitors (*clash*). The average frequency of the highest of the low-frequency competitors was 8.36 (SD = 4.32) with a range from 0 (*fret*) to 13 (*streak, fad*).

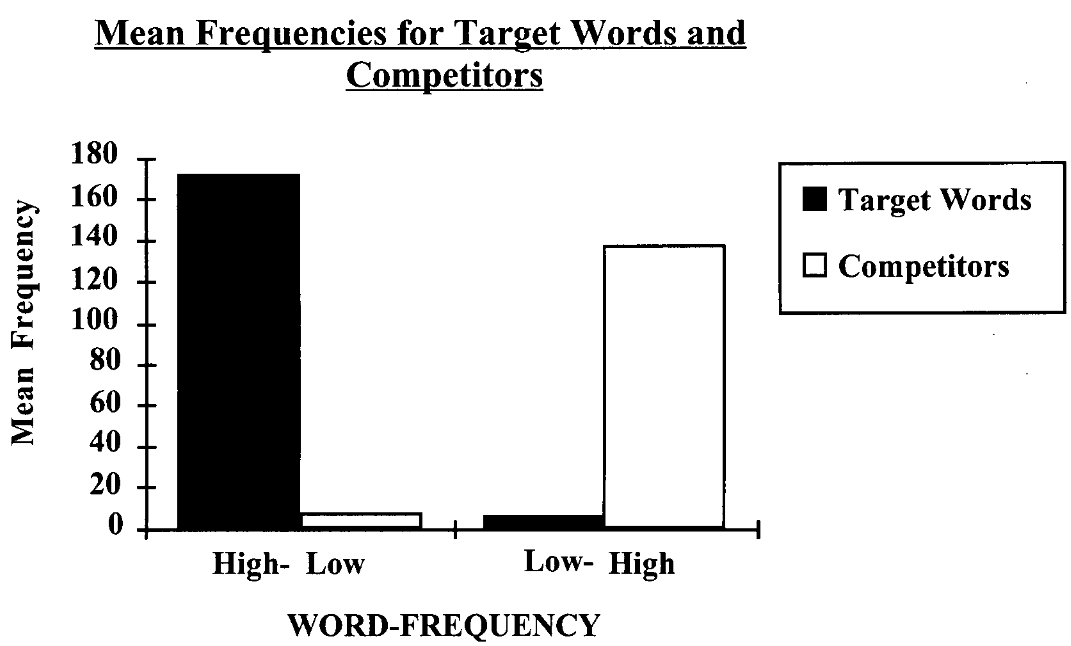


Figure 20: Mean Frequencies for Target Words and their Competitors presented in the gating task in Experiment 4.

4.2.4 Procedure

Using the Computerized Speech Research Environment (CSRE 4.5), the segments of digitized target words were presented diotically through headphones to the participant who sat in a sound-attenuated booth. The participant was instructed to listen to each segment and then verbally "guess" the identity of the word as well as indicating their confidence by choosing a number between 0 (low) and 10 (high). The verbal responses were transmitted via microphone to a research assistant who sat outside the booth and recorded them in written form on a response sheet. In addition, the verbal responses were recorded on audio cassette tape for reliability coding at a later date.

4.2.5 Scoring

The following results were scored:

1) Number of alternatives (incorrect and incorrect) generated to Low-High and High-Low target words by WMS [2 (WMS) x 2 (Word-Frequency) mixed ANOVA],

2) Number of high-frequency competitors generated to Low-High target words by WMS [2 (WMS) 1-way ANOVA],

3) Number of low-frequency competitors generated to Low-High and High-Low target words by WMS [2 (WMS) x 2 (Word-Frequency) mixed ANOVA],

4) Isolation Point of Low-High and High-Low target words occurred by WMS [2 (WMS) x 2 (Word-Frequency) mixed ANOVA]. The Isolation Point is the difference between the length of the segment at which identification of the target word was reported (without subsequent changes in identification) and the Alignment Point, irrespective of confidence rating (see Marslen-Wilson, 1990, p. 157, for advantage of disregarding confidence ratings).

5) Number of initial 50 ms gates to which the listener claimed she did not identify any word (missed) for High-Low and Low-High target words by WMS [2 (WMS) x 2 (Word-Frequency) mixed ANOVA],

6) Number of trials required before the initial consonant of the High-Low and Low-High target words was identified x WMS [2 (WMS) x 2 (Word-Frequency) mixed ANOVA],

7) Correlation analyses between acoustic measures (gap detection thresholds and point at which first phoneme in each word was identified, number of first trials that were missed, number of alternatives generated overall, or point at which target word was isolated).

4.3 Results

The number of alternatives generated to High-Low target words was significantly higher for the high than low WMS listeners, but not to the Low-High target words. The number of high-frequency competitors generated to Low-High target words was significantly higher for the high than the low WMS listeners. The number of low-frequency competitors generated to High-Low and Low-High target words was not significantly higher for the high than the low WMS listeners. The high WMS listeners isolated the target word significantly <u>sooner</u> in the gating task than low WMS listeners in the Low-High condition, but not in the High-Low condition. No significant relationships were found between the measured temporal auditory processing and performance in the gating task.

The number of alternatives generated to Low-High and High-Low target words by WMS were compared in a 2 (WMS) x 2 (Word-Frequency) mixed ANOVA. The analysis revealed a significant interaction between Word-Frequency and WMS ($F(1,30) = 4.312$, $p < .05$). Follow-up t-tests showed that the number of alternatives generated to High-Low target words was significantly higher for the high than low WMS listeners ($F(1, 30) = 7.05$, $p < .05$), but not to the Low-High target words ($p > .05$); see Figure 21.

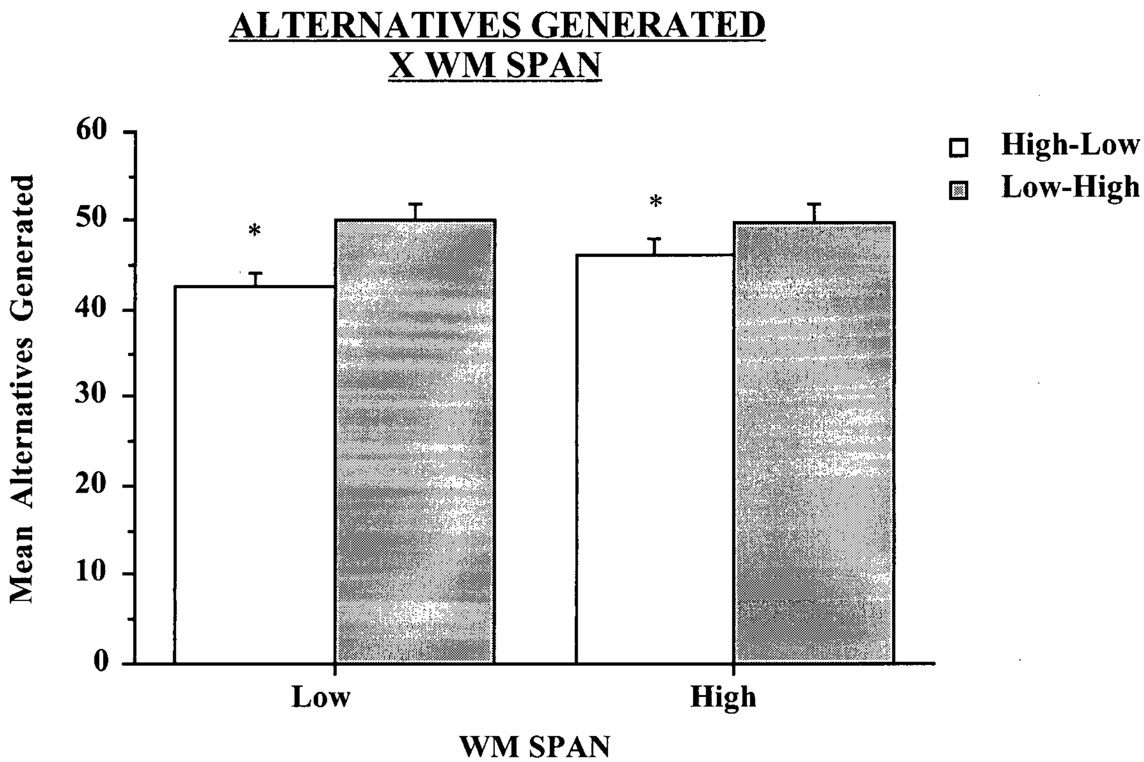


Figure 21: Number of alternatives generated to High-Low and Low-High target words by high and low WMS listeners.

The number of high-frequency competitors generated to Low-High target words by WMS was analyzed with a 2 (WMS) 1-way ANOVA. The number of high-frequency competitors generated to Low-High target words was significantly higher for the high than low WMS listeners [(F (1,30) = 6.305, p < .05; see Figure 22)].

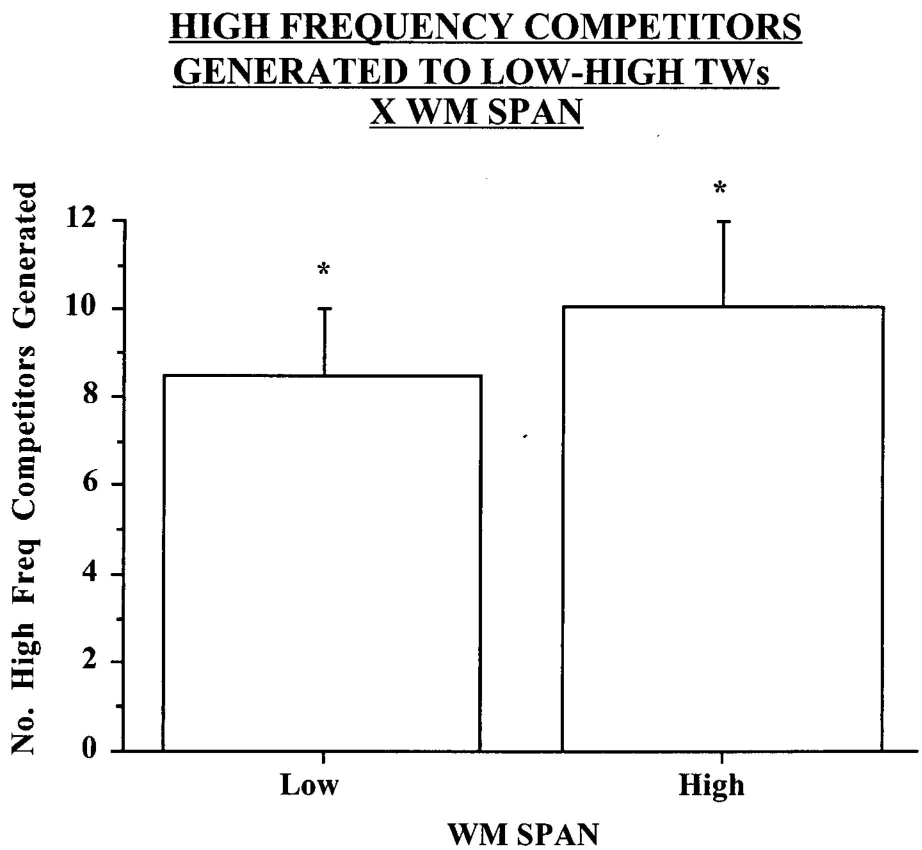


Figure 22: Number of high-frequency competitors generated to Low-High target words by high and low WMS listeners.

The number of low-frequency competitors generated to Low-High and High-Low target words by WMS was analyzed with a 2 (WMS) x 2 (Word-Frequency) mixed ANOVA. Planned protected t-tests showed that the number of low-frequency competitors generated to High-Low and Low-High target words was not significantly higher for the high than low WMS listeners ($p > .05$ in both cases; see Figure 23).

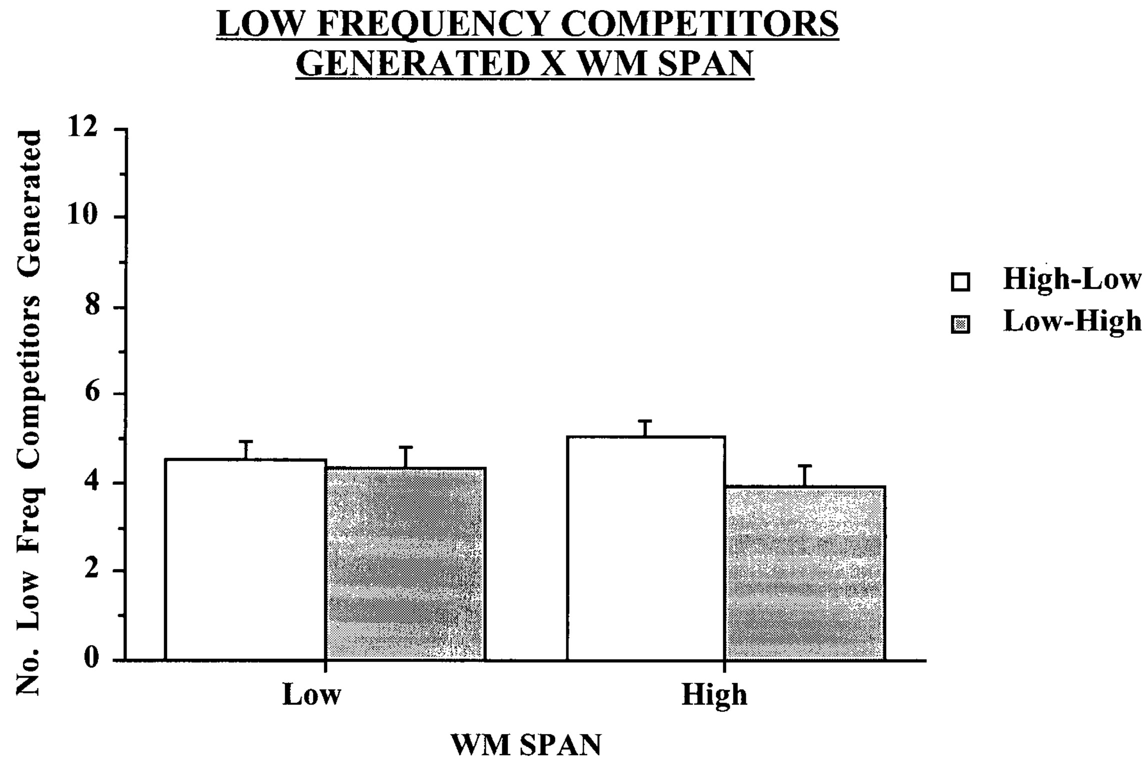


Figure 23: Number of low-frequency competitors generated to High-Low and Low-High target words by high and low WMS listeners.

To test whether the point in the gating task at which the Isolation Point of Low-High and High-Low target words differed for high and low WMS listeners, a 2 (WMS) x 2 (Word-Frequency) was conducted. Planned protected t-tests showed that the high WMS listeners isolated the target word significantly sooner in the gating task than low WMS listeners in the Low-High condition ($F\ (1,30) = 4.22$, $p < .05$), but not in the High-Low condition ($p > .05$; see Figure 24).

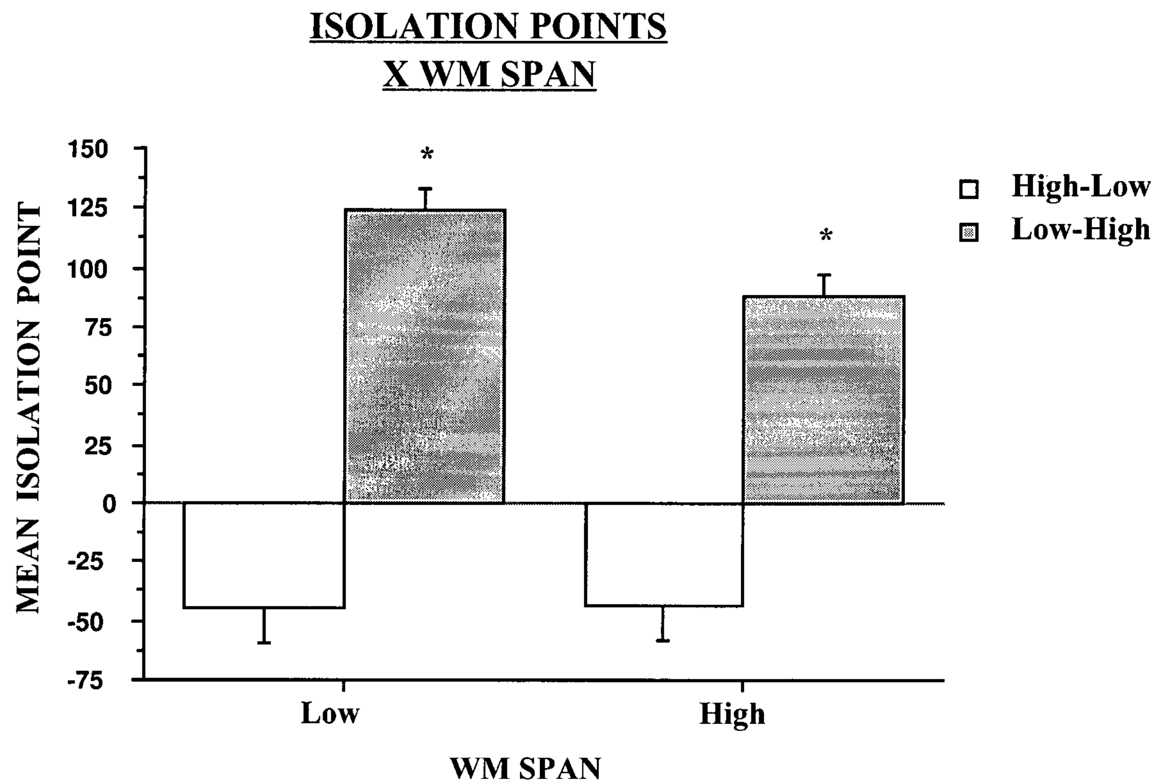


Figure 24: Mean Isolation Points for High-Low and Low-High target words by high and low WMS listeners.

Temporal auditory processing: The number of initial 50 ms gates missed for High-Low and Low-High target words by WMS was analyzed with a 2 (WMS) x 2 (Word-Frequency) mixed ANOVA. No significant differences were found between the WMS groups ($p > .05$). The number of trials required before the initial consonant of the High-Low and Low-High target words were identified by WMS were analyzed with a 2 (WMS) x 2 (Word-Frequency) mixed ANOVA, and, again, no significant differences were found. (The latter measure did not require that the listeners continued to choose the correct word to the last gate after they had chosen it once, as was a requirement of the Isolation Point measure.)

Correlational analyses between auditory measures (gap detection thresholds and point at which first phoneme in each word was identified, number of first trials that were missed, number of alternatives generated overall, or point at which target word was isolated) were conducted and no significant relationships were found ($p < .05$ in all cases).

4.4 Discussion

Hypotheses 6 and 7:

The high WMS listeners generated significantly more alternatives to the High-Low target words than the low WMS listeners. When only the high and low-frequency competitors were analyzed, the high WMS listeners generated significantly more high-frequency competitors to the Low-High target words than the low WMS listeners. The high WMS listeners did not appear, however, to generate significantly more low-frequency competitors to the High-Low or Low-High target words than the low WMS listeners, although they did isolate the Low-High (but not the High-Low) target words significantly sooner in the gating task than the low WMS listeners. Taken together, these results suggest that high WMS listeners tend to generate more alternatives in general to High-Low target words, generate more high-frequency competitors to the Low-High target words, and, although they tend to generate the same number of low-frequency competitors to High-Low and Low-High target words, for the Low-High target words they do so faster and arrive at the target word sooner than the low WMS listeners.

The results showing that high WMS listeners generate more alternatives in general to High-Low target words (see Figure 25, below), and more high-frequency competitors to the Low-High target words, lends support to Just and Carpenter's (1987; 1992) CC Reader model. The results showing that the high WMS listeners isolate the Low-High target words sooner in the gating task than the low WMS listeners, even though they have generated the same number of low-frequency competitors, lends support to Hasher and Zacks' (1988; 1997) explanation that low WMS listeners may have poorer inhibitory mechanisms. Even though the low WMS listeners generated significantly fewer high-frequency target words to the Low-High target words (see Figure 26, below for pattern of results), the fact that they still took longer to isolate the Low-High target words suggests that their poorer inhibition of inappropriate alternatives may interfere with the generation of the correct Low-High target word. Perhaps low WMS listeners persevere in their focus of attention to the incorrect alternatives which interferes with the generation of the correct response.

PATTERN OF ALL ALTERNATIVES GENERATED TO HIGH-LOW TARGET WORDS

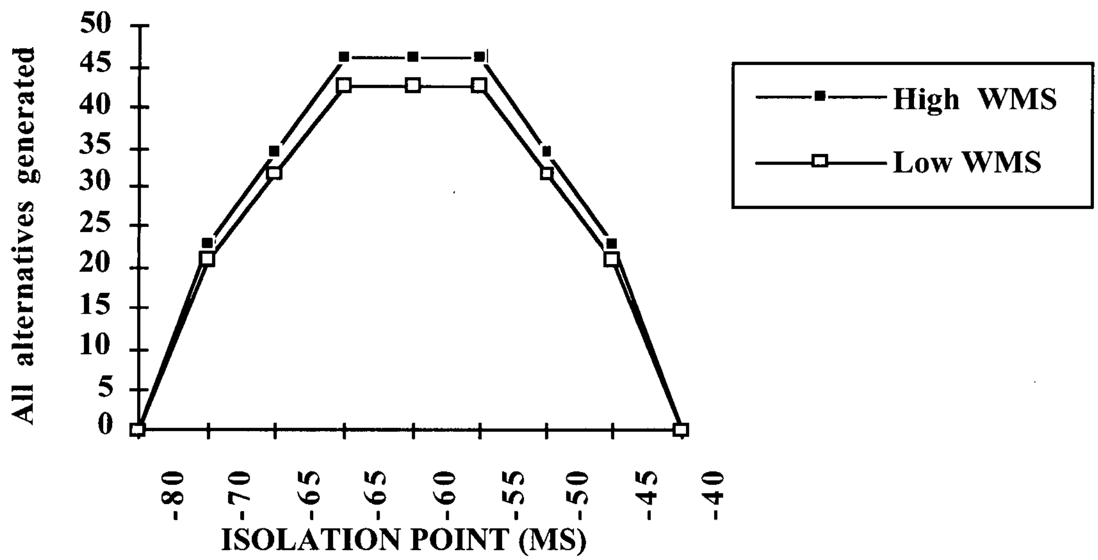


Figure 25: Pattern of results showing all competitors generated to High-Low target words and decay of incorrect alternatives by high and low WMS listeners.

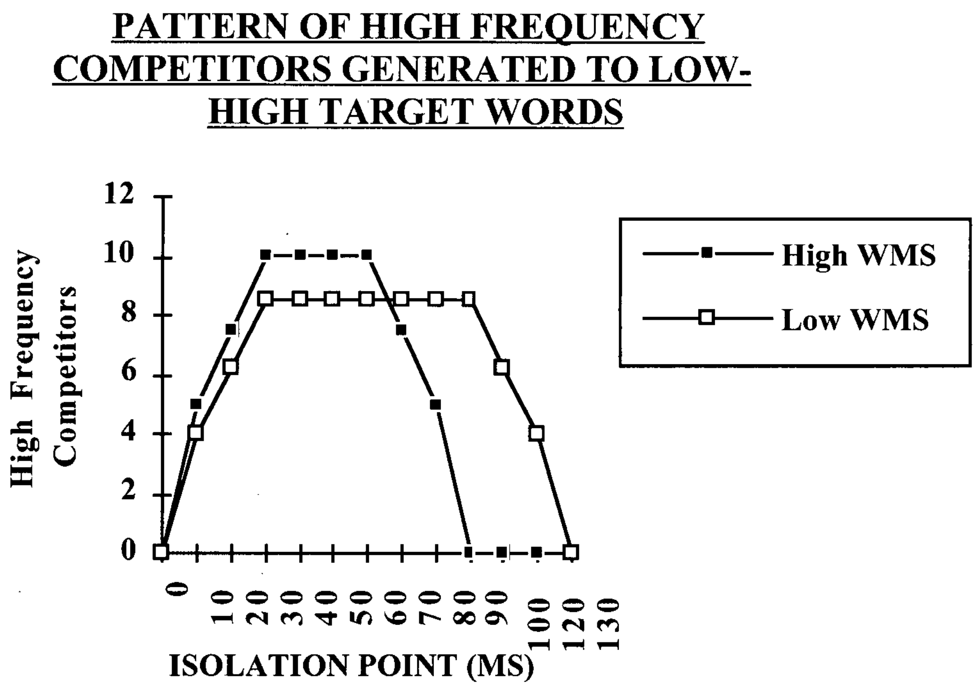


Figure 26: Pattern of results showing number of high-frequency competitors generated to Low-High target words and decay of incorrect alternatives by high and low WMS listeners.

These results also suggest that the high WMS listeners in Experiment 1 may have generated significantly more alternatives to the ambiguous target words (both high- and low-frequency) before responding, whereas the low WMS listeners may have been more influenced by the high-frequency words that were generated or made available on the response form without fully considering all possible alternatives for the target word. Thus, converging evidence is provided for the hypothesis that contextual information at the lexical level more directly interacts with phonological processing for low than high WMS listeners. Once again, temporal auditory processing appears to be unrelated to working memory in young, normal-hearing adults.

At present, it would be prudent to conclude that perceptual processing abilities do not explain individual differences in spoken language processing in a young, normal-hearing adult population. This does not preclude the possibility, however, that perceptual processing could account for spoken language processing differences in more diverse populations such as individuals with hearing loss, and/or the elderly who experience presbysusis. In future research it would be important to use methods in addition to the gap detection threshold measure to test auditory perceptual abilities. A test that measures listeners' abilities to detect a non-linguistic signal against a background of noise might be a more convincing measure of auditory processing.

V GENERAL DISCUSSION

The evidence from these studies favors the hypothesis that word-frequency cues at the lexical level interact with both phonological cues and working memory (Experiments 1 and 4), whereas contextual cues at the sentence level and temporal auditory processing do not appear to interact with working memory, nor interact as much with phonological cues as lexical level cues do in spoken language processing, although their influence cannot be ruled out (Experiments 2 and 3).

It appears that high WMS listeners are better at activating alternatives (Experiment 4), rehearsing the alternatives (Experiments 2 and 3) and that they are more efficient at inhibiting inappropriate alternatives, especially to low-frequency words (Experiment 4), but this does not appear to be related to temporal auditory processing.

Auditory processing

Because no clear correlation between gap detection thresholds and target word identification were found in Experiments 2, 3 and 4, it appears that temporal auditory processing, as measured by gap-detection, does not relate to spoken language processing in this population for these materials in these tasks, although it cannot be ruled out. Perhaps the sentences in Experiments 2 and 3 were spoken too slowly to yield significant correlations between temporal auditory processing and comprehension scores. Perhaps listeners who may have more pronounced temporal auditory processing deficits than the young, normal-hearing listeners tested in the above experiments may show positive correlations between temporal auditory processing and comprehension scores if they were tested in a task similar to that in Experiments 2 and 3, above. Such populations may include hard-of-hearing listeners (Tyler, Summerfield, Wood & Fernandes,

1982) and listeners with age-related declines in auditory processing (Pichora-Fuller, Schneider & Daneman 1995).

Phonological processing

The results of Experiment 1 clearly show that high WMS listeners are better phonological processors -- they are especially good at detecting voicing over place of articulation contrasts. In Experiments 2 and 3, the low WMS listeners identified the target words as well as or better than the high WMS listeners indicating that they perceived the phonological cues equally well (although the two-alternative forced-choice nature of the task may have constrained target word identification, making the task easier than Experiment 1 for the low WMS listeners). This difference could also mean that high WMS listeners are better at detecting more subtle cues such as differences in word-final consonants as in Experiment 1, but not necessarily clearer signals such as word-initial differences as in Experiments 2 and 3.

In Experiment 4, although there were differences between the WMS groups at the lexical level in terms of the number of alternatives generated, there were no significant group differences in phonological processing insofar as the number of initial 50 ms gates missed or the number of trials required before the initial consonant of the target words was identified. This finding is not as surprising, however, because the target word segments were presented in quiet only.

Phonological-lexical interaction

The results of Experiments 1 and 4 clearly indicate that lexical support in terms of word-frequency interacts with the processing of phonological cues significantly more for low than high WMS listeners. Although word-frequency influenced both groups' target word identification in Experiment 1, the effects were significantly greater for the low WMS listeners, suggesting that lexical information has a stronger influence for them.

The low WMS listeners may have been able to identify the target words as well as the high WMS listeners in Experiments 2 and 3 because word-frequency was controlled and the task was a two-alternative forced choice task for the target word, constraining the possible responses. In addition, it is worth noting that transitive verbs rather than concrete nouns were presented as target words in Experiments 2 and 3 and that these word classes are known to be stored and accessed differently as evidenced in pathological cases (e.g., Bird, Howard, & Franklin, 2001; Shapiro & Caramazza, 2001). Nouns are relatively independent lexical items which are free to serve a number of different syntactic functions, whereas transitive verbs are closely associated with expected noun subclasses (e.g., animate vs inanimate) which act as patients of agents. Furthermore, because the nouns used in Experiment 1 were concrete nouns, it is possible that they were imaged as independent items during lexical access, whereas it would not be possible to image a transitive verb without also imaging a subject and object. Therefore, because lexical access of the nouns in Experiment 1 was less reliant on context compared to access of the verbs in Experiments 2 and 3, a relatively high degree of the influence of context on lexical access may account for the absence of group differences in target word identification in Experiments 2 and 3.

The issue of whether the two groups access lexical items differently was explored more directly in Experiment 4. The findings show that high WMS listeners generated the same number of low-frequency competitors to High-Low target words as the low WMS listeners, but identified the High-Low target words more quickly, and, although more high-frequency competitors were generated to Low-High target words by the high WMS listeners, the Low-High target words were assessed in the same amount of time. Such results suggest that low WMS listeners activate and evaluate potential candidates in a slower manner, and that high WMS listeners do so in a faster manner.

Sentence Context

The effect of semantic context was examined in Experiments 2 and 3 by varying the consistency of the sentence meaning with the target word. The results from Experiments 2 and 3 indicate that consistency of sentence context does not directly influence the interpretation of ambiguous target words for either high or low WMS listeners. Contrary to expectation, even though there was a clear effect of word-frequency on word identification in Experiment 1, no consistent influence of sentence consistency on accuracy of target word identification emerged in either Experiment 2 or 3. This contrasts with findings that readers and listeners who differ in

WMS differ in their ability to compare the sentence context with the alternatives generated (e.g., Miyake, Just, & Carpenter, 1994; Wingfield, 1996). Another possibility is that differences in the processing of ambiguous target words may lie more in the ability of the listener to generate and/or inhibit alternatives than to the use of context. The relative roles of alternative generation, inihibition, and the influence of context no doubt depend on the specific nature of the contraints characterizing the materials and tasks, as well as on degree to which the listener is engaged in on-line vs off-line processing. Speed as well as accuracy needs to be considered to address this issue.

The findings in Experiments 2 and 3 that high WMS listeners took longer to consider the target word and context word before responding provides convergent support for Connine, Blasko and Wang's (1994) thesis that high WMS listeners may take longer to respond to ambiguous vs unambiguous speech tokens than low WMS listeners as a result of generating multiple alternatives to the target word. Such results can be reconciled with faster identification scores for high WMS listeners for the Low-High target words in Experiment 4 because of the differing demands of the experiments. Whereas Experiments 2 and 3 focused on the selection of words in sentence context, the gating task in Experiment 4 tested generation of isolated words. Therefore, compared to low WMS listeners, high WMS listeners are apparently faster at phonological processing during word generation at the lexical level but they take more time to select target words in sentence context. Experiment 4 better approximates on-line processing compared to Experiments 2 and 3 which entail a greater degree of off-line processing.

Rehearsal Capacity

The results from Experiments 2 and 3 showing longer response latencies for high WMS listeners suggest that, when there is opportunity and/or need for off-line processing, this group may be able to take advantage of better rehearsal capacity by considering multiple interpretations of the target word while considering the sentence context for a longer period of time than low WMS listeners. This provides new evidence that rehearsal capacity within the phonological loop (Baddeley & Logie, 1999) may vary depending on individual working memory capacity. This superior rehearsal capacity may also explain why, in a complex task like Experiment 1, the high

WMS listeners were able to correctly identify more target words across all conditions, interpret more sentences correctly in high noise, and interpret sentences with low-frequency words more accurately than low WMS listeners. (Note that RT was not measured in Experiment 1 so it is not possible to disentangle speed from accuracy.)

Generation vs Evaluation (Inhibition)

The results of Experiment 4 provided direct evidence that high WMS listeners both generate and evaluate alternatives to the target word differently than low WMS listeners. They generate more high-frequency competitors to Low-High target words, yet process them in the same amount of time, and they generate the same number of low-frequency competitors to High-Low target words, yet process them faster than the low WMS listeners. Therefore, evidence for superior generation (Carpenter et al., 1995) and evaluation (or better inhibition mechanisms for the inappropriate responses; Hasher & Zacks, 1988; 1997; Stoltzfus et al., 1996) for high WMS listeners is provided.

Kintsch's construction-integration model

Recall that Kintsch's construction-integration model (Kintsch, 1988; 1998; Kintsch & van Dijk, 1978) states that reading comprehension is based on the integration of text with previously-learned knowledge from long-term memory, and that inferential reasoning is an important linking component of the two sources of information. Ericsson and Kintsch (1995) thus view working memory as an activated portion of long-term memory, because information from the text is compared to related items activated in long-term memory.

The results of the four experiments presented in this paper show that phonological and lexical level information interact with working memory, but sentence context may not influence accuracy of responses. While the findings of the present experiments do not encompass discourse-level processing, they may be relevant to Kintsch's model insofar as processing demands due to perceptual-lexical and lexical ambiguities may be an underestimated aspect of uptake of text information. For example, when given the sentences *John traveled by car from the bridge to the house on the hill. A train passed under the bridge.* (Kintsch, 1998, p. 108), high

WMS listeners may generate multiple images for the term *bridge*, enabling them to more easily imagine a bridge over a railway (rather than a more common interpretation of a bridge over a river), and inhibit the incorrect interpretation more readily than low WMS listeners. WMS limitations may, therefore, explain individual differences in time to process alternative hypotheses and then retain only relevant meanings in reading tasks (e.g. Bever, Garrett, & Hurtig, 1973; Kintsch, 1988, 1998). These limitations are likely to play an even greater role in listening than in reading, especially because the quality of listening conditions are more often unfavourable.

Carpenter et al.'s CC Reader model

According to Carpenter et al.'s (1995) CC Reader model, working memory capacity is limited such that when processing and/or storage demands reach capacity, then either errors or processing time will increase. Capacity limitations are expected to be reached sooner for low than high WMS listeners. In the present experiments, processing demands were stressed at different levels (perceptual, lexical and sentence level) and group differences were examined.

In Experiment 1 the expected group differences were demonstrated by more errors on the part of low compared to the high WMS listeners, as Carpenter et al.'s (1995) model would predict. Specifically, they identified fewer ambiguous target words correctly across all conditions, interpreted fewer sentences correctly in high noise, and interpreted sentences with low-frequency words less accurately than high WMS listeners. Importantly, no group differences were observed when processing demands were reduced either because signal conditions were less noisy or word-frequency was higher. It follows from these results that both perceptual and lexical processing demands consume shared working memory capacity.

Experiments 2 and 3, however, did not reveal group differences when context was varied at the sentence level. Thus, processing of lexical level information and storage of phonological information appear to share a common pool of resources, whereas processing of sentence level and phonological information do not appear to share common resources (although caution is in order when interpreting null results).

The results of Experiment 4 support the claim that high WMS listeners generate more alternatives to target words than low WMS listeners, at least under some conditions, consistent with Just and Carpenter's (1987; 1992) CC Reader model [although the results showing that high WMS listeners generate the same number of alternatives but evaluate them faster under other conditions lends support to Hasher and Zacks' (1988; 1997) explanation that low WMS listeners may have poorer inhibitory mechanisms].

Because clear, positive correlations between temporal auditory processing and comprehension performance were not found in Experiments 2, 3 or 4, no evidence is provided for the hypothesis that temporal auditory processing plays a role within the general working memory system of Just and Carpenter's (1987, 1992) model. (Recall that temporal auditory processing scores were not measured in Experiment 1.)

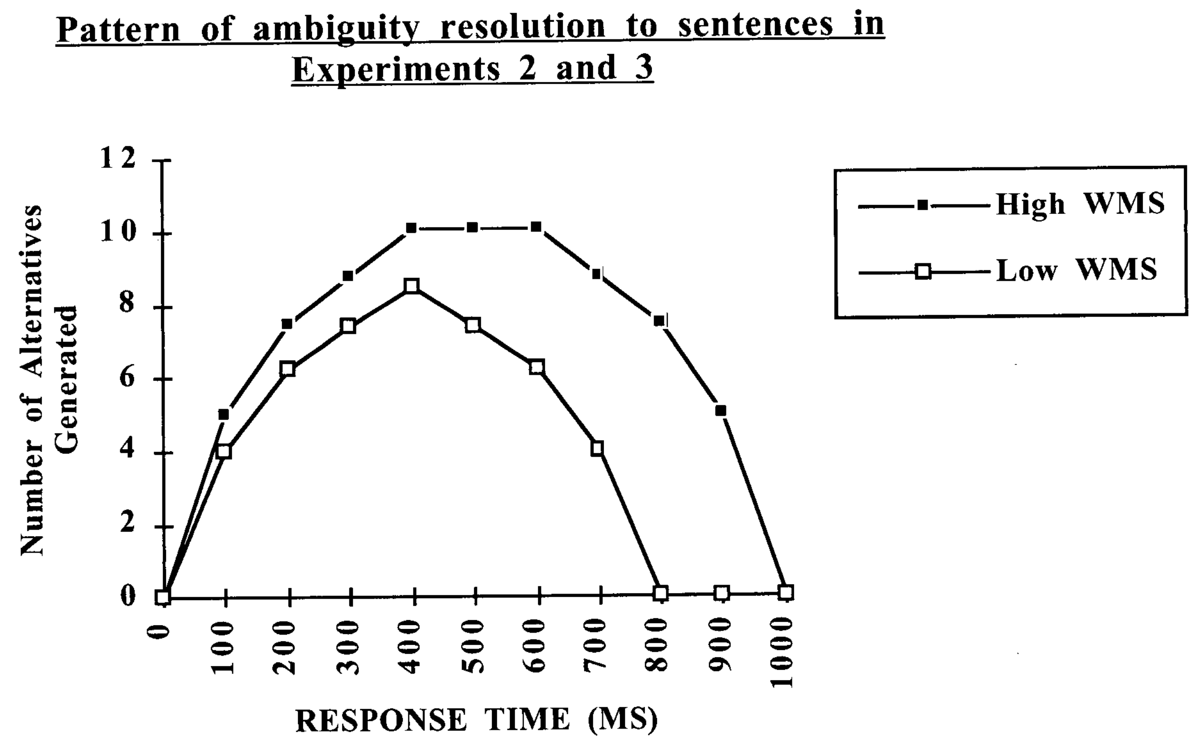
Hasher and Zacks

Stoltzfus et al.'s (1996) view that lack of inhibitory mechanisms explains individual differences in language comprehension is partly supported by the results from Experiment 4 showing that low WMS listeners may have difficulty inhibiting high-frequency competitors to low-frequency target words. In addition, the finding that it took longer for the low WMS listeners to isolate the low-frequency target words in Experiment 4 could mean that the high-frequency competitors that they generated interfere with further generation of low-frequency competitors due to poorer inhibition of the high-frequency competitors.

Baddeley and Logie's model of Working Memory

The findings from Experiments 1 - 4 all indicate that lexical level information interacts with phonological processing, especially for low WMS listeners, whereas evidence for interaction between sentence level and phonological processing was not found for either WMS group. These findings have important implications for Baddeley and Logies's (1999) model of working memory.

Because high WMS listeners showed longer response times than low WMS listeners in Experiments 2 and 3, but no clear differences in accuracy, Baddeley and Logie's (1999) argument that individual differences in spoken language processing may be dependent on activation and rehearsal capacity within the phonological loop is supported. Thus the high WMS listeners may be more able to <u>generate</u> and hold multiple interpretations of the signal, rehearsing them while considering the sentence context better than low WMS listeners (although the high WMS listeners appeared also to ultimately run out of processing time, as shown by more context word errors for the high WMS listeners in Experiment 2; see Figure 27, below).



<u>Figure 27:</u> Hypothesized pattern of results for Experiments 2 and 3 showing possible differences in generation of alternatives to target words, and response time for high and low WMS listeners.

Summary

An overview comparison of spoken processing by high and low WMS listeners is depicted in Figure 28. In summary, the results from the four experiments described above suggest that high and low WMS listeners do not appear to differ in temporal auditory processing of the incoming speech signal. (Although clear evidence for temporal auditory processing as an explanatory factor for WMS differences in spoken language processing was not found, it cannot be ruled out.) Clear differences do begin to be observed when the signal is encoded at the phonological level. Lexical processing interacts more with phonological processing for low WMS listeners, and they generate fewer alternatives to the ambiguous target words. The results from Experiment 4 provide direct evidence that high WMS listeners generate more alternatives to ambiguous target words and are more efficient in inhibiting incorrect alternatives. The inefficient performance by the low WMS listeners may be due to perseveration of incorrect alternatives, or poorer rehearsal capacity, preventing them from generating appropriate alternatives. Sentence context interacts less than lexical level processing with phonological processing for both WMS groups. Although low WMS listeners take longer to make lexical decisions than high WMS listeners, they take less time to process the sentence level information when the word-frequency of the ambiguous target words is controlled. This suggests that high WMS listeners may generate multiple sentence meanings as well as generating multiple alternatives to ambiguous target words. Thus, although temporal auditory processing appears to be similar for low and high WMS listeners, lexical processing interacts more with phonological processing for low than high WMS listeners, and sentence level processing interacts less than lexical processing with phonological processing for both WMS groups. The extent of interaction between types of processing for the two groups seems to crucially depend on time course.

While the experiments reported above provided informative results in many respects, there were several methodological shortcomings that prevented drawing stronger conclusions, particularly for Experiments 2 and 3. First, measuring storage simultaneously with processing, rather than comparing results to WMS measures taken independently of the experimental task is

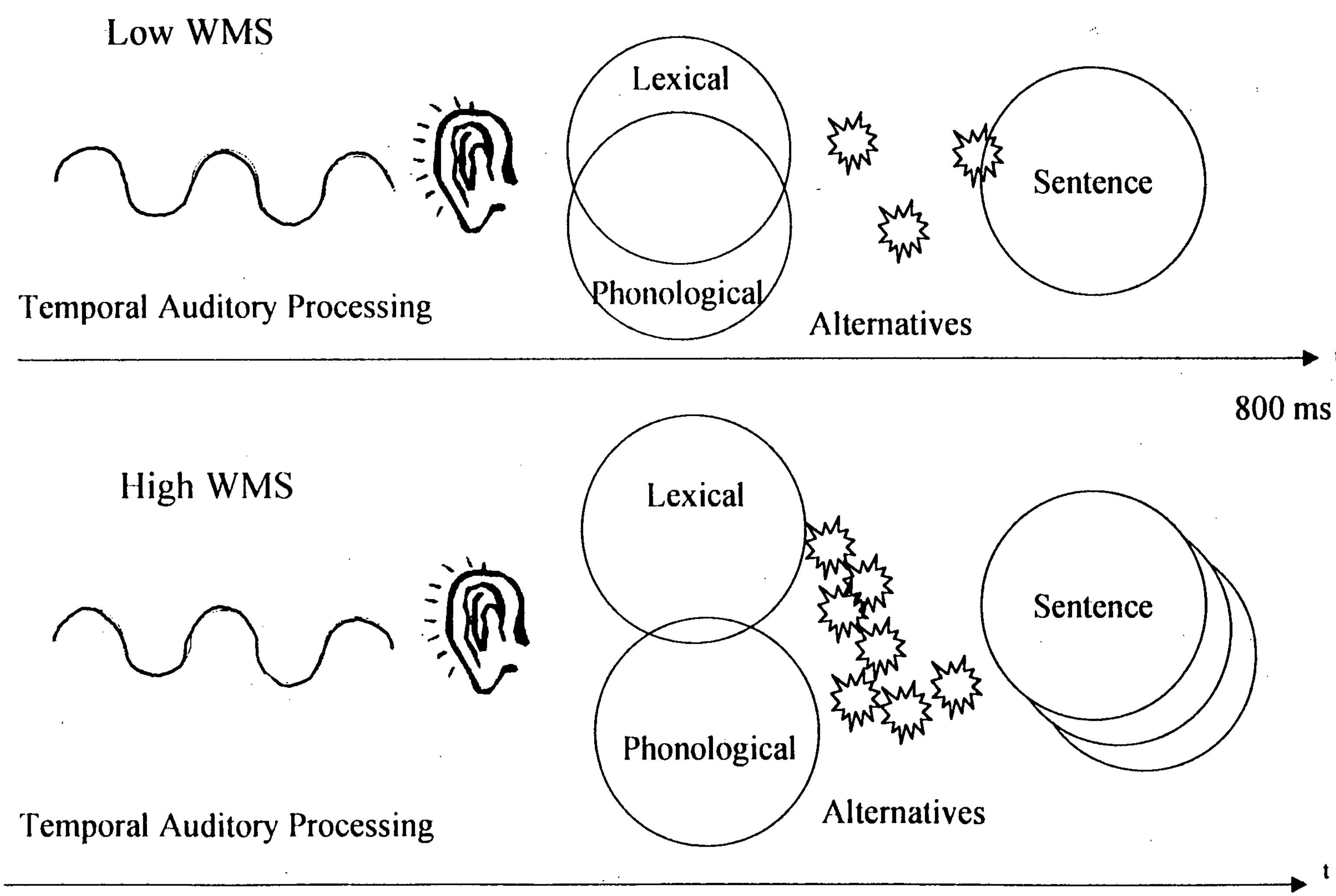


Figure 28: Diagram contrasting auditory, phonological, lexical and sentence processing by high and low WMS listeners

recommended in future studies. Second, instead of measuring gap-detection thresholds separately, it would be ideal to integrate a measure of temporal auditory processing into the task, such as varying temporal properties of the signal, for example, by presenting speeded speech. Third, to more directly test the influence of sentence context on the generation of alternatives to ambiguous target words, instead of taking off-line measurements of processing, as in Experiments 1 - 3, it would be preferable to design an on-line task. In a future version of Experiments 2 and 3, the following changes would be recommended: 1) Control word properties such as word-frequency, animacy, word class and phonological properties, 2) Sequence the word before the sentence context, and 3) Measure word recognition and generation of alternatives during presentation.

The research reported here is the first to show how different levels of language processing interact differently for groups that differ in WMS. These differences seem to crucially depend on time course of processing. It has been shown that differences in language processing vary depending on the time delay, as illustrated by differences in on-line vs off-line measures. Therefore, the methodological refinements recommended above would be valuable in further research because they would illuminate how the interaction of levels of processing may change over time for high and low WMS listeners.

REFERENCES

Altmann, Gerry T. M. (1990). Cognitive models of speech processing: An introduction. In Gerry T. M. Altmann (Ed.), Cognitive models of speech processing: Psycholinguistic and computational perspectives, (pp. 1-23). Cambridge, MA: Mit Press.

Andruski, Jean E., Blumstein, Sheila E., & Burton, Martha (1994). The effect of subphonetic differences on lexical access. Cognition, 52, 163-187.

AVAAZ Innovations, Inc. (1994). Computerized Speech Research Environment (CSRE v4.5). (London, ON).

Baddeley, A. D. (1996). Exploring the central executive. Quarterly Journal of Experimental Psychology, 49A, 5-28.

Baddeley, A. D., Eldridge, M., & Lewis, V. (1981). The role of subvocalisation in reading. Quarterly Journal of Experimental Psychology, 33A, 439-454.

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. H. Bower (Ed.), The psychology of learning and motivation, 8, (pp. 47-90). New York: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). WM: The multiple-component model. In Akira Miyake and Priti Shah (Eds.), Models of WM: Mechanisms of active maintenance and executive control, (pp. 28-61). Cambridge, UK: Cambridge University Press.

Baddeley, A. D., Vallar, G., & Wilson, B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 509-529). London: Erlbaum.

Berndt, R.S., & Caramazza, A. (1980). A redefinition of the syndrome of Broca's aphasia: implications for a neuropsychological model of language. Applied Psycholinguist, 1, 225-78.

Bever, T. G., Garrett, M. F., & Hurtig, R. (1973). The interaction of perceptual processes and ambiguous sentences. Memory and Cognition, 1, 277-286.

Bernstein, Lynne E., & Auer, Jr., Edward T. (1996). Word recognition in speechreading. In D.S. Stork, & Marcus E. Hennecke (Eds.), Speechreading by humans and machines: Models, systems, and applications, (pp. 17 -26). Berlin: Springer-Varlag.

Bird, H., Howard, D., & Franklin, S. (2001). Noun-verb differences? A question of semantics: A response to Shapiro and Caramazza. Brain and Language, 76, 213-222.

Broadbent, D. E. (1967). Word-frequency effect and response bias. Psychological Review, 74, 1-15.

Caplan, David. (1987). Neurolinguistics and linguistic aphasiology : An introduction. Cambridge [Cambridgeshire]; New York: Cambridge University Press.

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. Cognition , 21, 117-175.

Caplan, D., & Dufault, C. (2000, April). Working memory and on-line sentence processing efficiency in patients with dementia of the Alzheimer's type (DAT). Poster presented at the Cognitive Aging Conference, Atlanta, GA.

Caplan, D., & Waters, G. S. (1995). Aphasic disorders of syntactic comprehension and working memory capacity. Cognitive neuropsychology, 12, 637-649.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. Behavioral and Brain Sciences, 22, 77-126.

Carpenter, Patricia A., Miyake, Akira, & Just, Marcel (1995). Language comprehension: Sentence and discourse processing. Annual Review of Psychology, 46, 91-100.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Chomsky, N. (1964). Current issues in linguistic theory. The Hague: Mouton.

Chomsky, N. (1957). Syntactic structures. The Hague: Mouton.

Connine, Cynthia M. (1990). Effects of sentence context and lexical knowledge in speech processing. In Gerry T. M. Altmann (Ed.), Cognitive models of speech processing: Psycholinguistic and computational perspective, (pp. 281-294). Cambridge, MA: Mit Press

Connine, C., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in audoitory word recognition: Temporal and linguistic constraints. Journal of Memory and Language, 30, 234-250.

Connine, Cynthia M., Blasko, Dawn G., & Wang, Jian (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. Perception & Psychophysics, 56, 624-636.

Connine, Cynthia M., & Clifton, C., Jr. (1987). Interactive use of information in speech perception. Journal of Experimental Psychology: Human Perception & Performance, 13, 291-299.

Cook, V. J. (1975). Strategies in the comprehension of relative clauses. Language and Speech, 18, 204-212.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Crystal, David (1985). Linguistics. Markham, ON: Penguin Books.

Cutler, A., & Clifton, C., Jr. (1999). Comprehending spoken language: a blueprint of the listener. In C. M. Brown and P. Hagoort (Eds.), The neurocognition of language, (pp.123-166). NY, NY: Oxford University Press.

Daneman, Meredyth, & Carpenter, Patricia A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning & Verbal Behavior, 19, 450-466.

Daneman, Meredyth, & Merikle, Philip M. (1996). Working memory and language comprehension: A meta-analysis. Psychonomic Bulletin & Review, 3, 422-433.

Daneman, Meredyth, & Tardif, T. (1987). Working memory and reading skill reexamined. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading, (pp. 491-508). Hillsdale, NJ: Erlbaum.

Dillon, L. (1995). The effect of noise and syntactic complexity on listening comprehension. Special Collections Div., University of British Columbia. School of Audiology and Speech Sciences. Thesis. (M.Sc.).

Ebbinghaus, H. (1985/1913). Memory: A contribution to experimental psychology (H. A. Ruger, & C. E. Bussenius, Trans.). New York: Columbia University, Teacher's College. (Reprinted 1964, New York: Dover.)

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. Journal of Memory and Language, 27, 143-165.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 972-992.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In Akira Miyake and Priti Shah (Eds.), Models of WM: Mechanisms of active maintenance and executive control, (pp. 257-297). Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. Psychological Review, 102, 211-245.

Fodor, J. (1983). The modularity of mind. Cambridge, MA: MIT.

Francis, W. Nelson, & Kucera, Henry with the assistance of Andrew W. Mackie (1982). Frequency analysis of English usage: lexicon and grammar. Boston, MA: Houghton Mifflin.

Fry, D. B. (1979). The physics of speech. Cambridge University Press: Cambridge, UK.

Ganong, W. F. III. (1980). Phonetic categorization in auditory word perception. Journal of Experimental Psychology: Human Perception and Performance, 6, 110-125.

Garfield, J. L. (Ed.). (1989). Modularity in knowledge representation and natural-language understanding. Cambridge, MA: MIT Press.

Gathercole, S. E., & Baddeley, A. D. (1989). Development of vocabulary in children and short-term phonological memory. Journal of Memory and Language, 28, 200-213.

Gathercole, S. E., & Baddeley, A. D. (1993). Working memory and language. Hove, UK: Erlbaum.

Gernsbacher, M. A., & Faust, M. (1991). The mechanism of suppression: A component of general comprehension skill. Journal of Experimental Psychology: Learning, Memory, & Cognition, 17, 245-262.

Greenberg, S. (1996). Auditory processing of speech. In Norman Lass (Ed.), Principles of experimental phonetics (pp. 362-407). St. Louis, MO: Mosby.

Grosjean, Francois (1980). Spoken word recognition and the gating paradigm. Perception and Psychophysics, 28, 267-283.

Grosjean, Francois (1996). Gating. Language and Cognitive Processes, 11, 597-604.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 163-169.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Haubert, N., & Pichora-Fuller, M. K. (submitted). Gap detection and word discrimination by young and old adult listeners. Manuscript sumbitted for publication.

Howes, D., & Solomon, R. L. (1951). Visual duration thresholds as a function of word probability. Journal of Experimental Psychology, 41, 401-410.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. Journal of Memory and Language, 30, 685-701.

Just, Marcel A., & Carpenter, Patricia A. (1987). A computer simulation of reading: The READER model. In M. Just and P. Carpenter (Eds.), The psychology of reading and language comprehension, (pp. 261-283). Boston, MA: Allyn & Bacon.

Just, Marcel A., & Carpenter, Patricia A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99, 122-149.

Kemper, S., & Anagnopoulos, C. (1993). Adult use of discourse constraints on syntactic processing. In J. Cerella, J. Rybash, W. Hoyer, and M. L. Commons (Eds.), Adult information processing: Limits on loss, (pp. 489 –507). San Diego: Academic Press.

Kintsch, Walter (1988). The use of knowledge in discourse processing: A construction-integration model. Psychological Review, 95, 163-82.

Kintsch, Walter (1998). Comprehension: A paradigm for cognition. Cambridge, UK: Cambridge University Press.

Kintsch, Walter, & van Dijk, T. A. (1978). Towards a model of text comprehension and production, Psychological Review, 85, 363-94.

Ladefoged, P. (1975). A course in phonetics. NY, NY: Harcourt Brace Jovanovich, Inc.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. Cognition, 38, 245-294.

Lam, W. (in preparation). Adaptation to compressed speech in younger and older listeners. Unpublished master's thesis, University of British Columbia, Vancouver, British Columbia, Canada.

Lewis, R. L. (1996). Architecture matters: What soar has to say about modularity. In David M. Steier, & Tom M. Mitchell (Eds.), Mind matters: A tribute to Allen Newell (pp. 75-84). Mahwah, NJ: Lawrence Erlbaum Associates.

Lieberman, P., & Blumstein, S. (1988). Speech physiology, speech perception, and acoustic phonetics. Cambridge, UK: Cambridge University Press.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. Word, 30,384-422.

Lloyd, V.L. (1998a, June). The role of working memory span in spoken language comprehension. Poster presented at the 59th Annual Convention of the Canadian Psychological Association, Edmonton, AB, Canada.

Lloyd, V.L. (1998b, July). The roles of WMS, contextual and acoustic cues in spoken language comprehension. Poster presented at the NATO Advanced Study Institute on Computational Hearing, Il Ciocco, Italy.

Lloyd, V.L., & Pichora-Fuller, M.K. (1996, May) Inter-task reliability for a sentence comprehension test. Poster presented at the Third International Conference on Communication, Aging and Health, Kansas City, Missouri.

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. Doctoral dissertation,Indiana University, Bloomington, Indiana.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. Science, 189, 226-227.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. Cognition, 25, 41-102.

Marslen-Wilson, William (1990). Activation, competition, and frequency in lexical access. In Gerry T. M. Altmann (Ed.), Cognitive models of speech processing: Psycholinguistic and computational perspective, (pp. 148-172). Cambridge, MA: Mit Press

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. Cognition, 8, 1-71.

Marslen-Wilson, W., & Welsh, A. (1978). Processing interaction and lexical access during word recognition in continuous speech. Cognitive Psychology, 10, 29-63.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. Discourse Processes, 22, 247-288.

McClelland, J., & Elman, J. (1986). The trace model of speech perception. Cognitive Psychology, 18, 1-86.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Miller, G. A., & Nicely, P. (1955). An analysis of perceptual confusions among English consonants. Journal of the Acoustical Society of America, 2, 338-352.

Moore, Brian C. J. (1989). An introduction to the psychology of hearing. London; New York: Academic Press.

Miyake, Akira, Just, Marcel, & Carpenter, Patricia (1994). Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. Journal of Memory and Language, 33, 175-202.

Miyake, Akira, & Shah, Priti (1999). Emerging consensus, unresolved issues, future directions. In Akira Miyake and Priti Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control, (pp. 1-27). Cambridge, UK: Cambridge University Press.

Norris, D. G. (1993). Bottom-up connectionist models of 'interaction'. In G. T. M. Altmann and R. Shillcock (Eds.), Cognitive models of speech processing, (pp. 211-234). NJ: Erlbaum.

Norris, D. G., McQueen, J. M., & Cutler, A. (1998). Merging phonetic and lexical information in phonetic decision-making. (Manuscript.)

Perfetti, C. A., & Roth, S. (1981). Some of the interactive processes in reading and their role in reading skill. In A. M. Lesgold, & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, HJ: Erlbaum.

Pichora-Fuller, M. K. (1996). Working memory and speechreading. In D.S. Stork, & Marcus E. Hennecke (Eds.), Speechreading by humans and machines: Models, systems, and applications, (pp. 257-274). Berlin: Springer-Varlag.

Pichora-Fuller, M. K., Schneider, Bruce A., & Daneman, Meredyth (1995). How young and old adults listen to and remember speech in noise. Journal of the Acoustical Society of America, 97, 593-608.

Pickles, James O. (1988). An introduction to the physiology of hearing. London; NY: Academic Press.

Raven, J. C. (1938). The Mill Hill Vocabulary Scale. London: Lewis.

Roth, S. F., Perfetti, C. A., & Lesgold, A. M. (1979, May). Reading ability and children's word identification processes. Paper presented at the Midwestern Psychological Association Meeting, Chicago.

Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. Journal of Memory and Language, 24, 210-231.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. Psychological Review, 103, 403-428.

Samuel, A. (1981). Phonemic restoration: Insights from a new methodology. Journal of Experimental Psychology: General, 110, 474-494.

Samuel, A. (1990). Using perceptual-restoration effects to explore the architecture of perception. In Gerry T. M. Altmann (Ed.), Cognitive models of speech processing: Psycholinguistic and computational perspective, (pp. 295-314). Cambridge, MA: Mit Press.

Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., & Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. Journal of the Acoustical Society of America, 95, 980-991.

Selfridge, O. (1959). Pandemonium: A paradigm for learning. In D. Blake & A. Utteley (Eds.) Symposium on the mechanization of thought processes. London: H. M. Stationery Office.

Shapiro, K., & Caramazza, A. (2001). Sometimes a noun is just a noun: Comments on Bird, Howard, and Franklin (2000). Brain & Language, 76, 202-212.

Shah, Priti, & Miyake, Akira (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. Journal of Experimental Psychology: General, 125, 4-27.

Shah, Priti, & Miyake, Akira (1999). Models of working memory: An introduction. In Akira Miyake and Priti Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control, (pp. 1-27). Cambridge, UK: Cambridge University Press.

Slobin, D. I. (1979). Psycholinguistics. Glenview, Ill.: Scott, Foresman.

Small, J. A., Kemper, S, & Lyons, K. (2000). Sentence repetition and processing resources in Alzheimer's disease. Brain and Language, 75, 232-258.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 16, 32-71.

Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. Memory and Cognition, 7, 77-85.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. Journal of Experimental Psychology: Human Perception and Performance, 7, 658-672.

Stoltzfus, E. R., Hasher, L., & Zacks, R. T. (1996). Working memory and aging: Current status of the inhibitory view. In John T. E. Richardson, Randall W. Engle, Lynn Hasher, Robert H. Logie, Ellen R. Stoltzfus, & Rose T. Zacks (Eds.), Working memory and human cognition, (pp. 66-88). New York, New York: Oxford University Press.

Strouse, Anne, Ashmead, Daniel H., Ohde, Ralph N., & Grantham, D. Wesley (1999). Temporal processing in the aging auditory system. Journal of the Acoustical Society of America, 104, 2385-2399.

Tartter,Vivien C. (1986). Language processes. New York: Holt, Rinehart and Winston.

Tyler, L. K., & Marslen-Wilson, W. D. (1981). Children's processing of spoken language. Journal of Verbal Learning and Verbal Behavior, 20, 400-416.

Tyler, Richard S., Summerfield, Quentin, Wood, Elizabeth J., & Fernandes, Mariano A. (1982). Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. Journal of the Acoustical Society of America, 72, 740-752.

Vallar, G., & Shallice, T. (Eds.). (1990). Neuropsychological impairments of short-term memory. Cambridge, UK: Cambridge University Press.

van Noorden, L. P. A. S. (1975). Temporal Coherence in the Perception of Tone Sequences. Unpublished doctoral dissertation, Eindhoven University of Technology, The Netherlands.

Warren, R. M., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word-recognition. Perception and Psychophysics, 41, 262-275.

Warren, R. M., & Marslen-Wilson, W. D. (1988). Cues to lexical choice: Discriminating place and voice. Perception and Psychophysics, 43, 21-30.

Warren, R. M., & Warren, R. P. (1970). Auditory illusions and confusions. Scientific American, 223, 30-36.

Willott, James F. (1991). Aging and the auditory system: Anatomy, physiology, and psychophysics. San Diego, CA: Singular.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. Journal of the American Academy of Audiology, 7, 175-182.

Wingfield, A., Lindfield, K. C., & Goodglass, H. (2000). Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition. Journal of Speech, Language, and Hearing Research, 43, 915-925.

Zacks, R. T., & Hasher, L. (1997). Cognitive gerontology and attentional inhibition: A reply to Burke and McDowd. Journal of Gerontology: Psychological Sciences, 52, 274-283.

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

APPENDIX A

Three sentence types for Experiments 2 and 3: Context-Consistent, Context-Neutral, and Context-Inconsistent:

Context-Consistent:

The kiddies **pat** the doggies and chase toy rockets.
The kiddies **bat** the windows and shatter the glass.
The athletes **pat** their foreheads and drink some water.
The athletes **bat** the average and secure the lead.

Context-Neutral:

The families **bat** the things and discuss the action.
The families **pat** the things and discuss the action.
The families **bat** the stuff and think of ideas.
The families **pat** the stuff and think of ideas.

Context-Inconsistent:

The kiddies **bat** the doggies and chase toy rockets.
The kiddies **pat** the windows and shatter the glass.
The athletes **bat** their foreheads and drink some water.
The athletes **pat** the average and secure the lead.

APPENDIX B

<u>Sample response sheet for "Pen and Paper" task response form used in Experiment 1:</u>

Sample Response Sheet:

1. cab cap cat dock dog dot duck mouse pick pig pip

2. cab cap cat dock dog dot duck mouse pick pig pip

3. cab cap cat dock dog dot duck mouse pick pig pip

4. cab cap cat dock dog dot duck mouse pick pig pip

5. cab cap cat dock dog dot duck mouse pick pig pip

APPENDIX C

The forced-choice responses which appeared on the computer screen for Experiment 2:

Bat (Converse)	Pat (Converse) (3)*
Bat (Destroy)	Pat (Destroy) (48)
Bat (Follow)	Pat (Follow) (97)
Bat (Guzzle)	Pat (Guzzle) (1)
Bat (Invent)	Pat (Invent) (7)
Bat (Tighten)	Pat (Tighten) (3)

* Word Frequency of synonyms for final verbs in each sentence as verb form is shown in brackets (Francis & Kucera, 1982)

APPENDIX D

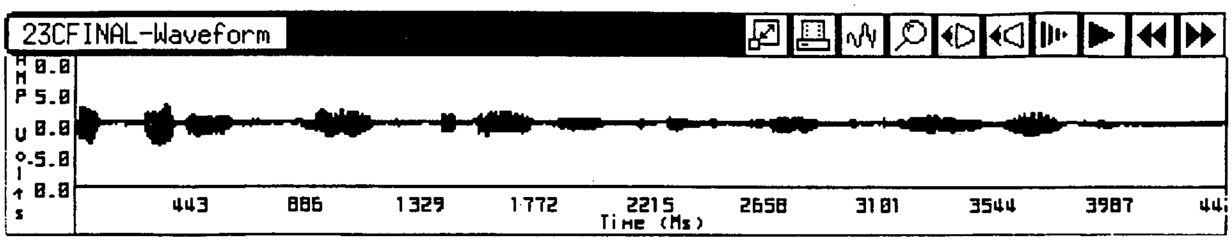
The forced-choice responses which appeared on the computer screen for Experiment 3:

Bat (Chase)	Pat (Chase) (4)*
Bat (Discuss)	Pat (Discuss) (28)
Bat (Drink)	Pat (Drink) (25)
Bat (Secure)	Pat (Secure) (16)
Bat (Shatter)	Pat (Shatter) (2)
Bat (Think)	Pat (Think) (433)

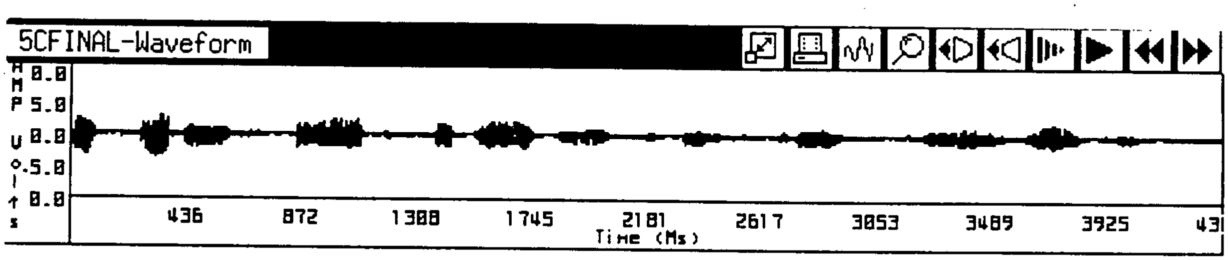
* Word Frequency for final verb in each sentence as verb form is shown in brackets (Francis & Kucera, 1982)

APPENDIX E

<u>Time waveforms of two sentences from Experiments 2 and 3: Top figure with *bat* and bottom figure with *pat*:</u>



<u>Figure E1:</u> Time waveform of the sentence *The kiddies <u>pat</u> the doggies and chase toy rockets.*



<u>Figure E2:</u> Time waveform of the sentence *The kiddies <u>bat</u> the doggies and chase toy rockets.*

APPENDIX F

Time waveforms of two target words from Experiment 4: Top figure for the High-Low word*quick* and bottom figure for the Low-High word *streak:*

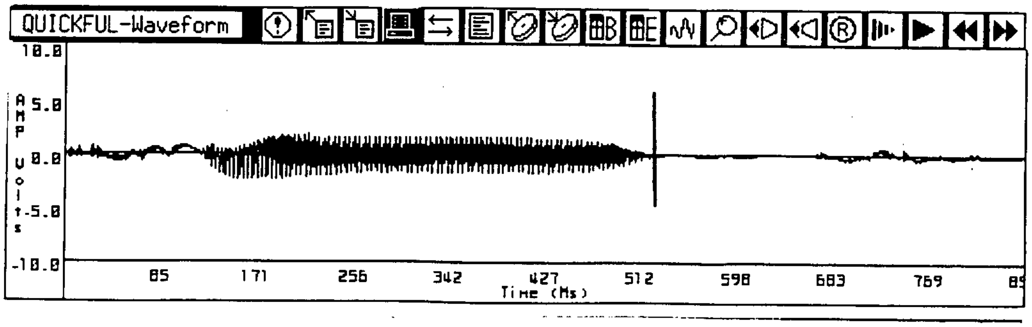


Figure F1: Time waveform of the High-Low target word *quick* showing alignment point marker.

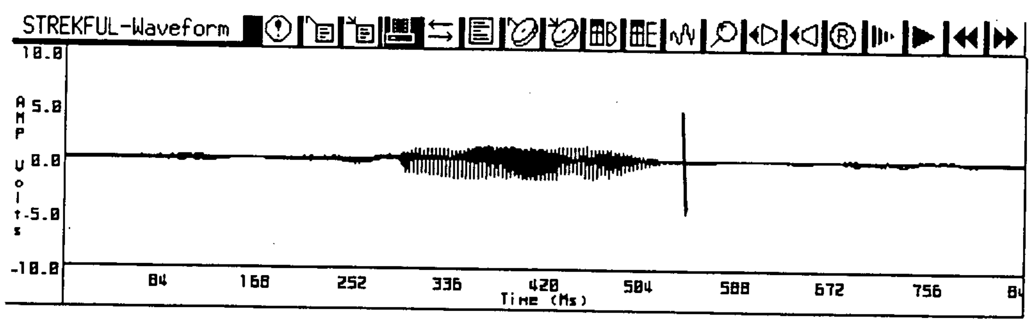


Figure F2: Time waveform of the Low-High target word *streak* showing alignment point marker.