DOCTORAL THESIS

Room Acoustics and Cognitive Load when Listening to Speech

Robert Ljung

Room Acoustics and Cognitive Load when Listening to Speech

Robert Ljung

Department of Building, Energy and Environmental Engineering

Faculty of Engineering and Sustainable Development

University of Gävle

and

Division of Technical Psychology

Department of Human Work Sciences

Luleå University of Technology

Printed by Universitetstryckeriet, Luleå 2010

ISSN: 1402-1544
ISBN 978-91-7439-099-5

Luleå 2010

www.ltu.se

Abstract

The present thesis investigated the effects of background noise or a long reverberation time in learning situations. All studies used spoken texts that were presented above the speech intelligibility threshold, but were degraded enough to make listening more effortful. The basic hypothesis for the whole project was that listening to speech in a bad acoustic environment should increase the cognitive load for the listener, which should impair memory of the text. In Paper I the auditory stimuli were lists of words and sentences that were degraded by a background noise. Paper II was a replication of the experiment from paper I, but the independent variable was changed from the level of the background noise to reverberation time. Paper III included two experiments where the stimulus material was 10 minutes lectures. Paper IV included two studies. The first experiment investigated whether serial recall performance is affected when words are presented in long reverberation time (Exp 1a). In experiment 1b word lists were presented in long or short reverberation time or with a background noise. The stimuli were recorded in one classroom with extremely good and one with very bad acoustic design. In experiment 2 word lists with many or few phonological neighbours were presented with long or short reverberation time. In all studies some measure of working memory capacity was included. Taken together, the overall results could be summarized in two sentences: Hearing what is said is a necessary but not a sufficient criterion for people to remember what is said, which means that spoken information should be heard without special effort, otherwise proper learning is jeopardized. No consistent relation was found between working memory capacity and the learning effect in the unfavorable listening conditions.

Acknowledgements

I would like to express my sincere gratitude to the people who have supported me in the completion of this work and in particular:

Professor Anders Kjellberg, supervisor, for bright and skilful scientific guidance and for unfailing support.

Professor Staffan Hygge, assistant supervisor, for your endless support and enthusiasm for my research work.

Professor Håkan Alm, formal supervisor, for all administrative help at Luleå University.

Stig Arlinger for useful comments on an earlier draft of this thesis.

Patrik Sörqvist, Helena Jahncke, David Hallman, Anne-Marie Green and Niklas Halin, for all interesting discussions and for being good travel mates.

Örjan Johansson and Johan Odelius at Luleå University, for help and guidance with recordings and acoustic settings.

All staff at Örebro Audiological Research Centre (Ahlséns) for technical assistance during the data collection.

Tommi Syri for technical assistance and programming.

All colleagues at BMG for the discussions during the coffee breaks.

Mum and Dad!

Finally, I would like to thank my wife Susanna and our children William, Anton and Alma. You give me so much joy and love every day!

This work was made possible by financial support from Swedish Council for Working Life and Social Research (FAS), and the Department of Building, Energy and Environmental Engineering, Faculty of Engineering and Sustainable Development, University of Gävle.

List of Papers

This Thesis is based on the following papers.

I. Kjellberg, A., Ljung, R., & Hallman, D. (2008). Recall of words heard in noise. *Applied Cognitive Psychology, 22,* 1088-1098

II. Ljung, R., & Kjellberg, A. (2009). Long reverberation time decreases recall of spoken information. *Building Acoustics, 16*, 4, 301–312.

III. Ljung, R., Sörqvist, P., Kjellberg, A., & Green, A. M. (2009). Poor listening conditions impair memory for intelligible lectures: implications for acoustic classroom standards. *Building Acoustics, 16*, 3, 257–265

IV. Ljung, R., Sörqvist, P., and Kjellberg, A. (Submitted). The locus of effects of degraded speech signals on memory for spoken materials: Categorical but not serial processing is affected.

Table of contents

Introduction

There are mainly two acoustical parameters that determine our ability to understand speech in addition to the absolute speech level: the signal-to-noise (S/N) ratio and the reverberation time (RT). The S/N ratio is a measure of the signal's strength in relation to the noise from surrounding sources, which could be babbling colleagues, ventilation noise, fans etc. RT is a measure of how long time it takes for a sound to decline 60 dB after it has been turned off. A too long RT, which e.g. may be heard in large churches, impairs speech perception by slurring the speech and making it less clear.

When we are listening to a degraded message we have to add other information to the speech signal, for example from the context and redundant information in the sentence, to be able to understand what is said. In such unfavorable acoustical conditions we have to interpret the signal rather than just listen to it. An easy automatic bottom-up process like listening is thereby transformed into a resource demanding top-down process. When we for example listen to a lecture it is essential to hear, understand and remember the spoken message. These cognitive tasks require the use of our working memory, which could be viewed as a pool of cognitive recourses that handles temporary information, and which has a limited capacity. When we are listening to spoken information in a bad acoustic environment, like a low S/N ratio or long RT, we have to spend more cognitive resources just to identify the words, which leave fewer cognitive resources to further elaboration of the speech. Thus, even if the acoustical environment allows us to hear what is said, it may make it more difficult to understand the message, and even if we have understood the message, our ability to remember the information may be degraded. This kind of reasoning is underpinning this thesis, and hopefully, this work can bring some new knowledge that could be of interest when guidelines and norms are set for buildings where communication and learning are of special importance.

Speech

The speech sounds are divided into two main groups: vowels and consonants. The vowels are made with the windpipe open, whereas the air-flow through the windpipe is throttled or closed when a consonant is produced. As an obvious consequence the vowels generally are louder and more audible than the consonants. The difference between these parts of speech is significant for understanding speech perception problems for listeners with significant hearing impairment (Arlinger, 2007). The sounds of human speech fall mostly in the range of about

100Hz to 8,000Hz, and the main energy of vowels is generally in a lower frequency range than that of consonants. Furthermore, there are more consonants than vowels, and consonants therefore carry more information than vowels. There are many examples that prove this fact, e.g., it is possible to read a text when all vowels have been removed, but the other way around makes the text unreadable.

Room acoustics

In a free field the sound pressure level drops 6 dB for each doubling of the distance to the sound source. Free field means that only the direct sound waves reach the listener, and no reflected sound reinforces the signal. If all sound waves that hit a surface in a room are absorbed (e.g. in an anechoic chamber), the room acoustics will show the same pattern as in a free field. A building where oral communication is of importance, e.g. a classroom, must have some reflecting surfaces that amplify the teacher's speech enough to make it intelligible for the students in the far end of the class room. On the other hand, if the surfaces are too reflective, the RT will be too long and the speech will be indistinct and lose intelligibility. The reason for that phenomenon is that the reflected sound has travelled a longer distance and arrives later to the listener than the direct sound. Reflected speech sounds that arrive at the ear within 35-40 ms after the direct sound subjectively summate with the direct sound and thus improve the intelligibility (Rossing, 1990). If the RT is longer a reflected phoneme may mask next phoneme instead of amplifying the original phoneme, thus distorting the speech signal and making the listener lose some information. A shorter RT therefore gives a clearer signal and better speech intelligibility given a constant S/N ratio (but, as mentioned, parts of the reverberating sound may improve the S/N ratio). The masking effect of the reverberating sound becomes larger when its level is high. Since vowels generally are much louder than consonants in verbal communication, a long RT affects consonants more than vowels. This is critical since consonants are more informative and crucial for understanding than vowels. Another acoustical phenomenon that is of interest for this effect is called "Upward spread of masking" and refers to the fact that a low-frequency sound masks a high-frequency sound more effectively than a high-frequency sound masks a low-frequency sound. Since many of the vowels (e.g. *a* and *o*) have their main energy in a lower frequency range than most of the consonants (e.g. *f* and *s*) this effect also means that consonant perception is especially difficult when the RT is too long. Thus, in conclusion, long RT increases the overall speech level and makes it easier to detect it, but seldom improves perception.

The relation between room acoustics and speech intelligibility is even more complex. A room usually has various RTs in each octave band. For example, a carpet has an absorption coefficient (*a*) of approximately 0.02 at 125 Hz and 0.65 at 4000 Hz, and an acoustical tile on concrete has a absorption coefficient of approximately 0.14 at 125 Hz and 0.37 at 4000Hz. An open window is assumed to “absorb” all the sound that reaches it (nothing is reflected); thus its absorption coefficient is 1, and a surface that absorbs half of the sound energy will have an absorption coefficient of 0.5 (Rossing, 1990). Thus, in general, there is more effective absorption in higher frequencies; consequently, the RT is longest in the lower frequency range.

School buildings are made to provide a good learning environment and to achieve that aim the acoustics are of special importance. Research has shown that acoustical problems are common in new as well as older schools, and that this may affect a child’s ability to understand what is said (Crandell & Smaldino, 2000a, 2000b; Shield & Dockrell, 2008). Nelson & Soli (2000) point out that poor listening conditions are especially critical for children with auditory disorders, but that also children with normal hearing are negatively affected. Nowadays teaching often involves workshops and other interactive moments where adequate communication is vital to obtain high-quality performance.

There is a general consensus about acoustical criteria for good classrooms, but supporting data from studies performed in actual classrooms are limited. The renewed interest in classroom acoustics is related to our growing understanding of the negative effects of noise and poor room acoustics on children’s ability to learn in schools (Bradley & Sato, 2008).

Speech perception

Speech is basically a stream of sound waves that we mentally split up in a row of single items (words), and each item is also mapped to one of tens of thousands of items stored in our mental lexicon (Cluff & Luce, 1990). This procedure is rapid and more or less automatic, and many researchers have tried to explain this complex and efficient process (Lieberman et al, 1967; Luce, 1986; Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986). Marslen-Wilson & Tyler (1980) developed the original version of the “cohort model” based on the assumptions that: (i) an auditory presented word activates a set of words in our mental lexicon; this set of words is the “word-initial cohort”. (ii) Words belonging to this cohort are then reduced due to further semantic and contextual information. (iii) This reduction process continues until contextual information and information from the word itself are sufficient to eliminate all but one of the words in the “word-initial cohort”. This is named “recognition point”. The first heard phoneme will activate a set of possible words; the next phoneme will reduce the number of activated words and so on until one reaches the recognition point. This view of the speech recognition model has been revised a couple of times (Marslen-Wilson, 1990; Marslen-Wilson &Warren, 1994); the difference between the original and the modified model is basically the criteria for the words to be a candidate. In the old model (Marslen-Wilson & Tyler, 1980) the possible candidates were either in or out of the cohort word set, whereas they later proposed that each word has a degree of activation with the most plausible word being most activated. The effects of contextual cues on word recognition are more limited in the revised model and, thus, it puts more stress on the bottom-up processing (Eysenck & Keane, 2005).

Luce (1986) proposed another model of word recognition for auditory stimuli known as the *Neighborhood Activation Model* (NAM). Fundamentally, Luce proposed that a spoken word activates a similarity “neighborhood” of phonetically similar words within the mental lexicon. According to the NAM model all phonemes (the whole word) contribute to the initial neighborhood activation. There are basically two factors that characterize the structure of a similarity neighborhood. The first, neighborhood density, refers to the number of phonetically similar words in a particular neighborhood. Neighborhoods with many words are defined as “dense”, whereas those with relative few words are defined as “sparse”. The second main variable is called “neighborhood frequency “, and refers to the average frequency in the language of the words in the neighborhood (Cluff & Luce, 1990). This reasoning relies on the word frequency effect in recognition memory, which refers to the finding that more frequent words are recognized faster than less frequent words; thus, frequent candidates are confirmed earlier than less frequent candidates. This means that less sensory evidence is needed to identify a very frequent word than a less frequent word (Morrison & Ellis, 1995).The degree of activation of the acoustic-phonetic patterns depends on the degree of similarity between the item in memory and the stimulus input. Thus, the number of words in the neighborhood and their frequency of occurrence are the main parameters for word recognition according to the NAM model. Since speech perception is very dependent on stimulus quality, the acoustic environment is highly important. In addition to the speaker’s voice and pronunciation, room acoustics affect our ability to detect a spoken message. A low S/N ratio means that the spoken word is somewhat masked by the surrounding noise. The level of the background noise is independent of the speaker, and is the same for the whole word (beginning, middle, and the end of the word). However, phonemes are pronounced at different sound levels, and the low level phonemes thus are hit hardest by surrounding noise since their S/N ratios are lowest. However, reverberation is a silent acoustic parameter, one cannot hear echo without a signal. The reverberation prolongs each phoneme, which is perceived as a slurred speech signal. Thus, the first phoneme is prolonged and thereby masks next phoneme. This means that a word heard in a room with a long RT is mainly degraded in its last part. This is true for single words in e.g. a word list but since speech basically is a long stream of sounds the masking is degraded equally over the whole auditory message.

Memory organization

Memory basically concerns storage and retrieval of information; we can twist and turn the concepts but it is still just storage and retrieval. So why should we describe the human memory in so many different ways? One general memory structure might of course be conceived. Scientists worldwide have conducted thousands of experiments, and interpretations of the results have made it hard to argue for such an idea. For example, researchers in the field of cognitive psychology have showed loads of results that detect a difference between long- and short-term memories. To make the results understandable, various models of memory structures have been suggested, and several models have been tested and rejected.

Nowadays, cognitive researchers talk about working memory rather than short-term memory. The concept of short-term memory is basically a time limited memory system that handles temporary information. In a classical paper by Miller (1956) he proposed that the short-term- memory can hold 7± 2 units without elaboration processing; this phenomenon is more known as the “magical number seven”. A short-term-memory test is basically just a test of how many items one can recall after presentation. The working memory is not just a memory, but a whole cognitive information processing centre, and a working memory test differs from a short-term-memory test since it contains a simultaneous cognitive task during the storing period. Researchers have been interested in working memory since this is considered as an important mechanism in many cognitive tasks, for example learning, problem solving and information processing. Working memory can be simplified as a temporary storage centre, a stage point where deduction and induction converge and make the sensations understandable.

Working memory capacity

Researchers in the field of cognitive psychology have widely used results of working memory capacity tests as a correlation measurement (Engle, 2002). The views of working memory capacity differ in some respects; some researchers suggest that working memory capacity actually is one's ability to control attention (Kane & Engle, 2003), whereas others propose that working memory capacity is a our ability to inhibit confounding information (Lustig, Hasher, & Zacks, 2008). Cowan (2005) suggests that working memory capacity is just a measure of one's mental storage capacity. Traditional working memory tests such as Reading span, Operation span and Counting span measure the ability to recall a set of presented words. During the storing session the participant has to perform a simultaneous task. In a recent article, Sörqvist, Ljungberg & Ljung (2010) showed evidence that working memory capacity should be divided into sub-processes. They developed the size-comparison span test, which makes it possible to measure intrusions from items that are part of the task but never meant for recall. Measurements of current-list intrusions errors seem to be highly related to semantic auditory distraction.

Baddeley & Hitch (1974) presented a model of working memory that has survived critical reviews the last 30 years, so, obviously, we have to take it seriously. The initial model was introduced in 1974, and contained a central executive unit and two slave systems, the phonological loop and the visual sketchpad. Many researchers criticized the approach and argued that the model was unable to explain many results. The main arguments were that memory capacity was too small and that the model missed a link to long-term memory (Was & Woltz, 2007). Baddeley (2000) has listened to this criticism and has recently adjusted the model by introducing a new component, the episodic buffer.

The phonological loop is assumed to manage auditory perception and speech production. The capacity is limited and auditory information is wiped out fairly quickly unless the information is repeated. Apart from neurological evidence Baddeley points out some phenomena that he believes demonstrate the existence of the phonological loop, for example:

The similarity effect – written lists of words that sound similar are more difficult to remember while semantic or visual similarity has less effect. This is evidence for acoustical or phonological coding of written words (Conrad & Hull, 1964). *The word length effect* - series of short words are easier to reproduce than series of long words. Baddeley interprets this in terms of time; basically it takes longer time to repeat long words and consequently rehearsal takes longer time and for a series with long words this increases the risk for them to be wiped out (Baddeley et al., 1975).

The effect of articulatory suppression - when subjects are prevented from repeating the words, reproduction performance deteriorates, and, moreover, the word length effect is wiped out. This supports the proposed importance of a rehearsal process (Baddeley et al., 1975).

The transformation of information - when a series of words are presented visually, a silent mental articulation is supposed to be made, which transforms the text data into auditory storage. When articulatory repetition is prevented, it is impossible to transfer data and the similarity effect disappears. This phenomenon is not true of auditory information because it is stored directly in the phonological loop, which was shown before Baddeley and Hitch presented their model of working memory (Murray, 1968).

The effects of a degraded speech signal could be interpreted in terms of this model since auditory information does not need to be mentally transformed to the phonological loop; competing with sounds such as noise that also have direct access to the phonological store. This means that degraded auditory information, with a low S/N ratio or a too long RT, gets into the phonological store without transformation. However, information is degraded and partly needs further processing to be identified. Speech recognition thus becomes more effortful since the speech is not perceived directly but has to be interpreted. As a consequence less capacity is left for further cognitive processing like storage, chunking and elaboration.

The visuo-spatial sketchpad handles and stores visual and spatial information, and it has basically the same role as the phonological loop has for auditory and acoustic information (Baddeley, 2007). Since this thesis not aim to study the visuo-spatial sketchpad that part of the model will not be developed further here.

The central executive was initially described with vague words as a limited capacity pool of general processing resources (Baddeley, 2007). Later Baddeley adopted Norman & Shallices' (1986) attention model, which basically explains executive function and resources in terms of schemata and automatic and controlled processes. An automatic process requires few resources and is based on past experience, while a controlled process is more costly and contains deductive thinking. Since controlled processes rely on previously stored information the next problem comes up, namely the link between working memory and long-term memory. As a solution to this problem Baddeley introduced the episodic buffer as a fourth component of the working memory model (Baddeley, 2000). The episodic buffer is assumed to operate as a link mechanism that allows perceptual information from the subsystems and information from long-term memory to be integrated into episodes.

Baddeley’s model of working memory has been extremely popular and influential in the field of cognitive research. Although it has been received with applause, it has also been highly criticized. The main criticism has been leveled at the extremely limited capacity (Was & Woltz, 2007).

Comprehension and memory of spoken information is heavily dependent on resources from working memory, a system that is responsible for temporary storage and cognitively demanding tasks (Baddeley, 2001). Integrating words and sentences into a context requires access to previously stored information. While these processes are ongoing, also recently heard information should be stored, which requires a variety of resources from the working memory, which, however, provides a limited amount of resources.

Recently, Rönnberg with colleagues have development a combined perception and working memory model named Ease of Language Understanding (ELU) (Rudner, Foo, Sundewall-Thorén, Lunner & Rönnberg, 2008; Rudner, Foo, Rönnberg & Lunner, 2009). The ELU model describes a role for working memory in language understanding, and suggests that the linguistic perceptual information input, at a cognitive level, automatically bounds together into phonological streams of information. However, in adverse listening conditions, and for persons with hearing impairment, explicit processing and storage capacity are required to infer the meaning of the message. This is related to cognitive capacity and speech processing under mismatch conditions, but not necessarily under matching conditions (Stenfelt & Rönnberg, 2009).

Effects on memory performance of the auditory environment

The effects of background noise on memory have been demonstrated in a large number of experimental studies (Beaman, 1998; Evans & Lepore, 1993). There are also several studies that indicate that long-term exposure to noise may impair cognitive performance of school children (Haines et al., 2001; Shield & Dockrell, 2003). Experiments on acute noise effects and studies of the effects of long-term exposure have almost exclusively used text-based test material (Smith, 1989), although many tasks in schools and workplaces require processing and storing of orally presented information. The obvious reason for using text-based tasks in such studies is that impaired performance in noise could otherwise be an effect of noise making it impossible to hear parts of what is said, which would make an effect rather trivial.

One common assumption is that if words are identified correctly, there is no negative effect of the acoustic environment, such as noise or a long RT (Surprenant, 1999). This thesis questions that assumption, and gives several examples of cognitive deterioration caused by degraded listening conditions. Auditory stimuli can be degraded by external factors such as noise and long reverberation (Rossing, 1990; Everest, 2001), or by individual factors like hearing loss or ability to understand speech with an accent or in a second language (Arlinger, 2007). It is easy to see the similarity between the deteriorated stimuli caused by either external or internal factors, and of course all possible interactions.

Listening to speech in unfavorable acoustic environments

Listening to speech in high background noise level and/or long reverberation puts higher demands on our working memory. The phonological coding of the speech, which is largely processed automatically in good listening conditions, turns into a top-down process when it becomes harder to hear what is said, and information processing is more dependent on previously stored information (Repovs & Baddeley, 2006). More items in the mental lexicon are activated by the ambiguous speech signal and an analysis of the context is necessary to choose among them. An additional noise effect on learning and understanding speech is that the listener needs to concentrate more to hear the speech, and more resources are directed towards the listening, while fewer cognitive resources are available to storing and understanding the meaning of the information. These processes can take place without difficulty during a short period, but if attention is detained for a long time it can lead to cognitive fatigue, and as a consequence task performance is likely to deteriorate (Kjellberg, 2004).

This kind of resource/capacity explanation was suggested by Rabbitt (1966; 1968; 1991) and Pichora-Fuller et al. (1995). The effect of the noise thus would be to cause difficulty in encoding the stimulus, which makes it necessary to spend more of a limited pool of resources just to distinguish and understand each word. Rabbitt (1968) suggested that when a person is listening to speech in noise this means essentially performing in a dual-task situation. Since the cognitive processing of the degraded stimulus is no longer automatic, some resources are being allocated to identify and understand the auditory stimulus, and those resources cannot therefore be used in high order cognitive processing. Rabbitt (1966) argued that the extra effort required for identifying words in noise would reduce the spare capacity left for the further processing of the speech. In two experiments he studied recognition of a series of 100 words, which had been presented with or without a background noise, and found some support for his hypothesis. To check that they had identified the words correctly, subjects had to shadow their presentation. He found no effect of the noise on the number of correctly recognized words but an increased number of false alarms, leading to a lowering of the signal detection parameter *Beta*. Several features of Rabbitt's design may probably have reduced the effects in these experiments. The task may become more sensitive to the detrimental effects of noise if the recognition task is substituted with the more difficult recall task. If the recognition task is used it is also likely that response times would be more sensitive than the number of recognized words; even if one reaches a correct decision about a word, a less efficient encoding of the word list may prolong the matching process necessary to reach this decision. Another feature of Rabbitt's study that may have contributed to the relatively weak effect was that he excluded all subjects who did not identify all words when they were presented. This means that the selected group consisted of the persons that best managed to hear words in noise. However, Rabbitt's interpretation of his results means that this effect should not be limited to noise. All types of degradations of the speech signal should have the same effect. This was supported by a later study (Rabbitt, 1991), where persons with a mild hearing impairment remembered fewer words than those with normal hearing although there was no difference between the groups in the number of correctly heard words. Rabbitt (1968) also studied the effect of a background noise on serial recall of a series of eight digits where there was a short pause between the first and second group of four digits. After the presentation of a series, subjects were asked to recall either the first or the second group of four digits in correct order. A background noise was added either to one or both of the two parts of the list, or to none of them. The recall of the first group was found to be better when it was followed by a group presented without noise than with noise. Rabbitt's interpretation was that the increased processing demands caused by the noise inhibited the rehearsal of the first four digits.

Rabbit (1968) showed that poor listening conditions (+5 dB S/N) impair memory of spoken prose. He let the participants listen to two prose passages and answer questions about the text afterwards. Each passage was divided in two halves. One of the prose passages was heard with noise in the second part, thus the conditions were *No Noise* and *No Noise/Noise*. The recall score for the second halves was significantly higher for the *No Noise* groups than for the *No Noise/Noise* groups. However, it is possible that noise impaired task performance simply because the noise made it impossible to hear what was said. More interesting is that the same pattern was shown for the first half. The recall score was significantly lower for the first half when the second half was heard in noise. His interpretation of the results was that increased difficulty of recognition of speech in noise may interfere with other activities (rehearsal), which may be necessary to retain data in memory.

Murphy, Craik, Li, & Schneider (2000) proposed that many factors could explain the adverse effect that babble noise has on memory for spoken words. First, babble could degrade the perceptual representation of words to such an extent that it compromises their subsequent processing. Second, the presence of babble noise in the period between word presentations could interfere with rehearsal, and thirdly, the top-down processes needed to extract the words from the babble could draw on resources that otherwise would be used for encoding. Heinrich, Schneider, & Craik (2008) tested all these hypotheses by presenting babble either only during word presentation or rehearsal, or by gating the babble on and off 500 ms before and after each word pair. Their results showed that only the last condition led to a decline in memory. They propose that this decline in memory occurred because participants were focusing their attention on the auditory stream rather than on remembering the words they had heard.

Effects of aging on auditory stimuli, cognitive or perceptual deterioration

Research about aging is another field within cognitive psychology that gives support to the view that degraded stimuli allocates more cognitive resources of our limited capacity. There is a debate if the age-related deterioration is due to slower cognitive functions or a perceptual decline. Both explanations mean that the age-related deterioration in memory performance is interpreted in a similar way as the effect of a degraded speech signal. In order to investigate age-related differences in identification and recall of sentence-final words heard in a babble background, Pichora-Fuller et al. (1995) conducted a couple of experiments. In experiment 1, the level of the babble was varied to determine psychometric functions (percent correct word identification as a function of S/N ratio) for old adults with near-normal hearing and young normal-hearing adults, when the sentence-final words were either predictable (high context) or unpredictable (low context). Differences between the psychometric functions for high- and low-context conditions were used to show that the old listeners derived more benefit from supportive context than did the young listeners. In experiment 2 a working memory task was added. Specifically, after listening to and identifying the sentence-final words for a block of sentences, the subjects were asked to recall the last words that they had identified. Old subjects recalled fewer of the items they had perceived than did young subjects in all S/N conditions, even though there was no difference in the recall ability of the two age groups when sentences were read. Furthermore, the number of items recalled by both age groups was reduced in adverse S/N conditions. The results were interpreted as supporting a processing model in which reallocable processing resources are used to support auditory processing when listening becomes difficult either because of noise or because of age-related deterioration in the auditory system. Because of this reallocation, these resources are unavailable to more central cognitive processes such as the storage and retrieval functions of working memory, so that "upstream" processing of auditory information is adversely affected.

Murphy, Schneider, Speranza, & Moraglia (2006) performed three experiments in order to explore age-related changes in auditory processing capabilities. In each case, they compensated for the age-related decline of hearing by always using stimuli that were clearly audible to all participants. By doing so they could ensure that the performance of older adults on the tasks was not compromised by ubiquitous age-related losses in discriminative capacity (Schneider, 1997; Schneider & Pichora-Fuller, 2000). Their results gave no evidence of age-related decline.

Schneider, Daneman, Murphy, & Kwong (2000) investigated how younger and older adults comprehend and remember complex spoken passages in quiet and noisy listening situations. Younger adults (< 30 years of age) and older adults (> 65 years of age) with good hearing listened to passages in quiet or in noise. After listening to a passage, the participants answered a set of questions based on information presented in the passage. The passages were presented at the same sound pressure level (dB SPL) to all participants, and the noise was the same for all participants. The number of questions answered correctly decreased in the presence of noise for both age groups, but the older listeners answered fewer questions correctly than younger listeners. These results indicate that, independent of whether the passages are presented in quiet or are embedded in a background of babble noise; older adults are unable to recall as much detail as younger adults. They performed an additional experiment, where the S/N was adjusted so that identification of the spoken passages was equally difficulty for all individuals. The participants’ performance still decreased in background noise, but the age differences in performance disappeared. The interactions between the perceptual and cognitive factors that support language comprehension are an ongoing discussion, and the results by Schneider et al. (2000) indicate that a perceptual factor (noise) negatively affects memory of an auditorily presented text.

In summary, older people recall less information from the auditory information, even though identification is at equal level as younger participants. When auditory stimuli are easy to detect or when stimuli are written, no extra effort is required for perception, and no significant age effect on recall performance is found. But when the auditory stimuli are degraded, the effortful listening impairs memory performance more for the older than for the younger participants. The conclusion is that the poorer performance of the elderly is due to their decline of working memory capacity (Cohen, 1979, 1981). Obviously, degraded auditory stimuli lead to effortful listening for all, independent of the participants' age, but since the older people's pool of cognitive resources is smaller, they have fewer resources left for rehearsal and elaborative encoding.

Summary of speech comprehension

The consonants carry more information than vowels and therefore are more important for speech understanding. But when it comes to phonology the consonants are inferior to the vowels; we pronounce the least important parts (vowels) of the language with more acoustic energy. Another significant difference between vowels and consonants is their frequency range; the consonants are characterized by a higher frequency range than vowels. This combination is not fruitful for speech understanding, since the basic acoustics law called "upward spread of masking" suggests that lower frequencies are more effective to mask higher frequencies than the opposite. Furthermore, RTs are usually longer in the low frequency range, which means that phonemes with primarily low frequency energy will dominate the echo and interfere most with succeeding phonemes. With above facts in mind, it is logical to assume that room acoustics is an important issue for facilities where communication is of importance, such as schools and conference rooms. For older and people with hearing impairment it is even more important. This is also likely to be true when the spoken message is in the listener's non-native language.

In conclusion, there are some indications that it is not enough to hear what is said; listening must be unproblematic, otherwise learning and memory is likely to be hampered. Today's standards for acceptable S/N ratios and RTs in buildings do not consider the difference between hearing and remembering. The standards are solely based on a hearing criterion (Shield & Dockrell, 2003). Since the goal is to remember a message rather than solely hear it, it would be better to base the acoustic standards on memory criteria rather than on intelligibility criteria.

Research questions

Research concerning speech comprehension of degraded stimuli has had two directions, external stimuli degradation (e.g. noise and reverberation) and internal degradation (e.g.

hearing impairment). In order to explain research findings, we have to take both perceptual and cognitive factors into account. There are many studies of the ability to identify speech or words in unfavorable listening situations, usually expressed as a percentage of words or sentences correctly identified by a listener; this seems to be true for both external and internal degradation (Kryter, 1994). When it comes to cognitive effects of degraded stimuli some studies have been done with hearing impaired participants (Akeroyd, 2008). Regarding external environmental factors that affect the ability to hear and understand spoken information, there are not many published studies. Rabbitt (1966, 1968, 1991) has published some articles regarding low S/N ratio and memory performance, but his research method was not optimal. On the subject of acoustics, very little has been done concerning whether long RT affects cognitive functions like recall (Kjellberg, 2004). Thus, it is important to improve our knowledge about the relation between perceptual and cognitive factors when we listen to auditory information in environments with long RTs in order to make our acoustic standards more valid.

The aim of this thesis was to study if a low S/N ratio or a long RT affects one’s ability to identify, understand and remember auditory information. The study presented in paper I was designed to investigate if background noise affects memory performance (recall and recognition) of auditory presented word lists and sentences. A further aim was to study if the participants’ working memory capacity was related to the noise effect. The hypothesis was that recall of words should be better when they were presented without background noise. The recognition measure was expected to be less sensitive to noise than the recall measure, although background noise was expected to prolong response times. Regarding the relation between working memory and the expected noise effect on recall and recognition, it was assumed that the noise effect should be stronger for participants with low working memory capacity. The study presented in Paper II aimed to examine if a long RT affects recall and recognition of auditory word lists and sentences. A very similar paradigm with the same hypothesis as in paper I was used, but the independent variable (background noise) was replaced by a long RT. Paper III presented two studies; the aim with experiment one was to investigate whether background noise impairs the listener’s memory for a spoken lecture, even though it is possible to hear what is said. Experiment two investigated if reverberation has unfavorable effects on memory for spoken lectures. Paper IV contains two experiments with several substudies. Experiment one explored the effect of a long RT on serial recall of auditory stimuli. Experiment two dealt with three research questions. The first part aimed to investigate if ecological RT and ecological background noise affects free recall of auditory presented words. The next part of the experiment investigated if the effects of ecological RT interact with phonological familiarly in the stimulus material. Finally was studied if ecological RT and ecological background noise affects the listener's ability to remember spoken lectures.

Paper I

Purpose

The aim of the first paper was to investigate if recall of words and recognition of sentences orally presented was affected by a background noise. Secondly, the relationship between working memory capacity and performance in these conditions was studied. A further aim was to study how the noise affected the recall of different parts of the list of items-to-be-remembered by comparing the serial position curves with and without a background noise.

Method

The study included 32 participants, 23 women and 9 men with an age range of 18–34 years. All participants were native speakers of Swedish and all reported their hearing to be normal. A within-subject design was used with two conditions, with or without background noise (Noise and Control condition).

The experiment was conducted in a sound insulated climate chamber with the subject seated at a desk in the middle of the room. The speech had an equivalent sound pressure level of 64 dB(A). In the Noise condition the speech signal was mixed with a noise of 60 dB(A), thus giving a S/N ratio of 4 dB. Two sets of phonetically balanced word lists each with 50 monosyllable words and two sets of five word sentences were presented orally to the subjects. The task was to memorize the words for later recall and the sentences for later recognition. One set was presented without and one with the background noise, in a counterbalanced order. Directly after presentation the subjects were asked to write down on a paper all the words that they could recall. Working memory capacity was assessed with a reading span test taken from the cognitive test battery TIPS (Hällgren, Larsby, Lyxell, & Arlinger, 2001).

Results & Discussion

In line with the hypothesis subjects remembered less of the word list when it had been presented with a background noise. The effect of the noise was most pronounced for words in the beginning and the end of the word list. The results also showed that there was a significant association between working memory capacity and the effects of noise on recall of the recency part of the list. Recognition of sentences was expected to be less sensitive to noise but was found to be totally unaffected, and the performance of this task was unrelated to reading span. The hypotheses thus were confirmed regarding performance of the word recall task, but not for the sentence recognition task. The effect of the noise was apparent both on the recall of the primacy and recency part of the word list, which indicates that both the retrieval from a longer-term store and the short-term storage were impaired by the noise. The less efficient retrieval is probably the result of fewer resources or less time being left for encoding of the words in the noise condition where more resources are needed for word identification. In the same vein, the effect on the recency part of the list indicates that less resources or less time was left for rehearsal of the words in the noise condition. The absence of any effect in the middle part of the list was probably the result of a floor effect; so few of these items were remembered that there was little room for an effect. In the paired associates test used by Murphy et al. (2000) recall of the last of the to-be-remembered words was extremely easy, which probably explains that they, in contrast to the present study and Surprenant (1999), found no effect of the noise on the last part of the list.

Paper II

The aim of the second study was to explore if a long RT has the same effect on recall of spoken words as background noise was shown to have in a paper 1. A further aim was to study the influence of working memory capacity on performance in these conditions.

Method

The study included 32 participants, all were native speakers of Swedish and all reported their hearing to be normal. This was close to a replication of the study described in paper 1, but the independent variable was changed from noise/control to long RT/short RT. The experiment was conducted in an anechoic chamber with the participant seated in a chair in the middle of the room. The speech was presented by 12 loudspeakers placed in a circle around the subject. The stimulus material used in this study was the same as in paper I.

Two virtual classrooms were designed in CATT-Acoustics 8.0 software. All geometrical values were common for the two classrooms; both were the same size (length 10 m, width 6 m, height 3 m) and were furnished with 30 desks. The sound source was placed 1 m in front of the blackboard in the centre of the classroom at a height of 1.7 m, and the receiver was placed 6.6 m in front of the source at a height of 1 m. The classroom with short RT had various absorbing panels on the walls and the ceiling, and 30 pupils were seated in the desks. In the classroom with long RT only 15 pupils were seated and some absorbing panels were replaced by concrete walls. In the short RT condition, mean RT 0.25-4 kHz was 0.53 s (with max RT 0.58 s at 0.25 kHz) and in the long RT condition it was 1.17 s (with max RT 1.41 s at 0.125 kHz). The STI (Speech Transmission Index) values indicate that the short and long RTs should produce very good (73.5) and fair (56.1) intelligibility, respectively.

Results & Discussion

In line with the hypothesis the participants recalled fewer words when the word list had been presented with a long RT, and extra list errors were also more common in the long RT condition. The RT effect was pronounced at the beginning of the word list but was absent in the last part of the list. Recognition of sentences was, as expected, less sensitive to the long RT, but measurements of response time revealed faster responses when sorting out irrelevant sentences in the short RT condition. Contrary to expectations, reading span performance was unrelated to performance in both the word list and sentence task as well as to the effect of RT in these tests.

Paper III

In the third paper two experiments are reported. The aim with experiment 1 was to investigate whether broadband noise impairs the listener's memory for a spoken lecture, even though it is possible to hear what is said. Experiment 2 investigated if reverberation has unfavorable effects on memory for spoken lectures.

Method

Experiment 1 was conducted in a sound-insulated climate chamber with the participants seated at a desk in the middle of the room. They began by performing a hearing test followed by two tests of memory for spoken lectures, one presented with broadband noise (S/N +5 dBA) and one presented without broadband noise. After listening to the lecture, eight open-ended questions about the content should be answered. A within-subject design was used and the order between background noise and control condition was counterbalanced

Experiment 2 was conducted in an ordinary upper secondary school classroom. Each desk was rigged with a laptop, with attached headphones. Binaural recordings were presented to the participants through the headphones. The participants entered the classroom and sat down at their computers and performed the two lecture tasks. After listening to the lecture, twenty open-ended questions about the content should be answered. A within-subject design was used and the order between Reverberation and control condition was counter balanced.

Results and discussion

The results from Experiment 1 are consistent with earlier investigations (Rabbitt, 1966, 1968) and indicate that background noise in classrooms may be detrimental to listeners' memory for lectures even when they are able to hear what is said. In Experiment 2 the participants' memory performance was worse when the lecture was heard with a long RT than in the short RT condition. Taken together, these two studies show that background noise and long RT are detrimental to memory of spoken lectures, even when the listeners are able to identify the speech. These results are consistent with previous investigations into the effects of noise on memory for spoken word lists (Kjellberg et al., 2008; Pichora-Fuller et al., 1995; Surprenant, 2007) and for spoken prose (Rabbitt, 1968) and lead us to argue that acoustic standards for rooms meant for learning should be more stringent than previously suggested. The results suggest that poor listening conditions (resulting from background noise and/or long RT) impair memory and learning, even if the conditions allow the listeners to hear what is said. Since the goal for students and pupils attending lectures is to remember the lecture rather than just hearing what is said, the results presented here indicate that standards for acceptable S/N ratios and RTs in buildings designed for learning should consider the distinction between speech intelligibility and memory. Standards should be based on memory criteria instead of intelligibility criteria.

Paper IV

The aim of Paper IV was to examine if degraded listening conditions affect serial recall and/or free recall of auditorily presented stimuli. Experiment 1a) examined if long RT disrupts serial recall of meaningful words, and Experiment 1b) studied if long RT and low S/N ratio impair performance of free recall of a word list. Experiment 2 examined a possible interaction between types of words in the list (words with many or few phonological neighbors) and degraded listening condition (short or long RT). Furthermore, the effects of degraded listening condition and stimulus type were analyzed with regard to the participants working memory capacity.

Method

Experiment 1a had a repeated measure design with two factors: RT (three levels: 1.3 sec, 0.7 sec, and 0 sec) and Serial Position (eight levels). The participants listened to ten word lists in each RT condition, followed by a serial recall test. In Experiment 1b the participants were presented with nine word lists each containing 11 words (three lists heard in short RT, three with background noise, and three in long RT), with recall directly after each presented list. To obtain ecologically valid school recordings, the word lists were recorded binaurally in two ordinary classrooms (one with long RT and one with short). These binaural recordings were presented to the participants through headphones. In Experiment 2 the to-be-remembered material was 16 word lists with seven nouns in each list. The words in eight of these lists had many phonological neighbors, and eight lists had few phonological neighbors. The participants listened to 16 word lists, 8 lists in long RT and 8 lists in short RT. All participants in Exp 1b) and Exp 2 performed a working memory test (operation span) in silence before the main experiment session. All effects of degraded stimuli or list type were analyzed with regard to the participants' working memory capacity.

Results and discussion

The purpose of the series of experiments reported here was to investigate why poor listening conditions (low S/N and long RT) impair memory for spoken materials. Long RT did not disrupt serial recall of meaningful words (Experiment 1a), but it did disrupt free recall (Experiments 1b and 2). Moreover, more detailed analyses revealed that words presented in the beginning of the list were more impaired by the unfavorable listening conditions, which is in agreement with earlier findings. Furthermore, listeners who made many invention errors in operation span produced significantly more false recalls in low S/N ratio (Experiment1b). The same pattern was shown in Experiment 2; where the effect of phonological neighbor word lists was significantly related to the number of inventions made in the operation span test. This finding gives support for the hypothesis that degraded stimuli activate more candidates (Cluff & Luce, 1990), and people who make such specific errors have problem to suppress those candidates.

General Discussion

The aim of the first two experiments described in this thesis was to establish an effect of degraded stimuli on memory performance. In Experiment 1 (Paper I) a low S/N ratio was used to degrade the spoken messages. The degradation was enough to make listening more effortful for the listener, without making it impossible to identify the words correctly. The results showed that background noise impaired free recall of spoken words despite correct identification. Furthermore, the noise effect was negatively correlated to the participant's working memory capacity. This was true for items presented in the mid and recency parts of the word list. In Experiment 2 (paper II) long RT was used to make listening more demanding. The same pattern was shown; recall performance was worse when the participants had heard the words in the long RT condition. However, no significant relation to working memory capacity was found in this experiment.

From an ecological point of view, word lists and single sentences are not common in our everyday life. Longer speech and lectures are more usual, and the main purpose in the experiments described in paper III was to investigate if a poor acoustic environment affects understanding and memory of a lecture. Two experiments examined if background noise (Experiment 1) and long RT (Experiment 2) affect later memory performance. Both experiments showed that poor acoustics have negative effects on the participants’ memory for spoken lectures. The aim with Paper IV was to explore the relation between working memory capacity and the effect of degraded listening conditions. The first experiment (Exp 1a) revealed that long RT does not affect serial recall, which was in line with the hypothesis. Experiment 1b explored how low signal- to-noise ratio and long RT affects the different positions of the word lists, and the relation between working memory capacity and the effect of degraded listening. Experiment 2 investigate two independent variables; long vs. short RT and word list with or without phonological neighbors, with the hypothesis that words with many phonological neighbors should activate more phonological candidates in our mental lexicon and therefore impair correct recall. The predicted effect was shown in the mid positions of the word list.

In many workplaces speech communication is essential for good performance, and therefore room acoustics are of importance to make sure that work can be carried out efficiently. Two acoustic conditions may impair the ability to perceive speech: the relationship between the speech and background noise (S/N ratio) and the room's RT (Rossing, 1990). Many experiments have shown that speech signals are clearer and easier to perceive in rooms with short RT as long as the S/N ratio is constant. In most of these studies, the effect of RT has been assessed by measuring how well the participants succeeded in identifying individual words or sentences. However, the traditional word identification tests reflect the total effect of the impaired listening situation on word identification, and do not specify the effects on different stages of processing. To my knowledge no published paper has reported the implications of long RT on the listener's cognitive load and memory performance. The hypothesis that degraded listening conditions (e.g. too long RT or low S/N ratio) lead to poor understanding and decreased memory performance even though correct identification is possible are supported by the experiments presented in this thesis.

In summary, this thesis has shown that an acoustical condition should be evaluated based on memory criteria, not only on identification terms. It is not enough to hear what is said, one has to hear the message without effort to achieve appropriate learning. These findings are of interest when building standards are updated and as a guideline when buildings where communication is of importance are built or renovated. Furthermore, several of the experiments described in this thesis have shown that people with low working memory capacity (measured by complex span measures, invention errors or prior-list errors) are more susceptible to the stimulus degradation. Since working memory capacity is highly correlated with nearly every cognitive task, it is logic to say that people that perform poorly on cognitive tasks in general also are more affected by degraded listening conditions (Engle, 2002).

Remarks for future research

All studies presented in this thesis have been performed with young adult participants. All of them had Swedish as the first language and they had normal hearing status. There are reasons to believe that this particular age group is able to handle demanding listening conditions better than e.g. older people and young children, since older people have declining hearing and young children have limited language knowledge in general. Thus, it would be interesting to perform additional studies with those particular age groups. From an ecological point of view a study with young school children is preferable, since good learning is essential for their school achievement. Such studies should also include children with hearing loss and with other language background than Swedish, thereby illustrating the interaction between external and internal degradation of the spoken message. Furthermore, many school buildings have rather poor acoustic design, thus degraded stimuli is a part of school children’s everyday life.

Sammanfattning på svenska

Denna avhandling behandlar effekterna av lågt signal-brus-förhållande eller en lång efterklangstid för minne och lärande. Alla studier har använt auditivt stimuli material (ordlistor och texter) som har presenterats över tröskeln för taluppfattning, men under tillräckligt dåliga förhållanden för att göra lyssnandet mer kognitivt krävande. Den grundläggande hypotesen för hela projektet var att en dålig ljudmiljö ökar den kognitiva belastningen för lyssnaren, vilket försämrar minnet av presentationsmaterialet trots att det har identifierats korrekt. I Artikel I användes ordlistor och meningar som minnesmaterial, dessa fick försökspersonerna höra med och utan bakgrundsbrus. Studien visade att återgivningen var signifikant sämre när orden presenterats med bakgrundsbrus. Artikel II var en replikering av studien i Artikel I, med den enda skillnaden att den oberoende variabeln bakgrundsbrus byttes ut till lång efterklangstid vilket gav samma resultat. Försökspersonerna mindes alltså orden sämre när de hade hört dem i lång efterklangstid. Artikel III inkluderade två experiment där minnesmaterialet var ca 10 minuter långa texter. Hypotesen var att texter skulle påverkas på samma sätt som visats på ordlistor i Artikel I och Artikel II. Studie ett undersökte om lång efterklangstid påverkar minnet av en hörd text, och studie två prövade om bakgrundsbrus hade någon påverkan på minnet av texten. Hypotesen bekräftades; minnet av texterna försämrades i båda experimenten. Artikel IV inkluderar två studier, det första experimentet (1a) undersökte huruvida lång efterklangstid påverkar vår förmåga att återge seriell information. I experiment 1b presenterades ordlistor med lång efterklangstid eller med bakgrundsljud samt i en betingelse utan bakgrundsbrus och kort efterklangstid (kontroll). De stimuli som användes spelades in i ordinära klassrum, ett klassrum med extremt bra ljudmiljö och ett med mycket dålig akustisk design. I experiment 2 användes ordlistor med många eller få fonologiska grannar, dessa ordlistor presenterades med lång eller kort efterklangstid.

De samlade resultaten kan sammanfattas i två meningar: Att höra vad som sägs är en nödvändig men inte tillräcklig förutsättning för att komma ihåg vad som sagts, vilket innebär att talad information ska höras utan särskild ansträngning för att erhålla god inlärning. Detta är något som bör tas i beaktning vid utvärderingar av lokaler där kommunikation och inlärning är av central betydelse.

References

Akeroyd, M. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impair. *International Journal of Audiology, 47,* 125–143.

Arlinger, S. (2007). *Nordisk lärobok i audiologi*. First Edition, C A Tegnér, Bromma.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4,* 417–423.

Baddeley, A. D. (2001) Is working memory still working? *American Psychologist, 56,* 851–864

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–90). New York: Academic Press.

Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589

Baddeley, A.D. (2007). *Working memory, thought, and action*. Oxford University Press, Oxford.

Beaman, C., Jones, D. M. (1998). Irrelevant sound disrupts order information in free recall as in serial recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology. 51, 3,* 615-636.

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *Journal of the American Acoustic Society, 123,* 2078–2086

Cluff, M. S., & Luce, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 551-563.

Cohen, G. (1979). Language comprehension in old age. *Cognitive Psychology, 11,* 412-429.

Cohen, G. (1981). Inferential reasoning in old age. *Cognition, 9,* 59-72.

Conrad, R. and Hull, A.J. (1964) Information, acoustic confusion and memory span. *British Journal of Psychology* 55, 429–432

Cowan, N. (2005). *Working memory capacity.* New York, USA: Psychology Press.

Crandell, C., & Smaldino, J. (2000a). Classroom Amplification Technology: Theory and Practice. *Language, Speech, and Hearing Services in Schools; 31*, 4, 371-375.

Crandell, C., & Smaldino, J. (2000b). Classroom acoustics for children with normal hearing and with hearing impairment.*Language, Speech, and Hearing Services in Schools, 31,* 4 362-370.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19-23.

Evans, G. W., & Lepore, S. J. (1993). Nonauditory effects of noise on children: A critical review. *Children´s Environments 10,* 31-51.

Everest, F. A. (2001). *Master handbook of acoustics*, Fourth Edition, McGraw-Hill.

Eysenck, M. W., & Keane, M. T. (2005). *Cognitive psychology: A student's handbook*. Psychology Press (5th ed.). New York.

Haines, M. M., Stansfeld, S. A., Job, R. F. S., Berglund, B., & Head, J. (2001). Chronic noise exposure, stress responses, mental health, and cognitive performance in school children. *Psychological Medicine, 31,* 265-277.

Heinrich, A.; Schneider, B, A. & Craik, F, I. M. (2008). Investigating the influence of continuous babble on auditory short-term memory performance. *The Quarterly Journal of Experimental Psychology. Vol 61),* May 2008, pp. 735-751.

Hällgren, M., Larsby, B., Lyxell, B. and Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons, *Journal of the American Academy of Audiology, 12,* 357-370.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47-70.

Kjellberg, A. (2004). Effects of reverberation time on the cognitive load in speech communication: theoretical considerations *Noise & Health, 7*; 11-21

Kryter, K. D. (1994). *The handbook of hearing and the effects of noise: Physiology, psychology, and public health*. Academic Press, CA, US, San Diego.

Liberman, A. M., Cooper, F. S., Shankweiler, D. S., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431-461.

Luce, P. (1986). *Neighborhoods of words in the mental lexicon* (Research on Speech Perception Tech. Rep. No. 6). Bloomington, IN: Speech Research Laboratory, Department of Psychology, Indiana University.

Lustig, C., Hasher, L., & Zacks, R. T. (2008). Inhibitory deficit theory: Recent developments in a “new view”. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in*

Cognition (pp. 145-162). Washington, DC: American Psychological Association.

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

Marslen-Wilson, W. D. (1990). *Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.

Marslen-Wilson, W.D., &Warren, P. (1994). Levels of Perceptual Representation and Process in Lexical Access: Words, Phonemes, and Features. *Psychological Review Vol. 101, 4*, 653-675

McClelland, J. L., & Elman, J. L. (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

Morrison, C. M., & Ellis, A. W. (1995) Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 116-133.

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and Aging, 15*, 323-334.

Murphy, D. R., Schneider, B. A., Speranza, F., & Moraglia, G. (2006). A comparison of higher order auditory processes in younger and older adults. *Psychology and Aging 21, 4*, 763–773

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

Nelson, P. & Soli, S. (2000). Acoustical barriers to learning: Children at risk in every classroom. *Language, Speech and Hearing Services in Schools, 31*, 356-361.

Norman, D.A., & Shallice, T. (1986). Attention to action. In: Davidson, R.J., Schwartz, J.E., Shapiro, D. (Eds.), *Consciousness and self-Regulation, vol. 4.* Plenum Press, New York, pp. 1–18.

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

Rabbitt, P. M. A. (1966). Recognition memory for words correctly heard in noise. *Psychonomic Science, 6,* 383-384.

Rabbitt P. M. A. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology 20*: 241-248

Rabbitt, P. M. A. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngologica, Supplementum, 476,* 167–176.

Repovs, G., & Baddeley, A. D (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139*, 5-21.

Rossing, T. D. (1990). *The science of sound*. Reading,Mass. Addison-Wesley.

Rudner, M., Foo, C., Sundewall-Thorén, E., Lunner, T. & Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *International Journal of Audiology, 47*, 91–98.

Rudner, M., Foo, C., Rönnberg, J. & Lunner, T. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavian Journal of Psychology, 50,* 405–418.

Schneider, B. A. (1997). Psychoacoustics and aging: Implications for everyday listening. *Journal of Speech-Language Pathology and Audiology, 21,* 111-124.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of cognitive aging* (2nd ed.; pp. 155-219). Mahwah, NJ: Erlbaum.

Schneider, B. A., Daneman, M., Murphy, D. R., & Kwong See, S. (2000). Listening to discourse in distracting settings: The effects of aging. *Psychology and Aging, 15,* 110–125.

Shield, B. M., & Dockrell, J. E. (2003). The effects of noise on children at school: A review. *Building Acoustics, 10*, 97-116.

Shield, B. M., & Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *Journal of Acoustical Society of America. 123*, 133–144.

Smith, A. (1989). A review of the effects of noise on human performance. *Scandinavian Journal of Psychology, 30,* 185–206.

Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables, *International Journal of Psychology, 34,* 328-333.

Surprenant, A. M. (2007). Effects of noise on identification and serial recall of nonsense syllables in older and younger adults, *Aging*. *Neuropsychology and Cognition, 14*, 126-143

Stenfelt, S., & Rönnberg, J. (2009). The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology, 50*, 385–393

Sörqvist, P., Ljungberg, J. K., & Ljung, R. (2010). A sub-process view of working memory capacity: evidence from effects of speech on prose memory, *Memory, 18*, 310 – 326.

Was, C. A., & Woltz, D. J. (2007). Reexamining the relationship between working memory and comprehension: The role of available long-term memory. *Journal of Memory and Language, 56*, 86-102.

Paper I

APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. (2007)

(www.interscience.wiley.com) DOI: 10.1002/acp.1422

Recall of Words Heard in Noise

ANDERS KJELLBERG*, ROBERT LJUNG and DAVID HALLMAN

Centre for Built Environment, University of Gävle, Gävle, Sweden

SUMMARY

The aim of the study was to explore if recall of words and recognition of sentences orally presented was affected by a background noise. A further aim was to investigate the role of working memory capacity in performance in these conditions. Thirty-two subjects performed a word recall and a sentence recognition test. They repeated each word to ensure that they had heard them. A reading span test measured their working memory capacity. Performance on the word recall task was impaired by the background noise. A high reading span score was associated with a smaller noise effect, especially on recall of the last part of the word list. Copyright © 2007 John Wiley & Sons, Ltd.

The effects of noise on memory, learning and other cognitive functions have been widely researched (see e.g. reviews by Jones, 1990 and Smith, 1989). With few exceptions this research has used written texts or other visual material. Very little has been done concerning cognitive effects of noise on speech communication, except for studies on speech intelligibility, usually expressed as a percentage of words or sentences correctly identified by a listener (for a review see e.g. Kryter, 1994). The obvious reason for using text based tasks in studies of other cognitive functions is that impaired performance in noise could otherwise be an effect of noise making it impossible to hear parts of what is said. However, efficient speech communication requires further processing and it seems likely that noise may affect speech communication even when the noise allows identification of the words spoken. In addition to identifying the words spoken the listener must be able to integrate the current information with information that has recently been processed as well as with stored information. Furthermore, to be able to use the information it must be transferred to a more permanent store. Processing speech in unfavourable listening conditions, for example, with a high background noise, may require a listener to consider more alternative interpretations of the speech stimuli and speech understanding may rely more on stored information for successful performance than in good listening conditions. The perceptual coding of the speech signal, which in good listening conditions is largely automatic, becomes more of a controlled resource-demanding process. Under such conditions the identification of the spoken words thus places higher demands on the limited cognitive processing resources. If so, a background noise could leave fewer resources for the further processing of the speech. Understanding and memory of spoken information then may be impaired by the unfavourable listening conditions even

---

*Correspondence to: Anders Kjellberg, Centre for Built Environment, University of Gävle, SE 801 76 Gävle, Sweden. E-mail: akr@hig.se

when every word is heard correctly (for further discussions of these issues see Kjellberg, 2004).

In a similar way, Rabbitt (1966) argued that the extra effort required for identifying words in noise would reduce the spare capacity left for the further processing of the speech. In two experiments he studied recognition of a series of 100 words, which had been presented with or without a background noise, and got some support for his hypothesis. To check that they had identified the words correctly, subjects had to shadow their presentation. He found no effect of the noise on the number of correctly recognized words but an increased number of false alarms, leading to a lowering of the signal detection parameter *Beta*. Several features of Rabbitt's design may have reduced the effects in these experiments. The task may become more sensitive to the detrimental effects of noise if the recognition task is substituted with the more difficult recall task. If the recognition task is used it is also likely that response times would be more sensitive than the number of recognized words; even if one reaches a correct decision about a word, a less efficient encoding of the word list may prolong the matching process necessary to reach this decision. Another feature of Rabbitt's study that may have contributed to the relatively weak effect was that he excluded all subjects who did not identify all words when they were presented. This means that the selected group consisted of the persons that best managed to hear words in noise. An alternative way to control for the effect of the noise on word identification would be to use the percentage recalled of the correctly identified words as a performance measure. Another possibility would be to disregard whether the words were correctly identified when presented and instead use the subject's interpretation of these words as the items to be remembered. In these ways, the subjects for whom the task required most effort would also be included in the study group. In the present study, these two scores were calculated in a task measuring recall of a long word list.

Rabbitt's interpretation of his results means that this effect should not be limited to noise. All types of degradations of the speech signal should have the same effect. This was supported by a later study (Rabbitt, 1991), where persons with a mild hearing impairment remembered fewer words than those with a normal hearing although there was no difference between the groups in the number of correctly heard words. Murphy, Craik, Li and Schneider (2000) argued from a similar point of view that the decline of memory among the elderly and the effect of noise on memory both may at least partly be the result of degraded sensory information. They showed that a background noise made young adults perform like older adults in a paired associates task. Both aging and noise primarily affected the primacy part of the serial position curve, thus indicating that the short-term storage of the last items of the series was relatively unaffected. However, there was no control for the number of correctly heard words in these experiments, and the effect of noise on recall may at least partly be a result of fewer words being correctly identified in this condition. Rabbitt (1968) also studied the effect of a background noise on serial recall of a series of eight digits where there was a short pause between the first and second group of four digits. After the presentation of a series subjects were asked to recall either the first or the second group of four digits in correct order. A background noise was added either to one or both of the two parts both parts of the list, or to none of them. The recall of the first group was found to be better when it was followed by a group presented without noise than with noise. Rabbitt's interpretation was that the increased processing demands caused by the noise inhibited the rehearsal of the first four digits. This study differed from Rabbitt (1966) in two important aspects: what was tested was recall of different permutations of the order of a number of known items and the number of items was within the short-term memory span, which made it possible for the subject to rehearse the whole list during the presentation. In a more recent study, a similar effect of background noise was demonstrated on the serial recall of a list of two-phoneme nonsense syllables (Surprenant, 1999). The effect of noise on word identification was tested separately to examine the possibility that differences in recall could be explained simply by the reduction in the number of words correctly heard in the noise condition. With one exception the effect of noise on syllable identification was not significant. However, the possibility that there was such an effect when the lists were presented, or that the noise led to missed items by momentarily distracting the subject's attention cannot be excluded.

An important feature of Surprenant's study was that she analysed the effect of noise on the serial position curve. When the signal-to-noise (S/N) ratio was 5 dB the recall of all items of the series was equally impaired by the noise, whereas an effect of the 10 dB S/N condition only was apparent in the primacy part of the serial position curve. This analysis was made to differentiate between two explanations of a memory effect, the capacity model and the dual-trace model. The capacity model is that invoked by Rabbitt (1966). The dual code model suggests that impaired recall is an effect of degraded stored physical representation. In a task like Rabbitt's (1966) where subjects repeat each item aloud the dual code model seems to give a less likely explanation of the memory impairment. The repetition of the word should reduce or eliminate possible differences in the quality of the stored physical information. Furthermore, in the serial recall tasks it was possible to rehearse the whole list, which was impossible with the long lists used by Rabbitt (1966). Still, a serial position analysis of the free recall of a long word list could shed light on the nature of the effect of noise on recall. An effect on recall of the final part of the list (the recency effect) would indicate that the noise had interfered with some form of short-term storage of these words. If noise affected recall of the first part of the list (the primacy effect) this would indicate that the noise had made retrieval from a longer-term store more difficult, probably by having impaired the encoding of the words.

If the critical effect of the noise is that more resources must be allocated to the word identification task, a person's available resources should be of importance for understanding and remembering speech presented under impoverished listening conditions as well as for persons with a hearing impairment. The resources involved are generally dealt with in terms of working memory capacity most often measured with a working memory span test (Daneman & Carpenter, 1980), which requires simultaneous processing of a series of sentences and storing for later recall of last or first words of these sentences. Hällgren, Larsby, Lyxell, and Arlinger (2001b) showed that reading span performance was correlated with speech perception in a difficult listening task (dichotic listening), although the correlation became nonsignificant when age was partialled out. Pichora-Fuller, Schneider, and Daneman (1995) used a reading and a listening span task and found age differences in recall of the last words in the listening but not in the reading version of the task. With a noise that lowered the S/N ratio was set to 5 dB recall was impaired when the memory load was high and when word identification got no support from preceding words in the sentence. They interpreted their results as an indication of less resources being left for the further processing when the speech signal is degraded as it is for the hearing impaired and when the S/N ratio is unfavourable.

In the present study, free recall of long word lists was performed with and without a background noise, and working memory capacity was tested with a test of reading span. In addition a sentence recognition task was performed in the same conditions. This task was included as a less resource demanding memory task that also allowed

DOI: 10.1002/acp

the measurement of response times. The objective was to test the following three hypotheses:

- Recall of words is better when they are presented without background noise.
- Recognition of sentences is less sensitive to the noise than the recall of words, but background noise prolongs response times.
- The expected noise effect on recall and recognition, will be weaker for subjects with high working memory capacity.

A further aim was to analyse how the noise affected the recall of different parts of the list of items-to-be-remembered by comparing the serial position curves with and without a background noise.

METHOD

Participants and design

The study included 32 participants, (23 women and 9 men with an age range of 18–34 years). All participants were native speakers of Swedish and all reported their hearing to be normal. A within-subject design was used with two conditions, with or without background noise (Noise and Control condition). The order between conditions was counterbalanced.

Apparatus

The experiment was conducted in a sound isolated climate chamber (20°C) with the subject seated at a desk in the middle of the room. The speech and noise were presented with loudspeakers 1.5 m in front of the subject at an angle of 45 degrees. A Numark DCMIX1 CD mixing console was used to mix the speech and noise. For visual presentation (reading span) and recognition responses the display of a laptop was used.

The speech and noise

The speech stimuli and the noise were part of a package of standardized tests for speech audiometry (Hagerman, 1982). The speech had an equivalent sound level of 64 dB(A). In the Noise condition the speech signal was mixed with a noise of 60 dB(A), thus giving a S/N ratio of 4 dB. The noise was a broadband noise synthesized from the speech material to produce the same spectrum as the speech. It was also amplitude modulated by a low frequency noise to make it sound more natural. In the Control condition the background noise in the room gave a S/N ratio of 27 dB.

Memory tasks

Recall of words

Two phonetically balanced word lists each with 50 one-syllable words were presented orally to the subjects. The lists are a part of a package of standardized for speech audiometry tests (Hagerman, 1982). The task was to memorize the words for later recall. There were approximately 3 seconds of silence between each presented word when the subjects repeated the word aloud to check that they had identified it. One list was presented

DOI: 10.1002/acp

without and one with the background noise, with a counterbalanced order. Directly after presentation the subjects were asked to write down on a paper all the words that they could recall. Recall performance was measured in two ways. The first was the number of words correctly recalled of the words that they had stated when the list was presented irrespective of whether the word had been correctly identified or not (stated words). The second recall measure was the percentage words recalled of those that had been correctly identified (correct words).

Recognition of sentences

The Hagerman test used for assessing recognition of sentences is a part of a package of standardized speech audiometry tests (Hagerman, 1982). It contains lists of spoken Swedish sentences. The sentences all contain five words and have an identical structure (name, verb, number, adjective, noun) but within this structure the words are not predictable (e.g. Kim bought six white pillows). Two lists with 10 sentences each were used. Both lists contained exactly the same words but combined in different ways. One list was presented in the Noise and one in the Control condition in a counterbalanced order. There was approximately 7 seconds of silence between each sentence; the subjects repeated each sentence aloud to check that they had heard it correctly. The subjects' task was to memorize the orally presented sentences for later recognition. Directly after presentation the subjects were shown a series of 20 sentences, 10 of which had been presented previously. The task was to determine if the sentence was one of the old ones or a new one. To make the task easier the new sentences contained one word that did not appear in any of the original sentences. The number of correct answers and response times (including reading time) were measured. The alternative lists of words and sentences have previously been shown to be equally intelligible in terms of the number of correctly identified items (Hagerman, 1982; Magnusson, 1995).

Working memory capacity (reading span)

Working memory capacity was assessed with a reading span test taken from the cognitive test battery TIPS (Hällgren, Larsby, Lyxell, & Arlinger, 2001a). The subject's task was to decide whether the sentences were absurd or normal and to recall either the first or the final words of these presented sentences. The words were presented in a word-by-word fashion. Each word was shown on the screen for 0.8 seconds. The inter word interval was 0.075 seconds. Half of the sentences were absurd (e.g. 'the house read a newspaper'), and half were normal (e.g. 'the pupil came too late'). The subjects' task was to indicate, during a 1.75 seconds interval, whether the sentence was a normal or absurd sentence by pressing a key on the keyboard (the number of incorrect answers was extremely small, and could therefore not be used as an indicator of working memory capacity). After a sequence of sentences (three, four, five or six sentences), the experimenter indicated that the subject should start to report orally as many as possible of either the first or the final words of three to six just presented sentences. The subjects did not know beforehand if they should report the first or the last words. The number of correctly recalled words was used as the performance measure.

Rated effort

To validate the assumption that word identification became more effortful by the noise subjects rated the effort required to follow the speech using Borg's CR10 scale (Borg, 1998). This was done directly after the presentations of the word and sentence lists. The scale has range of 0–10 with verbal label on eight steps. The scale values of the verbal labels have been chosen with the aim to approximate ratings at a ratio scale level.

Procedure

All subjects first performed the Reading Span Test followed by auditory recall and recognition tests in the two conditions (the order between the two tests was counterbalanced).

The experimental sessions lasted for approximately 40 minutes and were conducted between 9 AM and 4 PM. At the outset subjects were informed that the study was about memory.

RESULTS

As a check of the noise effect on the difficulty of the task, the mean of self-reported effort and number of incorrectly repeated words and sentences were calculated. A two-way ANOVA (noise condition × order of conditions, i.e. noise-control vs. control-noise) revealed that both effort and the number of incorrectly repeated words and sentences were significantly higher in the noisy condition. There was no significant effect of the order between conditions or any significant interaction between conditions and order for any of these measures (Table 1).

Recall of words

Two two-way ANOVAs (condition × order of conditions) were performed on recall performance. The number recalled of correct and stated words were almost perfectly correlated (.992 and .997 in the Noise and Control condition, respectively). Therefore, only the analyses of correctly recalled stated words are reported. The number of words correctly recalled was significantly lower in the noisy condition (mean = 8.50 and 11.03 for the Noise and Control condition, respectively, $F(1,30) = 17.28$; $p < 0.01$, $\eta^2 = 0.365$). Neither the effect of order between conditions nor the interaction between conditions and order was significant ($p = 0.22$ and 0.97, respectively).

In order to explore the serial position effect the word lists were split up in three parts (first 10, middle 30, last 10). The reason for collapsing second, third and fourth groups of 10 words was that in these three groups many subjects (between 16 and 28 per cent) did not recall any item correctly. Further, there was no significant difference in the number of

Table 1. Mean values (standard deviations) of effort (Borg, 1998) and incorrectly repeated words and sentences in the Noise and Control conditions and results from analysis of variance of the effect of conditions

	Noise *M* (SD)	Control *M* (SD)	*F*	*p*	η^2
Effort-word	3.94 (1.76)	2.31 (1.48)	47.08	<0.001	0.61
Effort-sentences	4.82 (2.45)	2.32 (1.57)	31.06	<0.001	0.50
Incorrectly repeated words	3.34 (1.21)	1.00 (1.05)	110.71	<0.001	0.78
Incorrectly repeated sentences	0.63 (0.98)	0.19 (0.47)	4.76	0.04	0.13

recalled items between these three parts. Neither was there any significant difference between the noise and control conditions in these parts of the list. By collapsing these 30 words into one group a less extremely skewed distribution of recall scores was obtained. As shown by Figure 1, the noise effect had its major influence in the first and last part of the list. This was reflected in an interaction between condition and parts in the quadratic trend ($F(1,30) = 6.99$; $p = 0.013$, $\eta^2 = 0.189$).

Recognition of sentences

A two-way ANOVA (condition × order of conditions) was performed to explore the noise effect on recognition of Hagerman's sentences. The analysis showed no significant difference between the conditions, except for a significantly shorter response time for misses in Noise than in the Control condition (Table 2). There was no effect of presentation order or interaction between order and condition for any of these measures.

Relation between working memory and noise effect

Correlations were calculated for the relation between reading span score and recall score overall and for the three parts of the word lists (primacy part, midpart and recency part). As shown by Table 3 most of these correlations were significant with the strongest correlation for the overall score. However, none of the differences between correlations was significant. To test whether the effect of noise was related to working memory capacity a regression analysis was performed with recall of stated words in Noise as the dependent variable and performance in the Control condition and reading span performance as predictors. As shown by Table 3, a positive regression coefficient was obtained in all these analyses, indicating that a higher reading span score was associated with a smaller effect of noise. This association was strongest for the recency part, but was also significant for the

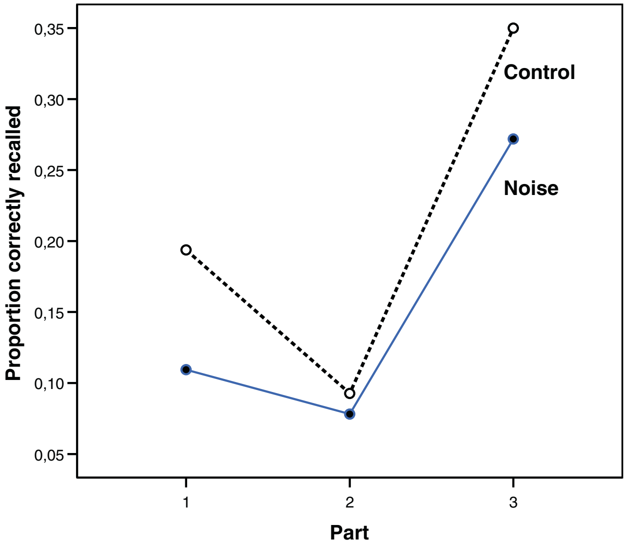


Figure 1. Proportion correct recalled words in the three parts of recall test in the Noise and Control conditions

Table 2. Mean value (standard deviation) for recognition of sentences in the Noise and Control conditions and results from ANOVA where consideration also was taken to presentation order

	Noise *M* (SD)	Control *M* (SD)	*F*	*p*	η^2
No. correct	16.34 (1.66)	16.38 (1.90)	0.01	0.94	<0.001
No. misses	2.47 (1.14)	2.34 (1.38)	0.16	0.69	0.005
No. false alarm	1.19 (1.18)	1.28 (1.25)	0.14	0.71	0.005
Response time (second)					
Correct responses ($n = 32$)	2.20 (0.60)	2.30 (0.56)	2.66	0.11	0.082
Misses ($n = 28$)	2.57 (0.83)	3.12 (1.26)	6.84	0.02	0.208
False alarms ($n = 18$)	2.82 (0.75)	2.71 (1.45)	0.14	0.72	0.008

Note: In the analyses of response time for misses and false alarms only subjects with these errors could be included.

midpart of the list. However, none of the differences between regression coefficients was significant. A corresponding analysis was made of the sentence recognition test where no significant relations were found.

A similar analysis was made of the relation between the effect of the noise on effort ratings and the recall of words. No such relation was found.

DISCUSSION

In line with the hypothesis subjects remembered less of the word list when it had been presented with a background noise. The effect of the noise was most pronounced for words in the beginning and end of the word list. The results also showed that there was a significant association between working memory capacity and the effects of noise on recall of the recency part of the list. Recognition of sentences was expected to be less sensitive to noise but was found to be totally unaffected, and the performance of this task was unrelated to reading span. The hypotheses thus were confirmed regarding performance of the word recall task, but not for the sentence recognition task.

The effect of the noise was apparent both on the recall of the primacy and recency part of the word list, which indicates that both the retrieval from a longer-term store and the short-term storage were impaired by the noise. The less efficient retrieval is probably the

Table 3. Correlations between reading span score and score in the test of recall of words in the Noise and Control conditions, overall and for the three parts of the word list

	Correlation with reading span score		Relation of reading span score to the noise effect		
Reading span score	r_{Control}	r_{Noise}	β^{a}	*t*	*p*
Overall score	.45**	.52**	.26	1.80	0.08
Primacy part	.38*	.30	.20	1.07	0.29
Midpart	.36*	.47**	.36	2.15	0.04
Recency part	.29	.50**	.44	2.68	0.01

*$p < 0.05$.

**$p < 0.01$.

[a] Standardized regression coefficients for reading span score as predictor of recall score in noise after control for the score in the Control condition.

result of fewer resources or less time being left for encoding of the words in the noise condition where more resources are needed for word identification. In the same vein, the effect on the recency part of the list indicates that less resources or less time was left for rehearsal of the words in the noise condition. The absence of any effect in the middle part of the list was probably the result of a floor effect; so few of these items were remembered that there was little room for an effect. In the paired associates test used by Murphy et al. (2000) recall of the last of the to-be-remembered words was extremely easy, which probably explains that they, in contrast to the present study and Surprenant (1999), found no effect of the noise on the last part of the list.

Rabbitt (1966) did not find any effect on the number of recognized words. The present experiment used a lower S/N ratio and a more difficult memory task (recall instead of recognition) and less selected subjects, and these differences obviously strengthened the noise effect.

In contrast to Rabbitt's study there was no effect of noise on the number of false alarms or any other negative noise effect in the sentence recognition task. The lists of sentences were much shorter than the word lists in Rabbitt's experiment. Further, all sentences had the same structure, which meant that the subjects knew, for example, that the fourth word in the sentence was an adjective, although the sentence was nonredundant in the sense that it was impossible to predict which adjective would come. The absence of any effect on sentence recognition therefore probably was the result of the task being too easy.

The noise effect was expected to be less severe for persons with a better working memory capacity as defined by their reading span performance. This hypothesis got its strongest support from the noise effect on the recall of the recency part of the word list. The effect on the primacy part of the list was unrelated to reading span performance, although working memory resources generally are supposed to be used also for encoding of the words. Thus, the results seem incompatible with an interpretation of reading span performance as an indicator of general working memory capacity. Alternatively, working memory capacity is not critical for the encoding of words. Two further aspects of the relation between reading span performance and the recall scores could be noted. First, the reading span and recall scores for all three parts of the list were correlated both in the noise and the control conditions, which supports the interpretation of reading span as an indicator of general working memory capacity. Second, the correlation between the recall of the recency part and the reading span score was higher in the noise condition than in the control condition (in spite of the fact that the standard deviation was somewhat larger in the control condition). This result is to be expected since working memory capacity should be more critical when the word identification task becomes more demanding. However, large sampling errors make all differences between correlations unreliable and no definite conclusions can be made without replications.

No noise was present during recall, which meant that the sound context was the same during presentation of the word list and recall testing in the control condition, but different in the noise condition. Given that recall is context-dependent, this difference may have contributed to the difference between the noise and control condition. However, when tested no such context effect has been observed (Hygge, Boman, & Enmarker, 2003).

As expected the word identification was rated as more difficult in the noise condition, but the individual differences in this effect were unrelated to the effect of the noise on recall. Probably these differences are not valid indicators of the actual differences in effort level.

A surprising result was the shorter response times with noise for misses in the sentence recognition task, which contradicts the hypothesis that decisions should take longer time in the Noise condition. There is no obvious explanation of this effect, but it might possibly be an effect of the noise leading to a higher priority given to speed of performance (Hockey, 1984).

The critical effect of the noise was assumed to be that it made the word identification more cumbersome. Another interpretation would be that it was the noise in itself that disturbed the short- or long-term storing of the words and that the effect would have been obtained also with a visual presentation of the words; the noise present between the words then might have been more critical than the noise during the word presentation. However, previous research indicates that a much higher level than 60 dB(A) is required for a continuous broadband noise to impair free recall of text items (Daee & Wilding, 1977) and even at these high levels the effect on the number of recalled items has been inconsistent (e.g. Breen-Lewis & Wilding, 1984; Smith, Jones, & Broadbent, 1981; Wilding, Mohindra, & Breen-Lewis, 1982). An effect of noise at lower levels has only been obtained with speech sounds or other sounds that give an irrelevant sound effect (Beaman & Jones, 1998; LeCompte, 1994).

An interpretation of the effect in terms of allocation of more of the available resources to word identification implies that all types of degraded speech signals should have this effect. Previous studies have found that this is true for hearing impairment (Pichora-Fuller et al., 1995; Rabbitt, 1991), but it remains to show that acoustic deterioration other than noise, like a long reverberation time, also has this effect.

The results of the present study are in line with and extend those obtained by Rabbitt (1966, 1968) and Surprenant (1999). This effect is important to keep in mind when discussing acoustical norms for classrooms and other premises where understanding and memory of spoken information is vital. It is therefore surprising that this possible effect is not mentioned in reviews of cognitive noise effects. However, to make conclusions about the practical relevance of this effect several questions should be answered. Does the effect appear also in less unfavourable S/N conditions and when speech intelligibility is impaired in other ways? How does noise affect the understanding and memory of longer spoken texts? What is the relation between the effect and other indices of acoustical conditions (primarily speech transmission index and reverberation time)?

REFERENCES

Beaman, C. P., & Jones, D. M. (1998). Irrelevant sound disrupts order information in free recall as in serial recall. *Quarterly Journal of Experimental Psychology*, *51A*, 615–636.

Borg, G. (1998). *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics.

Breen-Lewis, K., & Wilding, J. (1984). Noise, time of day and test expectations in recall and recognition. *British Journal of Psychology*, *75*, 51–63.

Daee, S., & Wilding, J. M. (1977). Effects of high intensity white noise on short-term memory for position in a list and sequence. *British Journal of Psychology*, *68*, 335–349.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scandinavian Audiology*, *11*, 79–87.

Hällgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2001a). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *Journal of the American Academy of Audiology*, *12*, 357–370.

Hällgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2001b). Cognitive effects in dichotic speech testing in elderly persons. *Ear & Hearing*, *22*, 120–129.

Hockey, G. R. J. (1984). Varieties of attentional state: V. Effects of environment. In R. Parasuraman, & D. R. Davies (Eds.), *Varieties of attention* (pp. 449–483). New York: Academic.

Hygge, S., Boman, E., & Enmarker, I. (2003). The effects of road traffic noise and meaningful irrelevant speech on different memory systems. *Scandinavian Journal of Psychology*, *44*, 13–21.

Jones, D. (1990). Recent advances in the study of human performance in noise. *Environment International*, *16*, 447–458.

Kjellberg, A. (2004). Effects of reverberation time on the cognitive load in speech communication: Theoretical considerations. *Noise and Health*, *7*, 11–22.

Kryter, K. D. (1994). *The handbook of hearing and the effects of noise. Physiology, psychology and public health*. New York: Academic Press.

LeCompte, D. C. (1994). Extending the irrelevant speech effect beyond serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1396–1408.

Magnusson, L. (1995). Reliable clinical determination of speech recognition scores using Swedish PB words in speech-weighted noise. *Scandinavian Audiology*, *24*, 217–223.

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory. *Psychology and Aging*, *15*, 323–334.

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

Rabbitt, P. (1966). Recognition memory for words correctly heard in noise. *Psychonomic Science*, *6*, 383–384.

Rabbitt, P. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology*, *20*, 241–248.

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngologica*, *479*, 167–176.

Smith, A. (1989). A review of the effects of noise on human performance. *Scandinavian Journal of Psychology*, *30*, 185–206.

Smith, A. P., Jones, D. M., & Broadbent, D. E. (1981). The effects of noise on recall of categorized lists. *British Journal of Psychology*, *72*, 299–316.

Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables. *International Journal of Psychology*, *34*, 328–333.

Wilding, J., Mohindra, N., & Breen-Lewis, K. (1982). Noise effects in free recall with different orienting tasks. *British Journal of Psychology*, *73*, 479–486.

DOI: 10.1002/acp

Paper II

Long Reverberation Time Decreases Recall of Spoken Information

by

Robert Ljung and Anders Kjellberg

Reprinted from

JOURNAL OF
BUILDING ACOUSTICS
Volume 16 · Number 4 · 2009

MULTI-SCIENCE PUBLISHING CO. LTD.
5 Wates Way, Brentwood, Essex CM15 9TB, United Kingdom

BUILDING ACOUSTICS · Volume 16 · Number 4 · 2009 Pages 301 – 312

Long Reverberation Time Decreases Recall of Spoken Information

Robert Ljung and Anders Kjellberg
Laboratory of Applied Psychology, Centre for Built Environment,
University of Gävle, Gävle, Sweden
Robert.ljung@hig.se

(Received 5 June 2009 and accepted 21 September 2009)

ABSTRACT

The aim of the study was to explore if a long reverberation time has the same effect on recall of spoken words as background noise was shown to have in a previous study. A further aim was to study the influence of working memory capacity on performance in these conditions. Thirty-two subjects performed a word recall and a sentence recognition test. They repeated each word to ensure correct identification. A reading span test measured their working memory capacity. Performance of the word recall task was impaired by the long reverberation time. The effect was most evident in the primacy part of the word list. No correlation was found between reading span score and recall performance.

Key words: Working Memory, Reverberation Time, Speech Intelligibility, Recall

1. INTRODUCTION

A major determinant of speech intelligibility is the signal-to-noise ratio. When the difference between signal and noise decreases, the listener loses information and has to rely more on information redundancy and contextual cues to understand the message. But the effect of the signal-to noise ratio is not restricted to hearing what words are spoken. Kjellberg, Ljung and Hallman [1] showed that a background noise can impair free recall of a list of spoken words although they had been correctly heard. Subjects listened to lists of 50 words with and without a background noise, and they repeated each word aloud to ensure correct identification. Free recall followed directly after the listening session. Fewer words were recalled in the noise condition and a further analysis revealed that the noise effect was found both in the first and the last part of the list which was interpreted as an indication of noise effects both on the short-term and the longer-term storage of the words. Kjellberg *et al.*'s interpretation of the noise effect was that the noise made word identification more difficult, which left fewer working memory resources for the further processing of the words. Their conclusion therefore was that effective learning requires that a message can be heard without excessive effort. Rabitt [2, 3] reported similar results and interpreted them in a similar way. Surprenant and Neath [4] and Surprenant [5] also referred to the extra effort required by listening during noise as one possible explanation of the effects of noise on the serial recall of nonsense syllables. As an alternative explanation, they also proposed that the noise degrades the sensory trace of the stimulus, thus making it more difficult to discriminate the stimuli from each other. However, the latter explanation appears more relevant for serial recall of the order of a fixed set of very similar nonsense syllables than for the free recall of a word list.

In all of these studies the noise was continuous and an alternative interpretation therefore would be that the background noise disturbed rehearsal and encoding processing between the presented items. If that is the case, other ways of making speech perception more effortful should not have the same memory effect. The other major acoustic parameter, beside the signal-to-noise-ratio, affecting speech intelligibility is the reverberation time (RT), which is a measure of the decay time of a sound and depends on how much of the sound that is reflected and how much is absorbed by surrounding surfaces. The sound that arrives at the listeners' ear is a mixture of direct sound from the source and reflected sound, which arrives later and is superimposed on the direct sound and may mask it. A shorter RT therefore gives a clearer signal and better speech intelligibility given a constant signal-to-noise ratio [6]. Like noise, a too long RT thus means that phonological coding becomes more resource demanding, which should leave less resources for the further processing of the speech [7]. Since a long RT distorts the signals and gives a short phonetic tail to the signal, the pauses are slightly shorter, but long RT does not affect the quiet pauses in a qualitative way. An effect of RT on recall of words is not open to the alternative interpretation of the effect of the signal-to-noise ratio, since the difference in interstimulus interval is small (max 0.3 s) relative to the total time (4 s) between items. In the present experiment the effect of a long RT on the recall of spoken words was studied by presenting the words in a virtual room. To our knowledge no study has been reported on the effect of RT on the memory of spoken items.

Researchers in the field of acoustic environments with a focus on RT have mainly been interested in music perception and speech intelligibility and have very seldom studied memory effects. However, Beaman and Holt [8] presented distracting irrelevant sounds with different RTs (5 s and no reverberation) during a memory task with visual stimuli. Their hypothesis was that the long RT would soften the distracting sound and therefore decrease its variability. The softening of the distracting sound was predicted to reduce the "irrelevant sound effect" on a serial recall task. Jones and Macken [9] showed evidence that it is the variability in the distracting sound that affects serial recall and that it does not matter if the irrelevant sound is speech or tones. Beaman and Holt [8] presented results that confirmed this hypothesis, but Perham, Banbury and Jones [10] performed a similar study with more realistic RT values (0.7 s and 0.9 s) and found no effect.

From a practical point of view beneficial effects of an extremely long RT on distracting irrelevant sounds are of less importance than the possible negative effects on the understanding and memory of relevant spoken information. This is a realistic risk since we know that many of today's classrooms are of very poor acoustic quality [11].

Many classrooms do not even meet the basic requirement that it should be possible for everyone in the room to hear what is said, and especially so for children, old persons and people with hearing impairment [12]. The situation is even worse if it turns out that understanding and memory of what is said may be impaired also with RTs that only make listening more effortful.

If the critical effect of bad listening conditions is that word identification requires a larger part of the available working memory resources, persons with a low working memory capacity should be especially vulnerable to bad listening conditions. Data supporting such a notion have been presented in a few papers. Pichora-Fuller *et al.* [13] showed that older subjects recalled fewer items than younger subjects, despite equal levels of identification, on a speech-perception-in-noise task. Rabbit [14] found similar age differences and effects of hearing impairment using a comparable word-list-recall task. For a broader overview of age effects on speech perception and psychoacoustic findings, see Pichora-Fuller [15]. A recent study by Surprenant [16] showed that even small changes in sensory processing between younger and older persons can lead to measurable declines in cognitive functioning as measured by a serial recall task.

Kjellberg *et al.* [1] also found a correlation between the noise effect (the difference in recall accuracy between no-noise and noise conditions) and reading span performance, which indicated that the effect was larger, the smaller the working-memory capacity. This correlation was, however, only found for the effect on recall of the first part of the word list. However, it should also be noted that Lustig, May and Hasher [17] proposed that the critical ability measured with reading span is the ability to withstand proactive interference. The correlation found by Kjellberg *et al.* [1] thus may be taken as an indication that the noise weakened the inhibition of competing extra list items (items that were strongly associated with the list items). If so, bad listening conditions should increase the number of intrusions, for example the number of reported words that were not included in the list to be remembered, and this effect should be negatively correlated with reading span performance [18]. Kjellberg *et al.* [1] did not register such intrusions and, therefore, could not test this hypothesis. However, Rabbitt [2] who studied the recognition of previously presented words gives some support for this hypothesis. He found no effect of noise on the number of correctly recognized words, but an increased frequency of false alarms (that is, words that had not been presented before and were reported as belonging to the original list.) The implications of this pattern of results were neither discussed in Rabbitt nor, to our knowledge, in later references to his study.

Kjellberg *et al.* [1] also tested the effect of noise on the recognition of sentences which were constructed to be easily mixed up. No effect was found on error rate in this task. However, it is possible that the noise prolonged the time needed to decide whether a sentence had been presented before or not, without affecting the error rate. The sensitivity of this task therefore might be improved by using response times as a performance measure.

The present study is a near replication of Kjellberg *et al.* [1] but using a long RT, instead of background noise, to degrade listening conditions. Words and sentences were presented with a long or a short RT. Working memory capacity was tested in a separate reading span test. The objectives of the present study were to test if recall of correctly heard words is better when they are presented with a short than with a long RT, and if this possible effect applies both to the first and last part of list of the items to be remembered. We expected that intrusions of words not belonging to the original list would be more frequent in the long RT condition. Furthermore, recognition of sentences was expected to be less sensitive to long RT than the recall of words, but long RT was expected to prolong the response times. However, a larger working memory capacity as measured with reading span, the less is the predicted effect of long RT on both recall and recognition.

2. METHOD

2.1 Participants and design

The study included 32 participants, 27 women and 5 men with an age range of 18-35 years. All participants were native speakers of Swedish and all reported their hearing to be normal. A within-subject design was used with two conditions, long RT and short RT. The order of conditions was counterbalanced across participants.

2.2 Apparatus

The experiment was conducted in an anechoic chamber with the participant seated in a chair in the middle of the room. The speech was presented by 12 loudspeakers placed in a circle around the subject. The stimulus material was mixed with a surround system to obtain a diffuse sound field in the anechoic chamber. A laptop was used for visual presentation of the reading span and recognition tests.

2.3 Speech and acoustic conditions

The speech stimuli were a part of a package of standardized tests for speech audiometry [19]. The speech had an equivalent sound level of 64 dB(A), and was mixed with broadband noise to get a signal-to-noise-ratio of 15 dB(A). Two virtual classrooms were designed in CATT-Acoustics 8.0 software. All geometrical values were common for the two classrooms; both were the same size (length 10 m, width 6 m, height 3 m) and were furnished with 30 desks. The sound source was placed 1 m in front of the blackboard in the centre of the classroom at a height of 1.7 m, and the receiver was placed 6.6 m in front of the source at a height of 1 m. The classroom with short RT had various absorbing panels on the walls and the ceiling, and 30 pupils were seated in the desks. In the classroom with long RT only 15 pupils were seated and some absorbing panels were replaced by concrete walls. In the short RT condition, mean RT 0.25-4 kHz was 0.53 s (with max RT 0.58 s at 0.25 kHz) and in the long RT condition it was 1.17 s (with max RT 1.41 s at 0.125 kHz). The STI (Speech Transmission Index) values indicate that the short and long RTs should produce very good (73.5) and fair (56.1) intelligibility, respectively.

2.4 Performance tests

Reading Span test. Working memory capacity was assessed with the reading span test, which was taken from the cognitive test battery TIPS [20]. The subject's task was to comprehend sentences and to recall either the first or the final words of the presented sentences. The sentences were presented visually, one word at a time. Each word was shown on the screen for 0.8 sec. The inter-word interval was 0.075 sec. Half of the sentences were absurd (e.g., "the house read a newspaper"), and half normal (e.g., "the pupil was late"). The subjects' task was to indicate, during a 1.75 sec interval, whether the sentence was normal or absurd by pressing a key on the keyboard. After a sequence of sentences (three, four, five or six sentences), the experimenter indicated that the subject should start to report orally as many as possible of either the first or the final words of the presented sentences. The number of correctly recalled words was registered.

Recognition of sentences. The Hagerman test [19] is a list of spoken Swedish sentences with the same grammatical structure and is a part of a package of standardized tests for speech audiometry. The subjects' task was to memorize orally presented sentences for later recognition. There was approximately 7 s of silence between each sentence, during which the subjects repeated each sentence aloud to check that they had identified each word. Two lists with ten sentences each were used. Each sentence contained five words and their structure was identical (name, verb, number, adjective, noun) but within this structure the words were not predictable (e.g. Kim bought six white pillows). Both lists contained exactly the same words but combined in new ways. One list was presented in the long RT condition and one in the short RT condition in a counterbalanced way. Directly after presentation the subjects were shown a series of 20 sentences, 10 of which had been presented previously. The task was to determine whether the sentence was one of the old ones or a new one. The number of correct answers and mean response time were measured.

Free recall of a word list. Two phonetically balanced word lists, each of 50 one-syllable words were presented orally to the subjects. The lists are a part of a package of standardized tests for speech audiometry [19]. The lists of words and sentences have previously been shown to be equally intelligible [19, 21]. The task was to memorize the words for later recall. There were approximately 4 s of silence between each presented word, during which the subject was asked to repeat the word aloud in order to check that they had identified it correctly. One list was presented with a long and one with a short RT, in a counterbalanced way. Directly after presentation, the subjects were asked to write down on paper all the words they could recall. Recall performance was measured in two ways. The first was the number of words correctly recalled of the words that they had stated when the list was presented, irrespective of whether the word had been correctly identified or not (stated words). The second recall measure was the percentage of words recalled of those that had been correctly identified (correct words). In addition, the proportion of the words written down that had not been part of the list (extra list errors), were registered.

2.5 Rated effort

To validate the assumption that word identification was made more effortful by the long RT, subjects rated the effort required to follow the speech using Borg's CR10 scale [22]. This was done directly after the presentations of the word and sentence lists. The scale has a range of 0-10 with verbal labels on eight steps. The scale values of the verbal labels have been chosen with the aim of approximating ratings at a ratio scale level.

2.6 Procedure

All subjects performed the Reading Span Test in silence followed by the tests with free recall of the auditorily presented words and the recognition of the auditorily presented sentences. The order between the two conditions (Long RT and Short RT) was counterbalanced across participants. Finally, subjects were interviewed about how the conditions were experienced and what strategy, if any, they had used in the memory tasks. Altogether, the experimental sessions lasted for about 40 minutes and were conducted between 9 AM and 4 PM. At the outset subjects were informed that the study was about memory.

3. RESULTS

As a check of the effect of RT on the difficulty of the task, the means of self-reported effort and of numbers of incorrectly repeated words and sentences were calculated (Table 1). Two two-way ANOVAs (RT Conditions x Order of Conditions) showed that both effort and the number of incorrectly repeated words and sentences were significantly higher in the Long RT condition There was no effect of presentation order for either incorrectly repeated words, F (1, 30) = .93, p =.34, or sentences, F (1, 30) = 1.52, p = .23. For incorrectly repeated sentences, there was a significant condition x order interaction, F (1, 30) = 4.38, p= .045, indicating that there was less difference in incorrectly repeated sentences between conditions when the long RT condition came last.

Table 1. Mean values (standard deviation) of effort and number of incorrectly repeated words in Long RT and Short RT conditions and results from two-way analyses of variance of the effect of conditions and order of presentation.

	Long RT M (s)	Short RT M (s)	*F*	*p*
Effort -word	4.64 (2.23)	2.53 (1.40)	50.25	<0.001
Effort -sentences	3.94 (2.69)	2.30 (2.26)	24.90	<0.001
Incorrectly repeated words	9.44 (3.05)	3.00 (2.02)	171.95	<0.001
Incorrectly repeated sentences	1.03(1.12)	0.25 (0.51)	16.19	<0.001

Recall of words. The numbers of correct and stated words recalled were almost perfectly correlated across participants (.981 and .996 in the Long RT and Short RT condition, respectively). Therefore, only the analyses of correctly recalled stated words are reported. Two two-way ANOVAs (condition x presentation order) were performed for recall performance. Significantly fewer stated words were recalled in the Long RT than in the Short RT condition (means = 12.97 and 10.78 respectively, F (1, 30) = 7.67, p=.01, η^2 = .20. The overall effect of presentation order was not significant, but recall in the Short RT condition was significantly better when it was performed as the second condition. No such difference was seen in the long RT condition. This was shown as an interaction between condition and presentation order, F (1, 30) = 4.32, p= .046, η^2=.25).

The interaction between order and RT was primarily the result of four subjects that performed very much better in the second condition. Three of them had long RT as their first condition, and the order effect therefore strengthened the hypothesized difference between conditions. The interviews after the experiment revealed that these subjects had changed to a more effective mnemonic strategy in the second condition. A two-way analysis of variance without these four subjects showed that the main effect of RT conditions remained significant $F(1, 25) = 6.31$, $p = .019$, $\eta^2 = .20$ despite the reduced mean difference between conditions (2.19 in the whole group and 1.58 in the reduced group). Furthermore, in the reduced group there was no significant interaction between order and condition.

In order to explore the serial position effect, the word lists were split up into five parts with ten words in each part. As shown in Figure 1, recall of the first two parts of the list only was affected by RT. This was reflected in an interaction between condition and parts in the linear trend $F(1, 30) = 10.16$; $p = .003$, $\eta^2 = .25$. A test of the difference between RT conditions in the first two parts of the list showed that this effect remained significant after exclusion of the four subjects with an extreme order effect $F(1, 26) = 16.31$, $p = .001$, $\eta^2 = .39$, although the mean difference between conditions was smaller than in the whole group (0.99 in the whole group and 0.74 in the reduced group).

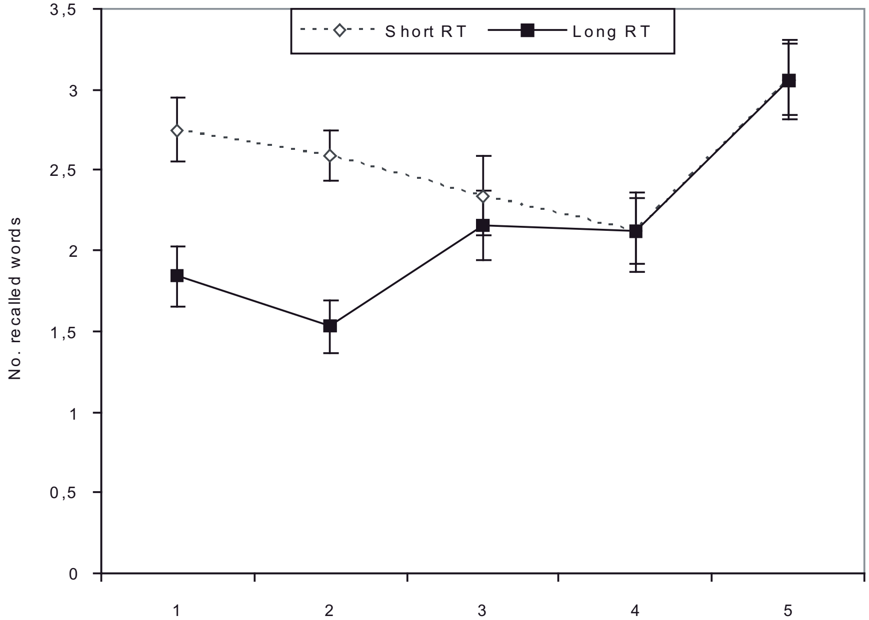


Figure 1. Number of correctly recalled words and standard errors in the five parts of the recall test in the Long RT and Short RT conditions. Standard errors calculated with consideration taken to the within-subjects design.

The distributions of extra list errors were very positively skewed, especially in the long RT condition. Thus the difference between conditions was tested using Wilcoxon's matched-pairs signed-ranks test. Extra list errors were significantly less frequent in the short than in the long RT condition (means: 2.2 and 3.1, $z= 2.22$, $p = .03$).

Recognition of sentences. A two-way ANOVA (condition x presentation order) showed no significant difference between the conditions regarding the number of correct responses of Hagerman's sentences. There was no effect of presentation order, or interaction between order and condition, for any of the accuracy measures. However, an analysis of response times on the recognition test revealed that when sentences were heard in the short RT condition the participants were faster to identify the sentences that not had been presented (Short RT M = 3.03 s, Long RT M = 3.42 s, $F\ (1, 31) = 4.62$, $p = .04$, $\eta^2 = .13$).

Relation between working memory capacity and recall. Correlations were calculated for the relation between reading span score and recall score overall and for the five parts of the word lists as well as for extra list errors. None of these correlations was significant. The highest, although nonsignificant, correlation obtained was with the proportion of extra list errors in the short RT condition ($r = -.31$, $p = .087$). Neither was there any correlation between reading span score and the effect of RT. In no case was there any clear deviation from linear relationships.

Corresponding analyses of the sentence recognition test did not show any significant relation with reading span performance.

In addition reading span scores were dichotomized and trichotomized to test differences in recall performance between these groups. In no case were the mean differences significant.

4. DISCUSSION

In line with the hypothesis the participants recalled fewer words when the word list had been presented with a long RT and extra list errors were also more common in the long RT condition. The RT effect was pronounced at the beginning of the word list but was absent in the last part of the list. Recognition of sentences was, as expected, less sensitive to the long RT, but measurements of response time revealed faster responses when sorting out irrelevant sentences in the short RT condition. Contrary to expectations, reading span performance was unrelated to performance in both the word list and sentence task as well as to the effect of RT in these tests.

Kjellberg *et al.* [1] found that background noise impaired recall of words from the more adverse listening condition, despite correct identification during the listening session. They assumed that the critical effect of the noise was to make word identification more inefficient. Similar reasoning is found in a major literature, for example [2, 14, 13, 23]. An alternative interpretation was that the noise between the words disturbed the short- or long-term storage of the word. In the present study such an interpretation was excluded since only the speech signal was affected by the RT, making the two conditions nearly identical in the pauses between the words.

In the previous noise experiment there was an effect both on the first and last part of the word list, whereas only the earlier parts of the list was affected in the present experiment. This indicates that the effect of the noise on the last part in the former study may have been the result of distraction effects of the noise and not by the degraded speech signal. The effect on the first part of the list that was obtained in both studies thus may be explained by the more resource demanding item identification becomes when the RT is too long or the background noise is too high, which should leave less time for the transfer to long-term storage and early consolidation in the long-term memory.

Kjellberg *et al.* [1] found significant correlations between reading span and performance of the recall task as well as with the effect of noise on recall of the mid and last part of the list, which supported this interpretation. In the present study none of these correlations was significant. A strong order effect might conceal such correlations, but this seems unlikely, since the exclusion of the four subjects with the strongest order effect did not change the results. Considering the wide confidence intervals of correlations obtained in a group of 32 persons, further studies are needed for a more definite conclusion.

As expected, extra-list errors were more common in the long RT condition. This replicates the findings of Rabbitt [2] in a recognition task and indicates that the inhibitory control was less efficient in bad listening conditions. It is also to be noted that extra-list errors also was the performance measure that was most strongly correlated with reading span which is in line with the proposal by Lustig *et al.* [17] that reading span performance primarily reflects inhibitory control. This indicates that it might be more fruitful to relate the effect of unfavourable listening conditions to measures of specific working memory processes rather than to try to find indicators of global working memory capacity. Miyake [24] showed examples of different tasks and measurements regarding higher cognitive functions, and an inclusion of a measure that reflects inhibitory and other aspects of executive control, for example, might help us to specify the nature of the effect of noise, long RTs and other factors that make speech perception more difficult. Wingfield *et al.* [25] studied interactions between age, hearing loss and speed and syntactic structure in speech. Their results suggest that neither age-related cognitive constraints, nor peripheral hearing acuity alone, will give the full picture for individuals' effectiveness in sentence comprehension.

As predicted, the effect of the long RT on recognition performance was restricted to the reaction time measures. This effect was shown as a longer processing time to sort out the sentences that had not been presented during the listening session. To decide if you have not heard a sentence requires that you search the entire to-be-remembered list of sentences before you are able to determine that it was not presented. That task therefore is more demanding than the recognition of the presented sentences and should be more vulnerable to bad listening conditions.

The present study demonstrated that a long RT may disrupt memory of spoken information, including for words that have been correctly identified. This is important to keep in mind when discussing acoustical norms for classrooms and other premises where understanding and memory of spoken information is vital. However, to get a basis for better acoustical norms it is necessary to test the effects on more realistic learning tasks, such as recall of longer spoken texts.

5. REFERENCES

1. Kjellberg, A., Ljung, R. and Hallman, D., Recall of words heard in noise, *Applied Cognitive Psychology*, 2008, 22, 1088-1098.
2. Rabbitt, P., Recognition memory for words correctly heard in noise, *Psychonomic Science*, 1966, 6, 383-384.
3. Rabbitt, P., Channel-capacity, intelligibility and immediate memory, *Quarterly Journal of Experimental Psychology*, 1968, 20, 241-248.
4. Surprenant, A. M. and Neath, I., The relation between discriminability and memory for vowels, consonants, and silent-center vowels, *Memory & Cognition*, 1996, 24, 356-366.
5. Surprenant, A. M., The effect of noise on memory for spoken syllables, *International Journal of Psychology,* 1999, 34, 328-333.
6. Hodgson, M. and Nosal, E., Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms, *Journal of Acoustical Society of America*, 2002, 111, 931-939.
7. Kjellberg, A., Effects of reverberation time on the cognitive load in speech communication: theoretical considerations*, Noise and Health*, 2004, 7, 11-22.
8. Beaman, C. P. and Holt, N. J., Reverberant auditory environments: the effects of multiple echoes on distraction by irrelevant speech, *Applied Cognitive Psychology*, 2007, 21, 1077-1090.
9. Jones, D. M. and Macken, W. J., Irrelevant tones produce an ”irrelevant speech effect”: Implications for phonological coding in working memory, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1993, 19, 1-13.
10. Perham, N., Banbury, S. and Jones, D. M., Do realistic reverberation levels reduce auditory distraction?, *Applied Cognitive Psychology*, 2007, 21, 839-847.
11. Seidel, J., Weber, L., Leistner, P. and Laschczok, S., Acoustic properties in German class rooms and their effect on the cognitive performance of primary school pupils, *Proceedings of Forum, Acusticum*, Budapest, 2005, 2065-2069.
12. Helfer, S. K. and Wilber, A. L., Hearing loss, ageing, and speech perception in reverberation and noise, *Journal of Speech and Hearing Research*, 1990, 33, 149-155.
13. Pichora-Fuller, M. K., Schneider, B. A. and Daneman, M., How young and old adults listen to and remember speech in noise, *Journal of the Acoustical Society of America*, 1995, 97, 593-608.
14. Rabbitt, P. M. A., Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ, *Acta Otolaryngologica, Supplementum*, 1991, 476, 167-176.
15. Pichora-Fuller, M. K., Cognitive aging and auditory information processing, *International Journal of Audiology*, 2003, 4 (2), 26-32.
16. Surprenant, A. M., Effects of Noise on Identification and Serial Recall of Nonsense Syllables in Older and Younger Adults, *Aging, Neuropsychology, and Cognition*, 2007, 14, 126-143

17. Lustig, C., May, C. P. and Hasher, L., Working memory span and the role of proactive interference, *Journal of Experimental Psychology: General*, 2001, 130, 199-207.
18. Sörqvist, P., Halin, N. and Hygge, S., Individual differences in susceptibility to the effects of speech on reading comprehension, *Applied Cognitive Psychology*, 2009, Published online in Wiley InterScience, DOI: 10.1002/acp.1543.
19. Hagerman, B., Sentences for testing speech intelligibility in noise, *Scandinavian Audiology*, 1982, 11, 79-87.
20. Hällgren, M., Larsby, B., Lyxell, B. and Arlinger, S., Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons, *Journal of the American Academy of Audiology*, 2001, 12, 357-370.
21. Magnusson, L., Reliable clinical determination of speech recognition scores using Swedish PB words in speech-weighted noise, *Scandinavian Audiology*, 1995, 24, 217-223.
22. Borg, G., *Borg's perceived exertion and pain scales*, Champaign, IL: Human Kinetics, 1998.
23. Surprenant, A. M., LeCompte, D. C. and Neath, I., Manipulations of irrelevant information: Suffix effects with articulatory suppression and irrelevant speech, *Quarterly Journal of Experimental Psychology*, 2000, 53, 325-348.
24. Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. and Wager, T., The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis, *Cognitive Psychology*, 2000, 41, 49-100.
25. Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A. and Cox, L. C., Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity, *Journal of the American Academy of Audiology*, 2006, 17, 487-497.

Paper III

Poor Listening Conditions Impair Memory for Intelligible Lectures: Implications for Acoustic Classroom Standards

by

Robert Ljung, Patrik Sörqvist, Anders Kjellberg
and Anne-Marie Green

Reprinted from

JOURNAL OF
BUILDING ACOUSTICS
Volume 16 · Number 3 · 2009

MULTI-SCIENCE PUBLISHING CO. LTD.
5 Wates Way, Brentwood, Essex CM15 9TB, United Kingdom

Poor Listening Conditions Impair Memory for Intelligible Lectures: Implications for Acoustic Classroom Standards

Robert Ljung, Patrik Sörqvist, Anders Kjellberg and Anne-Marie Green
Laboratory of Applied Psychology, Centre for Built Environment
University of Gävle, Gävle, Sweden
Robert.ljung@hig.se

(Received 26 April 2009 and accepted 28 August 2009)

ABSTRACT

This paper reports two experiments on the effects of degraded speech signals on memory for spoken lectures. Experiment 1 showed that broadband noise impairs university students' memory for a spoken lecture, even though the participants heard what was said. Experiment 2 showed that reverberation has detrimental effects to school adolescents' memory for spoken lectures, similar to broadband noise. The results suggest that poor listening conditions (resulting from background noise and/or long reverberation time) impair memory and learning, even if the conditions allow the listeners to hear what is said. Since the goal for students and pupils attending to lectures is to remember the lecture rather than just hearing what is said, the results presented here indicate that standards for acceptable signal-to-noise ratios and reverberation times in buildings designed for learning should consider the distinction between speech intelligibility and memory. Standards should be based on memory criteria instead of intelligibility criteria.

Keywords: broadband noise, reverberation time, signal-to-noise ratio, classrooms, acoustic standards, memory

1. INTRODUCTION

Acoustic standards for classrooms are based on intelligibility criteria [1-3]. The American National Standards Institute states that ”Normal adults typically require 0 dB signal-to noise ratios for high speech intelligibility when listening to simple and familiar speech material for short periods of time”[4]. An underlying assumption is that if speech is identified correctly, there are no detrimental effects of noise and reverberation. Still, to hear what the teacher says is only a necessary condition for understanding and remembering the lecture, but perhaps not a sufficient condition. Kjellberg [5] argued that poor listening conditions may impair memory for spoken materials even when it is possible to hear what is said. This hypothesis was tested by

Kjellberg, Ljung and Hallman [6] by presenting to-be-recalled spoken word lists with and without a background noise (signal-to-noise ratio 4 and 27 dB respectively). Participants were asked to repeat aloud each word immediately after its presentation. This repetition procedure made it possible to ensure complete intelligibility. After list presentation, the participants were asked to recall all words presented in the list. The results revealed that a broadband noise during presentation impairs memory, even when the words are correctly identified. Ljung and Kjellberg [7] later conceptually replicated these results, using the same procedure, but comparing memory for word lists spoken with a long reverberation time with memory for word lists spoken with a short reverberation time. They found that memory is impaired by a long reverberation time, even though the words are intelligible, just as with a small signal-to-noise (S/N) ratio. These results are consistent with several other studies showing that noise impairs memory for intelligible word lists [8-11]. We therefore argue that the acoustic standards for classrooms should be based on a memory criterion instead of intelligibility criteria.

A good memory of a spoken message seems to require a better S/N ratio than 0 dB, even though high speech intelligibility may be achieved at this ratio. Relatively good S/N ratios may therefore be detrimental to memory and learning in schools. However, studies looking into the effects of noise on memory for intelligible speech have almost exclusively been restricted to memory for word lists. Recall of spoken word lists is a task rarely encountered outside the laboratory, and it is therefore too rash to base acoustic standards for classrooms on these studies. Students and school children are challenged with the task of understanding and remembering spoken lectures rather than word lists, and background noise and reverberation may well interfere with memory for spoken lectures in ways that differ from memory for a word list.

Memory for spoken lectures can be said to consist of traces from many different levels of processing: perceptual processing, analysis of the words, phrases, sentences and the meaning of the message. Traces from any of these activities may be retained in memory [12] and poor listening conditions may interfere with any of these levels of processing. For instance, noise might impair perception and memory of single words, but still not affect the listener’s memory of larger parts of the lecture. Another factor to be considered when moving from memory of word lists to memory of lectures is the context which may promote intelligibility for spoken lectures, whereas this is not possible for lists of unrelated words presented with no context. This may make understanding and memory of the lecture less sensitive to the effects of noise.

In a pioneering study, Rabbitt [13] showed that poor listening conditions (+5 dB S/N) impair memory of spoken prose. He let the participants listen to prose passages and answer questions about the contents afterwards. One group of participants listened to the prose passage without noise in the background (No noise group). Another group listened to the passage without noise during the first half and with noise during the second half (Noise group). Rabbitt found that the Noise group scored lower on questions concerning the first half of the prose passage (which was heard without noise in both groups) as well as the second half. Rabbitt’s interpretation of the results was that the degraded speech in the second half in the Noise group impaired any additional processing of the first half of the lecture. However, the between-groups design used by

Rabbitt is open to other interpretations. One possible interpretation of his results is that the participants in the Noise group had a lower memory capacity in general, or poorer speech perception abilities, and therefore received a low score. This explanation is consistent with the finding that the Noise group received a lower score for the part of the prose passage that was heard without noise. Another problem with Rabbitt's design is that he tested the participants in groups of 11 to 22. Obviously, listening conditions may have varied with group size and this effect differed between the two groups (e.g., if the speech was presented in loudspeakers and not with earphones; the description of the method is not clear on this point). Hence, there are reasons to doubt that the difference between the two groups was caused by noise.

2. EXPERIMENT 1

In Experiment 1, we aimed to replicate Rabbitt's [13] results by testing the effects of broadband noise on memory for a spoken lecture, while controlling for the participants' ability to hear what is said during the lecture and using a within-subjects design.

2.1 Method

2.1.1 Participants

28 university students 19-35 years old were paid to participate in the experiment. All participants were native speakers of Swedish and reported normal hearing ability.

2.1.2 Materials

Spoken lectures and noise. The spoken lectures (eight minutes long) were studio recorded and taken from two reading comprehension tests previously used in the national university aptitude test. The participants listened to one lecture with recorded broadband background noise and another lecture without the background noise. One lecture concerned inductivism and scientific methods, and the other lecture was about acting. After listening to the lecture, the participants were given eight open-ended questions about the content of the lecture. In the noise condition, a broadband noise was presented simultaneously with the spoken lecture giving a S/N ratio of + 5dB(A), which was expected to make it difficult but possible to hear everything that was said during the lecture. In the control condition, the S/N ratio was +29 dB(A). The lectures and the noise were presented in two loudspeakers placed 1.5 m in front of the participant.

Hearing test. The hearing tests consisted of two lists of ten sentences each presented with and without the broadband noise. All sentences had the same structure (e.g. *Sean took eighteen old balls, Anna held three beautiful rings*), and were constructed so as to carry no redundant information (i.e., the context gave few cues to what word would follow). The participants immediately repeated each sentence aloud. The five first sentences in each list were considered as training, and only the results from the five last sentences were used to measure the hearing ability. The sentences were taken from a standardized hearing test [14].

2.1.3 Procedure and design

The experiment was conducted in a sound-attenuated climate chamber, with the participants seated at a desk in the middle of the room. All participants were tested individually. They began by performing the hearing test followed by the two tests of memory for spoken lectures, one presented with broadband noise and one presented without broadband noise. A within-subject design was used and the order between background noise and control condition was counterbalanced (i.e., half of the participants began with the background noise condition and half with the control condition), as well as the order of the two different spoken lectures.

2.2 Results

The participants' memory performance was worse when the lecture was heard in the noise condition ($M = 2.68$, $SD = 1.64$) than in the control condition ($M = 3.45$, $SD = 1.76$), as shown in Figure 1. A 2 (Noise condition: +5dB S/N ratio vs. +29dB S/N ratio) × 2 (Condition order: +5dB S/N ratio first vs. +29dB S/N ratio first) analysis of variance, with Noise condition as a within-subject variable and Condition order as a between-subject variable, revealed a significant difference between the two background conditions, $F(1, 26) = 6.71$, $MSE = 1.23$, $p < .05$, $\eta^2 = .20$, observed power = .70, but no significant main effect of condition order and no interaction between these two variables. An additional analysis including only participants with no errors in the hearing test ($n = 16$) showed consistent results. These results suggest that a background noise impairs memory for spoken lectures.

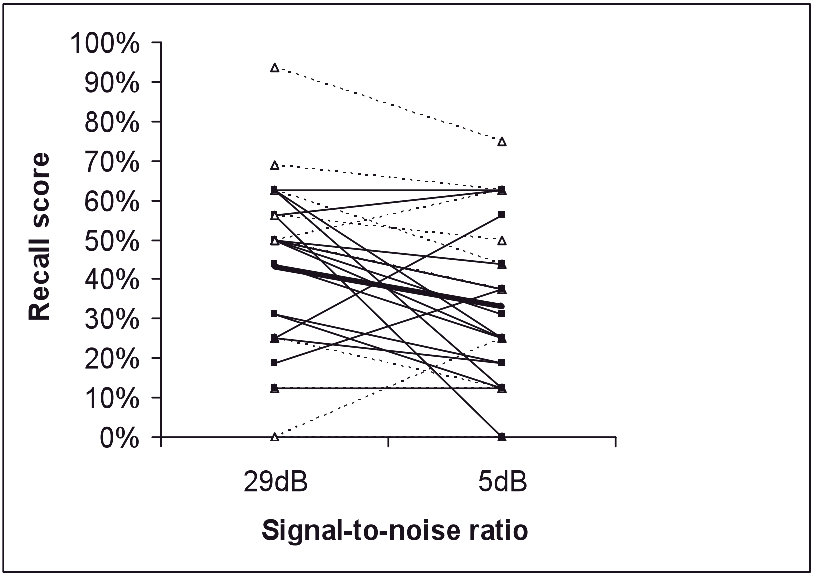


Figure 1. Each participant's recall score (in percent of total score possible) for lectures spoken with a +29 dB versus a +5 dB signal-to-noise ratio in Experiment 1. Dotted lines with triangle-marks represent participants who had at least one error on the listening test, whereas unbroken lines with square-marks represent participants who had no error on the listening test. The thick line with circle-marks represents the two condition means.

2.3 Discussion

The results from Experiment 1 are consistent with previous investigations [6, 9, 13] and indicate that background noise in classrooms may be detrimental to listeners' memory for lectures even when they are able to hear what is said, at least for university students.

Another factor that influences the transmission of speech from a speaker to a listener is reverberation. The reverberation time is the time it takes for a sound signal to drop 60 dB from its initial dB level [15]. Typically recommended values for classrooms are in the range of 0.3-0.8 s [3, 4, 16], but it is not unusual to find a reverberation time of as much as 1.3 s in classrooms [17, 18]. The effects of reverberation on the speech signal resemble those of a broadband noise. Based on the results obtained in Experiment 1, we assumed that a long reverberation time is detrimental to memory for a spoken lecture, even when the listener is able to hear what is said.

3. EXPERIMENT 2

The memory of spoken information presented with varying reverberation time has, to our knowledge, only been dealt with in one previous study [7] where the to-be-remembered material was word lists. The purpose of Experiment 2 was to conceptually replicate the effect found in Experiment 1, but testing the effect of reverberation time instead of broadband noise on memory for spoken lectures. Another aim of the experiment was to test the effect in more ecologically valid conditions. Two steps were taken towards reaching a higher ecological validity in Experiment 2. First, the spoken lectures were recorded in two ordinary classrooms; one with a long reverberation time and one with a short reverberation time. Second, school adolescents (rather than university students) were recruited as participants and tested in a group setting.

3.1 Method

3.1.1 Participants

A total of 20 adolescents from an upper secondary school class in Sweden served as participants in exchange for a cinema ticket. One reported a hearing impairment and was therefore removed before the analyses. The remaining 19 participants (17 females and 2 males) were about 17 years old and reported normal hearing and Swedish as their native tongue.

3.1.2 Materials

Contents of lectures. Two lectures were constructed. One lecture was about a fictitious culture called the "Timads" and the other was about a fictitious culture called the "Lobiks". The purpose of using fictitious cultures was to make sure that previous knowledge did not influence task performance. Each lecture consisted of 10 short paragraphs about different topics (e.g., religion, geography, history) and included two phases; one listening phase and one recall phase. In the first phase, the participants listened to the 10 paragraphs. In the second phase of the lecture task, 20 questions (2 for each paragraph) were presented in sequential order on a computer screen. The questions concerned facts explicitly stated in the lecture (e.g., "What did people wear in their afterlife?"). Answers were never longer than a single sentence (e.g., "In bird feathers") and scored as correct if they contained a specific keyword (e.g., "feathers") or described the accurate meaning of the keywords (e.g., "Bird suits"). The participants answered the questions by typing them on the computer keyboard and were given as much time as they needed to answer the questions. After each lecture, the participants scored their ability to hear the lecture on a 7-point scale. This made it possible to compare self-reported hearing of the two lectures.

Auditory recording of lectures. The two lectures were spoken in male voice at normal speed in an anechoic room and recorded. To get ecological school recordings, the lectures were played from a loudspeaker in two ordinary classrooms (one with long reverberation time and one with short) and recorded binaurally. These binaural recordings were presented to the participants through headphones. The classrooms were about the same size (length 10 m, width 6 m, height 3 m) and were furnished with desks. The loudspeaker was placed 1 m in front of the blackboard in the centre of the classroom at a height of 1.5 m, and an acoustical head was placed as a seated student at a desk in the back of the room, about 6 m in front of the loudspeaker. The classroom with short reverberation time had various absorbing panels on the walls and the ceiling. The reverberation time was 0.3 s in all octave bands from 125 Hz to 4 kHz. The classroom with long reverberation time had some absorbing panels, but the walls and the ceiling was mostly painted concrete, and the reverberation time was 1.84 s at 125 Hz, 1.46 s at 250 Hz, 0.94 s at 500 Hz, 0.77 s at 1 kHz, 0.78 s at 2 kHz and 0.68 s at 4 kHz.

3.1.3 Design and procedure

The computers, with headphones attached, were rigged in an ordinary upper secondary school classroom. The participants entered the classroom and sat down at their computers, which were selected without the influence of the experimenter. The participants were given a short oral statement of the purpose of the experiment (i.e., that the experiment was about memory for stories) and thereafter performed the two lecture tasks. Since the participants completed the lecture tasks at different paces, additional filler tasks followed after the last lecture. The purpose with these tasks was to engage the participants in a meaningful activity so as not to disturb the other participants still working on the lecture tasks. The participants did not interact during the experiment proper. After the experiment, the participants were debriefed and thanked. Experiment 2 had a within-subject design, and the order between the two reverberation-conditions was counterbalanced between participants (i.e., half began in the long reverberation time condition and half in the short reverberation time condition), as well as the order of the two different spoken lectures.

3.2 Results and Discussion

The participants' memory performance was worse when the lecture was heard with a long reverberation time ($M = 2.16$, $SD = 1.57$) than in the short reverberation time condition ($M = 4.00$, $SD = 2.03$), as shown in Figure 2. A 2 (Reverberation condition: long vs. short reverberation time) × 2 (Condition order: long reverberation time first vs. short reverberation time first) analysis of variance, with Reverberation condition as a within-subject variable and Condition order as a between-subject variable, revealed a significant difference between the two reverberation conditions, $F(1, 17) = 16.60$, $MSE = 2.12$, $p < .001$, $\eta^2 = .49$, observed power = .97, but no main effect of condition order and no interaction between these two variables. These results suggest that a long reverberation time impairs memory for spoken lectures. The participants' mean rating of how well they heard the lecture was 4.95 (*SD* = 1.75) in the short reverberation time condition and 4.68 (*SD* = 1.89) in the long reverberation time condition. These two means did not differ significantly, as shown by a paired *t*-test, $t(18) < 1$. An additional analysis was carried out based only on the participants who rated their hearing of the lecture spoken with long reverberation time equal to or better than their hearing of the lecture spoken with short reverberation time ($n = 8$). The results from this analysis were entirely consistent with the results from the analysis above. Taken together, the results from Experiment 2 indicate that a long reverberation time disrupts memory for lectures, even when the participants are able to hear the lectures equally well as lectures spoken with a short reverberation time.

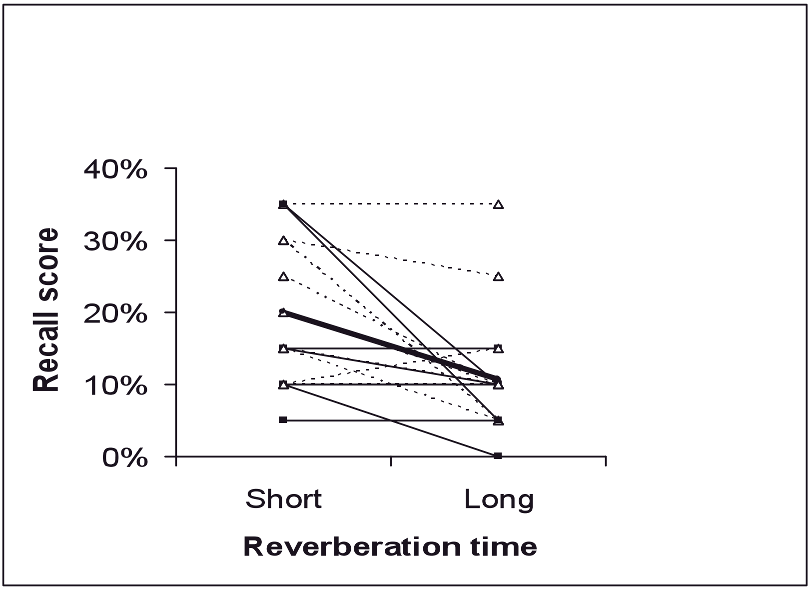


Figure 2. Each participant's recall score (in percent of total score possible) for lectures spoken with a short (0.3 sec) versus long (≈ 1.0 sec) reverberation time in Experiment 2. Dotted lines with triangle-marks represent participants who reported that they heard the lecture with short reverberation time better, whereas unbroken lines with square-marks represent participants who had reported that they heard the lecture with long reverberation time at least as good as the other lecture. The thick line with circle-marks represents the two condition means.

4. GENERAL DISCUSSION

Taken together, these two studies show that background noise and long reverberation time are detrimental to memory of spoken lectures, even when people are able to hear what is said. Experiment 1 found that a S/N ratio of +5 dB, which is higher than the 0 dB needed for high speech intelligibility [4], impaired memory for a spoken lecture compared with a S/N ratio of + 29 dB. Experiment 2 found similar results for long reverberation time compared with short. These results are consistent with previous investigations into the effects of noise on memory for spoken word lists [6, 8-11] and for spoken prose [13] and lead us to argue that acoustic standards for rooms meant for learning (e.g., upper secondary school classrooms) should be more stringent than previously suggested.

With few exceptions, studies on the effects of noise on learning have tested memory for written materials [19-22]. Little has been done concerning effects of noise on memory for speech, and to our knowledge, the present paper is the first to report effects of reverberation on memory for spoken lectures. We suggest two directions for future investigations. The effects of reverberation and background noise could be combined to test for interactions. Also, the effects of different levels of S/N ratios (and reverberation times) on memory should be compared to investigate the function between decreased S/N ratio (and reverberation time) and decreased memory performance.

In conclusion, hearing what is said is a necessary but not a sufficient condition for people to remember what is said. Today's standards for acceptable signal-to-noise ratios and reverberation times in buildings designed for learning do not consider this discrepancy between intelligibility and memory. The standards are solely based on a hearing criterion [1-4]. Since the goal is to remember the lecture rather than solely hear it, the results presented here and elsewhere [6, 9, 10, 13] suggest that acoustic standards should be based on memory criteria instead.

REFERENCES

1. Karabiber, Z., & Vallet, M., Classroom acoustics policies – An overview, *Acta Acustica (Stuttgart)*, 2003, 89, 12.
2. Shield, B. M., & Dockrell, J. E., The effects of noise on children at school: A review, *Building Acoustics*, 2003, 10, 97-116.
3. Smaldino, J. J., & Crandell, C. C., Speech perception in classroom, *Volta Review*, 2000, 101(5), 15-21.
4. American National Standards Institute., *Acoustical performance criteria, design requirements, and guidelines for schools*, 2002, (Standard S12.60-2002).
5. Kjellberg, A., Effects of reverberation time on the cognitive load in speech communication: Theoretical considerations, *Noise & Health*, 2004, 7, 11-21.
6. Kjellberg, A., Ljung, R., & Hallman, D., Recall of words heard in noise, *Applied Cognitive Psychology*, 2008, 22, 1088-1098.
7. Ljung, R. & Kjellberg, A., Long reverberation time decreases recall of spoken information. *Building Acoustics* (in press)

8. Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M., How young and old adults listen to and remember speech in noise, *Journal of the Acoustical Society of America*, 1995, 97, 593–608.
9. Rabbitt, P., Recognition memory for words correctly heard in noise, *Psychonomic Science*, 1966, 6, 383-384.
10. Rabbitt, P., Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngologica*, 1991, 479, 167-176.
11. Surprenant, A., Effects of noise on identification and serial recall of nonsense syllables in older and younger adults, *Aging, Neuropsychology, & Cognition*, 2007, 14, 126-143.
12. Kintsch, W, & Bates, E., Recognition memory for statements from a classroom lecture, *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 150-159.
13. Rabbitt, P., Channel-capacity, intelligibility and immediate memory, *Quarterly Journal of Experimental Psychology*, 1968, 20, 241-248.
14. Hagerman, B., Sentences for testing speech intelligibility in noise, *Scandinavian Audiology*, 1982, 11, 79-87.
15. Rossing, T. D., *The science of sound*, Reading Mass, 1990, Addison-Wesley.
16. Finitzo-Hieber, T. & Tillman, T., Room acoustics effects on monosyllabic word discrimination ability for normal-hearing and hearing-impaired children, *Journal of Speech & Hearing Research*, 1978, 21, 440-458.
17. Bradley, J. S., Speech intelligibility in classrooms, *Journal of the Acoustical Society of America*, 1986, 80, 846-854.
18. Crandell, C. C. & Smaldino, J. J., An update of classroom acoustics for children with hearing impairment, *Volta Review*, 1994, 96, 291-306.
19. Beaman, C. P., Auditory distraction from low-Intensity noise: a review of the consequences for learning and workplace environments, *Applied Cognitive Psychology*, 2005, 19, 1041-1064.
20. Evans, G.W., & Lepore, S. J., Nonauditory effects of noise on children: A critical review, *Children's Environments*, 1993, 10, 31-51.
21. Hygge, S., Classroom experiments on the effects of different noise sources and sound levels on long-term recall and recognition in children, *Applied Cognitive Psychology*, 2003, 17, 895-914.
22. Hygge, S., Boman, E., & Enmarker, I., The effects of road traffic noise and meaningful irrelevant speech on different memory systems, *Scandinavian Journal of Psychology*, 2003, 44, 13-21.

Paper IV

The locus of effects of degraded speech signals on memory for spoken materials: Categorical but not serial processing is affected

Robert Ljung, Patrik Sörqvist and Anders Kjellberg

Laboratory of Applied Psychology, Centre for Built Environment

University of Gävle, Gävle, Sweden

Correspondence to

Robert Ljung

Laboratory of Applied Psychology

University of Gävle

E-mail: Robert.Ljung@hig.se

Abstract

The aim of the present paper was to examine if degraded listening conditions affect serial recall and/or free recall of auditory presented stimuli. Experiment 1a showed that a long reverberation time (RevT) does not disrupt serial recall of meaningful words, whereas it decreases free recall performance (Exp 1b). More detailed analyses revealed that memory of words presented in the beginning of the list were more impaired by the unfavorable listening conditions, which is in agreement with earlier findings. Experiment 2 examined recall of words with few or many phonological neighbours in good and degraded listening condition (long RevT or a low S/N ratio). In addition, the relationship between working memory capacity and the effects of degraded listening conditions was studied. People who made many invention errors in the operation span test performed significantly more false recall in low S/N. The same pattern was shown in Experiment 2; where the effect of phonological neighbor word lists was significantly related to the number of inventions made in the operation span test. This findings give support for the hypothesis that degraded stimuli activate more candidate words, and people who make those specific errors are have problem to suppress those candidates.

Keywords: Reverberation time, noise, recall, working memory capacity

Researchers' understanding of how acoustical conditions influence classroom performance is compromised with misconceptions. One of those misconceptions is that the acoustic characteristics of the classroom are acceptable as long as it allows for complete word identification (see for instance Shield & Dockrell, 2003). However, hearing what the teacher says is only a necessary, not a sufficient condition for people to understand and remember what is said. There is by now considerable evidence that points towards the possibility that, even though participants are able to identify what is said correctly, subsequent recall of the message is impaired by poor listening conditions (Ljung, Sörqvist, Kjellberg, & Green, 2009; Pichora-Fuller, 2003; Rabbitt, 1966, 1968). This has been shown for free recall of word-lists presented with low signal-to-noise ratio (Kjellberg, Ljung, & Hallman, 2008) and with long reverberation time (Ljung & Kjellberg, 2009), and it has been shown for cued recall of spoken lectures in similar listening conditions (Ljung et al., 2009). Many contemporary classrooms and other buildings where teaching and learning takes place have bad acoustic design (Crandell & Smaldino, 2000a, 2000b; Shield & Dockrell, 2008) and poor listening conditions thus constitute a tangible problem in society. The purpose of the present series of experiments was to investigate the locus of the effects of poor listening conditions on memory.

The locus of the effects of degraded speech signals on memory is still a matter of debate. One possibility is that masking of the speech signal makes it more difficult to serially rehearse the memory materials. Surprenant (1999) let participants listen to nonsense-syllables in low signal-to-noise (S/N) ratio and requested them to recall the syllables in order of presentation (serial recall). The experiment showed that poor listening conditions impaired recall across all list positions. One interpretation of this finding is that masking noise impairs serial order memory of the target sequence of non-words. Alternatively, and more likely, the impaired recall may have been the result of more incorrectly identified syllables in the low S/N condition. It is unclear at present if poor listening conditions disrupt serial recall of meaningful words in a similar way and if the effect is still obtained when word identification presents no problem. When a spoken word is to be identified it may activate more than one candidate (e.g., McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000), especially when the speech signal is deteriorated by masking noise or reverberation (Luce & Pisoni, 1998). Those candidates are usually words with a similar sounding as the target word—so called *lexical neighbors*—which compete with one another for identification. High neighbor density usually impairs identification and prolongs lexical-decision response latencies (e.g., Ziegler, Muneaux, & Grainger, 2003). We propose that poor listening conditions caused by prolonged reverberation time (RevT) impair memory for spoken messages since the deteriorated speech signal activates more candidates for word identification which have to be inhibited not to interfere with the recall of the target items. Specifically, when identification of the speech signal is optimal (as with short RevT), only a few candidates for identification are activated in semantic memory. When identification is suboptimal (as with long RevT), the uncertainty activates several competing candidates for identification in semantic memory (Stenfelt & Rönnberg, 2009). Even when the target is accurately identified, the other candidates remain active in semantic memory and will intrude into the recall protocol if not deliberately inhibited. This inhibition process not only affects the alternative candidates, but also the candidate that won the competition for identification, and thereby impairs recall (see Marsh, Hughes, & Jones, 2008, 2009 for a similar explanation of cross-modal auditory distraction).

Experiment 1a

The purpose of Experiment 1a was to test if a long RevT influences serial recall of meaningful words when word identification presents no problem. To this end, we used an order reconstruction task in which the participants cannot capitalize on the identity of the words to complete the task. Rather, efficient task performance relies more directly on serial-order processes. We expected no effect of RevT on this task, since we argue that the effect only appears when lexical-identity of the words form the basis of retrieval rather than serial order.

Method

Participants

A group of 30 undergraduate students (19-30 years of age) at the University of Gävle were recruited as participants. They all reported normal hearing, and normal or corrected-to-normal vision. The participants received a small honorarium as compensation.

To-be-remembered lists

The to-be-remembered items were eight auditory (i.e. spoken) words with the structure consonant-vowel-consonant-consonant. All words were recorded in a male voice in an anechoic room and downloaded on a computer. Out of these words, 39 to-be-remembered sequences were created. The order of the words within each sequence was random and each word occurred once within each sequence. The offset-to-onset interval between two successive words was 1000 milliseconds. Three versions of each sequence were created with Catt-Acoustics v8.0 software: one with RevT = 1.3 sec, one with RevT = 0.7 sec and one with RevT = 0 sec. This procedure made it possible to keep the word-sequences constant between participants while counter-balancing the RevT of the lists across participants.

Design and Procedure

A repeated measure design was used with two factors: reverberation time (three levels: 1.3 sec, 0.7 sec, and 0 sec) and serial position (eight levels). There were 10 lists in each reverberation time condition. The participants sat alone in a silent room in front of a computer screen. The to-be-remembered lists were presented through headphones in different random orders for each participant. The participants were told that their task was to remember the order of the words and to later recall them in their order of presentation. The participants began with three practice trials, one with each list type, before the experiment proper. One second after the last list word was presented a recall box appeared on the computer screen. The recall box contained all the words from the to-be-remembered sequence presented in a random order. The participants' task was to reconstruct the order of the sequence, by clicking on the words with the computer mouse in their order of presentation. Recall was self-paced and they had the possibility to undo clicks. When all words were marked, the participants pressed a button allowing for the next trial to begin.

Results

The responses were scored according to a strict serial recall criterion (i.e., items placed in the right serial position was scored as correct). As can be seen in Figure 1, the recall scores were approximately equal across the three RevT conditions. A 3(Reverberation time) × 8(serial position) analysis of variance revealed no effect of Reverberation time, $F(2, 58) = 1.46$, $MSE = 5.03$, $p = 0.24$, but an effect of serial position, $F(7, 203) = 66.16$, $MSE = 4.96$, $p < 0.01$. The interaction between the variables was non-significant, $F(14, 406) = 1.41$, $MSE = 1.93$, $p = 0.15$.

Please, insert Figure 1 about here

Discussion

As expected, Experiment 1a could not reveal an effect of RevT on the capability to process the order between individual target items (sequence reproduction). These results lend support to our assumption that poor listening conditions during encoding of meaningful words impair subsequent recall due to inefficient target activation and automatic activation of non-target materials phonologically related to the target materials.

Experiment 1b

The purpose of Experiment 1b was to investigate how degraded speech signals influence lexical-categorical retrieval (free recall). To this end, we investigated how S/N-ratio and RevT affect free recall. Experiment 1b also considered individual differences in cognitive capability that may be expected to influence the magnitude these effects. One typical measure of individual differences in cognitive capabilities is working memory capacity (WMC). WMC is commonly operationalized with complex-span tasks (i.e., a task that combines a processing component with a short-term memory component). Originally, Daneman and Carpenter (1980) suggested that WMC measures a pool of capacity (or cognitive resources) that must be divided between the storage and processing of information. That view of WMC has been applied most often to the investigation of individual and developmental differences (Bunting & Cowan, 2005), but more recent views suggest that WMC measures the capability to control attention (Kane & Engle, 2003), the capability to inhibit irrelevant information (Lustig, Hasher, & Zacks, 2008), the capability to search for items in secondary memory (Unsworth & Engle, 2007), or a combination of several functionally different capabilities (Sörqvist, Halin, & Hygge, 2010; Sörqvist, Ljungberg, & Ljung, 2010). According to the view that poor listening conditions activate irrelevant candidates for identification in semantic memory (Luce & Pisoni, 1998; Stenfelt & Rönnberg, 2009), high-WMC individuals should be less distracted by the masking of the speech signal because they are better at inhibiting irrelevant candidates and search for the appropriate candidate in memory. To investigate if inhibitory control contributes to the effect of degraded speech signals on memory, it might be fruitful to consider the role of prior-list intrusions (i.e., recalled items that were targets for recall on earlier trials but subsequently became irrelevant for recall) in the complex-span task used to measure WMC. Several researchers argue that the capability to inhibit or suppress proactive interference from previous trials or outdated information is responsible for individual differences in reading comprehension and other higher-order capabilities known to depend on the ability to suppress activation of irrelevant materials (Gernsbacher, 1993; Lustig, et al., 2001; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001; Passolunghi & Pazzaglia, 2004; see also Sörqvist et al., 2010a, 2010b). Because of this, it seems reasonable to assume that individual differences in suppression capabilities should influence the effect of poor listening conditions on free recall. Indeed, Kjellberg et al. (2008) found a negative correlation between WMC (as measured with reading span) and the effects of low S/N-ratio on free recall of spoken word lists This finding supports the assumption that WMC contributes to individual differences in susceptibility to effects of masked target signals. Moreover, some studies indicate that poor listening conditions increase recall of non-presented items (false recall; Ljung & Kjellberg, 2009; Rabbitt, 1966). Therefore, we assume that the participants should make more false recalls in the experimental conditions than in the control condition. That effect is supposed to be related to the individual's WMC; especially to the number of inventions in operation span task (new words that do not belong to any to-be-remembered list). This measure has proven to be a strong predictor for our ability to suppress irrelevant stimuli (Sörqvist et al, 2010).

Kjellberg et al (2008) used word lists with 50 items when they performed a recall experiment, and in order to explore the serial position effect they split up the word list into three parts in their analysis. The results showed that the noise effect was only significant in the primacy and recency part. The absence of any effect in the middle part of the list was explained as a result of a floor effect. Ljung & Kjellberg (2009) used the same paradigm in long RevT and control condition, their results showed an effect of long RevT only in the primacy part.

A word presented in degraded acoustics is likely to activate more words in our mental lexicon than the target word. Instead of holding the list of target items in memory, the degradation thus makes the memory task to include both target words and a number of activated phonological neighbors. That the effect of degraded stimuli is most pronounced in the primacy part is may be possibly be explained by word activation overload, which leads to lower correct recall and a higher number of inventions. In Experiment 1b we used shorter word lists containing eleven items in each list. Therefore we expect no floor effect at the mid positions, however, we still assume the performance to be worse in the mid positions due to the temporal distinctiveness.

The temporal distinctiveness model proposes that retrieval of a word is determined by the uniqueness of its temporal context, which acts as a cue for retrieval (Oberauer & Lewandowsky, 2008). Furthermore, the temporal distinctiveness models propose that items that are temporally isolated from their neighbours during list presentation are more distinct and thus should be recalled better, which means that items in the beginning and the end of a word list are easier to recall. Since, the initial and the final words are most distinctive and the middle words least distinctive, the distinctiveness are a function of the position in the series. At recall, search for middle-list items should be more easily confused with irrelevant (but still active) recognition-candidates bound to the same temporal position as the middle-list target words.

Method

Participants

A total of 24 university students (19-27 years of age) at the University of Gävle were recruited as participants. They all reported normal hearing, and normal or corrected-to-normal vision. The participants received a cinema ticket as compensation.

Tasks

Operation Span. To measure the participants' WMC an operation span test was used. It consisted of 25 simple arithmetical operations (e.g. "is (4/2) +1 = 4?") that were displayed on the screen in groups of two to six operations. The participants had to respond "yes" or "no" to each operation depending on the validity of the calculation. The response time was registered by the computer program. After responding (Yes or No), a to-be-remembered word was presented on the screen for one second. After two to six operations the participants got the instruction to recall the words. All to-be-remembered words were taken from the same semantic category (animals), in order to make the task more complicated, and increase the level of interference between lists. The operation span test was scored with the partial-credit scoring method suggested by Conway et al (2005). Five measures were registered during the operation span test: (I) the number of correctly recalled words, (II) the number of correct calculations, (III) the processing time of the operations, (IV) the number of inventions (recalled words that not were presented), (V) the number of prior-list intrusions (recalled words presented in an earlier set).

Word list memory

The participants were presented with nine word lists each containing 11 words. All nine lists were phonetically balanced to make identification equally easy between lists. To get ecologically valid school recordings, the word lists were played from a loudspeaker in two ordinary classrooms (one with long reverberation time and one with short) and recorded binaurally. These binaural recordings were presented to the participants through headphones. The classrooms were about the same size (length 10 m, width 6 m, height 3 m) and were furnished with desks. The loudspeaker was placed 1 m in front of the blackboard in the centre of the classroom at a height of 1.5 m, and an acoustical head was placed as a seated student at a desk in the back of the room, about 6 m in front of the loudspeaker. The classroom with short reverberation time had various absorbing panels on the walls and the ceiling. The reverberation time was 0.3 s in all octave bands from 125 Hz to 4 kHz. The classroom with long reverberation time had some absorbing panels, but the walls and the ceiling was mostly painted concrete, and the reverberation time was 1.84 s at 125 Hz, 1.46 s at 250 Hz, 0.94 s at 500 Hz, 0.77 s at 1 kHz, 0.78 s at 2 kHz and 0.68 s at 4 kHz. The recordings from the class room with long reverberation time were used in the RevT condition, the recordings from the classroom with short RevT was condition used in the control; in the noise condition the recordings from the short RevT were mixed with broadband noise, to a signal-to-noise ratio of +4 dB.

Design and procedure

The computer, with headphones attached, was rigged in a sound isolated room. The participants were given a short oral instruction of the purpose of the experiment. All participants first performed the operation span test in silence. The participants listened to nine words lists each containing 11 words (three heard in short RevT, three with background noise, and three in long RevT), with recall direct after each presented list. Directly after the experiment, the participants were debriefed and thanked.

Results

Word list recall: the 11 words lists in the control, RevT and Noise conditions

Figure 2 shows the serial position curve for all three listening conditions. Based on these curves the primacy part was defined as the first three positions and the recency part as the three last positions. The remaining five positions constituted the mid part of the curve. Mean recall was calculated for these three parts of the word lists (Figure 3). A 3 (listening conditions) X 3 (list positions) ANOVA for repeated measurements was made of the data including tests of linear and quadratic trends.

Please, insert Figure 2 and Figure 3 about here

There was a significant quadratic trend over the three parts of the list, $F(1, 23) = 54.62$, $MSE = 0.32$, $p < .001$, $\eta^2 = 0.70$.

The main effect of listening condition was not significant, $F(2, 46) = 2.62$, $MSE = 0.16$, $p = .084$, but there was a significant interaction between condition and position in the quadratic trend, $F(1, 23) = 4.91$, $MSE = 0.09$ $p < .05$, $\eta^2 = 0.18$. Follow-up tests showed that the RevT condition differed from control in the primacy part of the list ($t = 3.91$, $p < .01$).

Please, insert Table 1 about here

The relation between operation span and word list recall

The strongest and most consistent correlations between operation span and word list recall were obtained between number of recalled words and prior-list intrusions in the operation span task and recall of the primacy part of the lists (Table 1). Inventions and response times were uncorrelated with word list recall. But inventions made in OSPAN were related to the noise effect on inventions. This was shown with a regression analysis where inventions in the noise condition was entered as dependent variable, and inventions in control as first independent variable, and inventions in OSPAN entered in the second step (b = -0.43, t = 2.13, p = .045). To analyse the relation between operation span and the effect of noise and RevT, analogous regression analyses were performed. In these analyses the score in the noise or RevT condition was entered as dependent variable and the score in the control condition as the first independent variable. In the next step one of the operation span scores was entered to test whether these scores could predict the residual score. These analyses showed that delayed intrusions were significantly related to the effect of noise on the primacy score (b = -0.51, t = 3.69, p < .001). No corresponding effect was obtained for the mid or recency part of the lists. Neither was there any significant relation between operation span and the effect of RevT.

Discussion

As predicted, unfavorable listening conditions only affected the primacy part of the list. This is in agreement with earlier findings (Kjellberg *et al* 2008; Ljung & Kjellberg, 2009). The effect of the background noise on the recency part obtained by Kjellberg *et al* (2008) therefore probably was a chance effect. Surprenant (1999) showed that low S/N affected recall equally over all serial positions, which is inconsistent with the present data. However, Surprenant used nonsense syllables as to-be-remembered stimuli material; nonsense syllables are not present in our mental lexicon and therefore no cognitive mismatch could be possible. Furthermore, it is possible that the obtained memory effect was a result of false identification.

The analyses regarding the relation between WMC and the effect of degraded stimuli showed that the delayed intrusions measure in the operation span test was significantly related to the effect of noise on the primacy score. As predicted, the number of inventions in OSPAN was related to the increase of false recall in the noise condition, which gives evidence for the reasoning that the degradation of stimuli activates more phonological candidates. Thus people that tend to make many inventions in OSPAN have problems to suppress the activated candidates from the target words. The result that recall was significantly worse in the mid part, in Experiment 1b give further evidence for the temporal distinctiveness phenomenon (Oberauer & Lewandowsky, 2008).

Experiment 2

As already mentioned, numerous experiments have shown that the number of phonological neighbours that are activated during recognition is a powerful predictor of performance on auditory targets (Ziegler, Muneaux, & Grainger, 2003; Cluff & Luce, 1990; Luce & Pisoni, 1998; Luce, Goldinger, Auer & Vitevitch, 2000; Rönnberg, 2003). However, to our knowledge no published study has explored the relationship of RevT, phonological activation and serial position.

The extents to which there are problems to hear what is said should vary between words depending on how easily identified their phonemes are (Cluff & Luce, 1990). The phonemes most vulnerable to masking effects, and thus the ones most open to alternative interpretations of the phoneme, are speechless consonants (e.g. *f, k, p* and *s*). The extent to which this ambiguity give room for alternative interpretations of the word (i.e. the number of *phonological neighbors*), determines how difficult and time consuming the identification of the word will be (Marslen-Wilson & Warren, 1994). When listeners are asked to classify the initial sound during presentation of series of spoken non-words, they tend to classify the initial sound as if it belonged to a real word. For example, the initial sound of the non-word ‘peef’ tends to be classified as ”*b*“ (as in ‘*beef*’), and the initial sound of the non-word ‘beace’ tends to be classified as ”*p*” (as in ‘peace’), suggesting that lexical categorization influence phonetic perception (Newman, Sawusch & Luce, 1997). This may suggest that there are similarities in how real words and non-words are perceived, especially when the phonetically similarity between the non-word (‘beace’) and its matching real words (‘peace’) is large. According to this account, degraded speech signals should impair categorical-lexical based free recall. Thus, the probability for word-recognition mismatch should be larger with words with many phonological neighbors. The fewer alternative interpretations the more easily and faster will the word be identified. One possibility is therefore that high neighborhood density impairs recall because fewer resources are available after identification for further processing and storage. Another possibility is that more neighbors activate more candidates which compete for identification, and an inhibition process must be employed to suppress them from the recall protocol. High neighborhood density thus requires more inhibition than low density and impairs recall because the inhibition processes accidently spread suppression to target items. The main question in experiment 2 thus was whether the effect of a long RevT on recall of words lists is larger when the list contains words with many phonological neighbors than when the words have no or few phonological neighbors. Additionally, according to temporal distinctiveness theories of short-term memory (Murdock, 1960; Neath, 1993), we expect that the effect should be most prominent at the mid positions in the word list. Furthermore, we expect low WMC participants to make more false recall in the RevT condition.

Method

Participants

A total of 24 participants (20-42 years of age) at the University of Gävle were recruited as participants. They all reported normal hearing, and normal or corrected-to-normal vision. The participants received a cinema ticket as compensation.

Tasks

Operation Span

This task was the same as in Experiment 1b.

Word lists

16 word-lists with 7 nouns in each list were used as to-be-remembered material. The words in eight of these had many phonological neighbors, and eight lists had no phonological neighbors. A word with a many phonological neighbors was defined as one where a change of a voiceless consonant to another consonant yielded at least three common Swedish words. The words were chosen with the help of crossword puzzle programs, by searching for fitting nouns when the positions of the voiceless consonants were left open. The participants listening to 16 word lists, 8 lists in long RevT and 8 lists in short RevT). Directly after the experiment, the participants were debriefed and thanked.

Results

Word list recall: the seven words lists with or without phonological neighbors in the control and RevT conditions

Figure 4 shows the results for both acoustic conditions (long RevT and control) and both types of word lists. Based on this serial position curve, the primacy part was defined as the first two positions and the recency part as the last position. Mean recall was calculated for the primacy and the mid part (positions 3-6) of the lists. Data were analyzed with a 3(Serial Positions: primacy vs. middle vs. recency) × 2(RevT: long vs. short) × 2(Phonological Neighbors: high density vs. low density) ANOVA including linear and quadratic trends. Apart from the general quadratic trend over the positions *F*(1, 23) = 31.89, *MSE* = 0.58, $p < .001$, $\eta^2 = 0.59$, the analysis showed that fewer words were recalled when the lists was presented with the longer RevT (means: 3.06 and 2.78, respectively; *F*(1, 23) = 13.25, *MSE* = 0.44, $p <.001$, $\eta^2 = 0.37$). This effect did not differ between positions of the curve as indicated by the non-significant interaction ($p = .40$) between Serial Position and RevT. There was no main effect of Phonological Neighbors and no significant interaction between Condition and Phonological Neighbors, but as shown by Figure 5, the trend over Serial Positions differed between lists containing words with few or many Phonological Neighbors, *F*(1, 23) = 6.47, *MSE* = 0.18, $p < .05$, $\eta^2 = .22$). Follow-up tests showed that the significant effect of neighbors was only obtained in the mid part of the list *t*(23) = 2.47, $p < 0.05$.

Please, insert Figure4 and Figure 5 about here

The relation between operation span and the effect of RevT and phonological neighbours

The same types of regression analyses as were done in Exp 1b were performed to test the relation between operational span and the effect of RevT and phonological neighbours. Mean performance with long and short RevT as well as with or without neighbours were calculated. The effect of phonological neighbours was significantly related to the number of inventions in OSPAN, *b* = -.29, *t*(21) = 2.12, $p < .05$. This was true both for the long and short RevT. The recall of the mid part and the recency part of the lists were not significantly related to any indictor of operation span performance (see Table 2).

The effect of long RevT on the primacy part of the list was significantly related to the number of correctly recalled words in the operation span task, *b* = .58, *t*(21) = 2.99, $p < .01$. This relation was also significant for the mid part of the lists task, *b* = .29, *t*(21) = 2.49, *p* <

.05. For the recency part the effect was significantly related to delayed intrusions, $b = -.55$, $t(21) = 2.90$, $p < .01$.

Please, insert Table 2 about here

General discussion

The purpose of the present series of experiments was to investigate why poor listening conditions (low S/N-ratio and long RevT) impair memory for spoken materials. Long reverberation time did not disrupt serial recall for meaningful words (Experiment 1a), but it did disrupt free recall (Experiments 1b and 2). Moreover, more detailed analyses revealed that words presented in the beginning of the list were more impaired by the unfavorable listening conditions, which is in agreement with earlier findings (Pichora-Fuller, 2003; Kjellberg et al. 2008; Ljung & Kjellberg, 2009). Furthermore, people who made many invention errors in OSPAN performed significantly more false recall in low S/N (Experimant1b), which gives support for the hypothesis that degraded stimuli activate more candidates (Cluff & Luce, 1990), and that people who make that specific errors are have problem to suppress those candidates.

The present investigation also considered individual differences in WMC to further tease apart the processes that contribute to the effect of poor listening condition on short-term recall. In Exp 1b the analyses showed that the prior-list intrusions measure in the operation span test was significantly related to the effect of the degraded listening conditions on primacy recall performance. In other words, people that fail to suppress or delete stored information from earlier tasks are more vulnerable to noise. Furthermore, the invention error measure in OSPAN was significantly related to the false recall in the low S/N condition. This finding supports the theory that (I) degraded stimuli activates more phonological neighbours, which shows up as a increase of false recalls, (II) people that make this certain kind of invention errors in OSPAN cannot inhibit those activates neighbours properly and therefore recall wrong word. That hypothesis was verified in Experiment 2, were the effect of phonological neighbour word lists was significantly related to the number of inventions made in the operation span test, this was true both for the RevT condition and the control condition. The effect of phonological neighbour word lists was shown by a significant trend over the serial positions. However the follow up test revealed that the effect was concentrated to the mid positions of the list. This effect was expected and the results are in correspondence with the temporal distinctiveness theory (Oberauer & Lewandowsky, 2008). Since, the mid positions are less distinct than the primacy and recency position, the recall curve is U-shaped. As far as the performance in the mid positions not reach a floor effect that part could be most vulnerable to additional lexical distinctiveness.

The main effect of low S/N ratio and long RevT on auditory to-be-remembered word lists are consistent with earlier findings (Kjellberg *et al* 2008; Ljung & Kjellberg, 2009; Rabbitt, 1966, 1968). In these papers this effect was explained in terms of limited capacity; when we listening to degraded auditory messages more mental recourses are allocated to listening and interpreting the message and therefore less spare capacity are left for storage. However, there are mixed result for the relation between WMC and the effect of bad acoustics on recall of speech (Ljung & Kjellberg, 2009). Sörqvist et al. (2010) showed that reading comprehension performance was related to prior-list intrusions errors in the WMC test, and people's susceptibility to auditory distraction was related to the participant's immediate intrusion errors in the WMC test. Thus, maybe there are other measures of WMC that give more reliable results for cognitive demanding listening condition.

Studies using auditory to-be-remembered stimulus material have one obvious problem all results of memory performance can be an effect of incorrect word identification. Rabbitt (1966) instructed his participants to repeat aloud all stimuli words and deleted all results from participants with any false identification, this procedure led to a sorted group of participants. Kjellberg et al. (2007) and Ljung & Kjellberg (2009) let their participants repeat all to-be-remembered words, and scored later recall in accordance with the perceived words. However, repeating the stimuli can lead to better recall performance and therefore there is reason to believe that the effects actually are stronger then reported. In the present study no repeating procedure was used, however, false identification should effect all list positions equally and therefore the interactions between position and condition supports our theory that all effects reported here is due to memory effect, and the effects are not confounded by missed identification.

References

Bunting, M, F., & Cowan, N. (2005) Working memory and flexibility in awareness and attention. *Psychological Research, 69*, 412–419.

Cluff, M. S. & Luce, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception & Performance, 16*, 551-563

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

Crandell, C., & Smaldino, J. (2000a). Classroom Amplification Technology: Theory and Practice. *Language, Speech, and Hearing Services in Schools; 31*, 4, 371-375.

Crandell, C., & Smaldino, J. (2000b). Classroom Acoustics for Children with Normal Hearing and with Hearing Impairment.Language, Speech, and Hearing Services in Schools, 31, 4 362-370.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading*. Journal of Verbal Learning & Verbal Behavior, 19*, 450–466.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science, 4*, 294-298

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47-70.

Kjellberg, A., Ljung, R., & Hallman, D. (2008). Recall of words heard in noise. *Applied Cognitive Psychology, 22,* 1088-1098.

Ljung, R., & Kjellberg, A. (2009). Long reverberation time decreases recall of spoken information. *Building Acoustics, 16,* 301-312.

Ljung, R., Sörqvist, P., Kjellberg, A., & Green, A. M. (2009). Poor Listening Conditions Impair Memory for Intelligible Lectures: Implications for Acoustic Classroom Standards. *Building Acoustics, 16*, 3, 257–265

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics, 62*, 615–625.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear & Hearing, 19*, 1–36.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General, 130,* 199-207.

Lustig, C., Hasher, L., & Zacks, R. T. (2008). Inhibitory deficit theory: Recent developments in a "new view". In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in Cognition* (pp. 145-162). Washington, DC: American Psychological Association.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic memory: A process- based approach. *Journal of Memory & Language, 58*, 682-700.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition, 110*, 23-38.

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 4*, 653-675.

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

Murdock, B.B. (1960). The distinctiveness of stimuli.*Psychological Review, 67*, 16–31.

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory and Cognition, 21*, 689–698.

Newman, R. S., Sawusch, J. R. & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception & Performance, 23*, 873-889.

Norris, D., McQueen, J. M., / Cutler, A. (2000). Merging information in speech recognition: Feedback in never necessary. *Behavioral and Brain Sciences, 23,* 299-325.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*, 3, 544–576

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*, 344–354.

Passolunghi, M. C., & Pazzaglia, F. (2004). Individual differences in memory updating in relation to arithmetic problem solving. *Learning & Individual Differences, 14*, 219–230.

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing, *International Journal of Audiology, 4*, 26-32.

Rabbitt, P. (1966). Recognition memory for words correctly heard in noise. Psychonomic Science, 6, 383 384.

Rabbitt, P. (1968). Channel-capacity, intelligibility and immediate memory, *The Quarterly Journal of Experimental Psychology, 20*, 241-248.

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngologica, Supplementum, 476,* 167-176.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology, 42*, S68–S76.

Shield, B. M., & Dockrell, J. E. (2003). The effects of noise on children at school: A review, Building Acoustics, 10, 97-116

Shield, B. M., & Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. J. Acoust. Soc. Am. 123, 133–144.

Stenfelt, S., & Rönnberg, J. (2009). The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology, 50*, 385–393

Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables. *International Journal of Psychology, 34,* 328-333.

Sörqvist, P., Halin, N., & Hygge, S. (2010). Individual differences in susceptibility to the effects of speech on reading comprehension. *Applied Cognitive Psychology, 24,* 67-76.

Sörqvist, P., Ljungberg, J. K., & Ljung, R. (2010). A Sub-Process View of Working Memory Capacity: Evidence from Effects of Speech on Prose Memory. *Memory, 18*, 310 - 326.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104–132.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language, 48,* 779-793.

Figure Captions

Figure 1. Recall scores for the three RevT conditions (RevT 0sec, RevT 0.7sec, RevT 1.3sec) across the eight serial positions.

Figure 2. Serial position curve for all three listening conditions (RevT, Noise, Control) together.

Figure 3. Mean recall for primacy part, mid part and recency part of the word list. Primacy part is the mean score for the first three items and the recency part is the mean score for the two last positions, and the remaining six positions constitute the mid part of the curve.

Figure 4. Mean recall score in the seven positions of the word list, both word list with or without phonological neighbours, and both conditions (long RevT and Control) together.

Figure 5. Mean recall for primacy part, mid part and recency part of the word list. Primacy part was defined as the first two positions and the recency part as the last position, and the remaining four positions constitute the mid part of the curve.

Figure 1

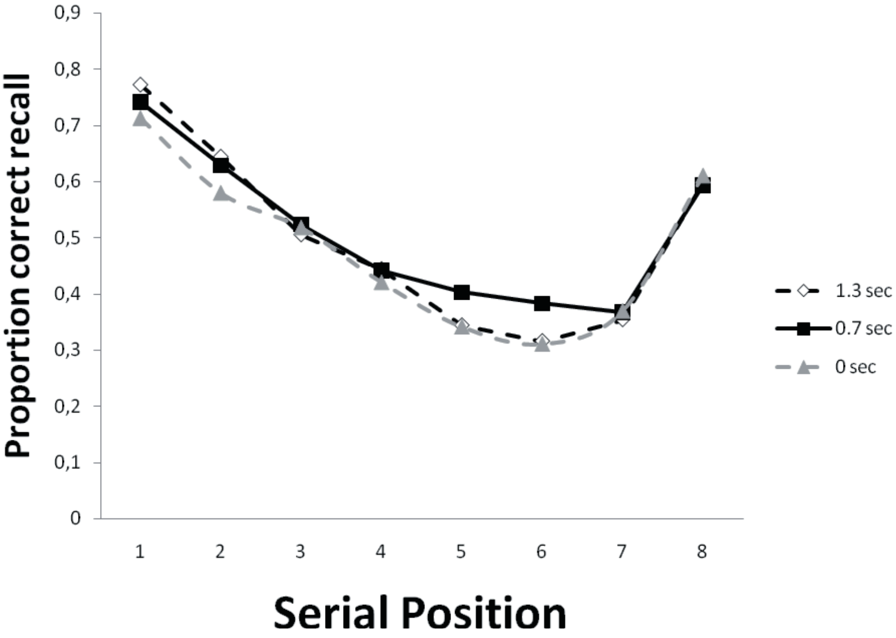

0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
Proportion correct recall
1
2
3
4
5
6
7
8
Serial Position
1.3 sec
0.7 sec
0 sec

Figure 2

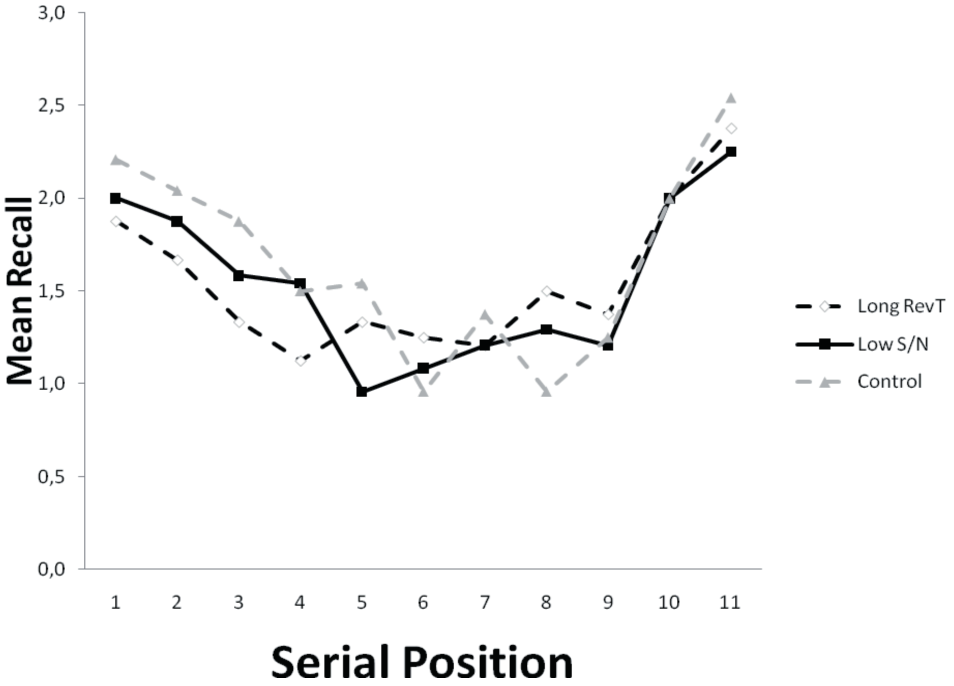

3,0
2,5
2,0
1,5
1,0
0,5
0,0
Mean Recall
1
2
3
4
5
6
7
8
9
10
11
Serial Position
Long RevT
Low S/N
Control

Figure 3

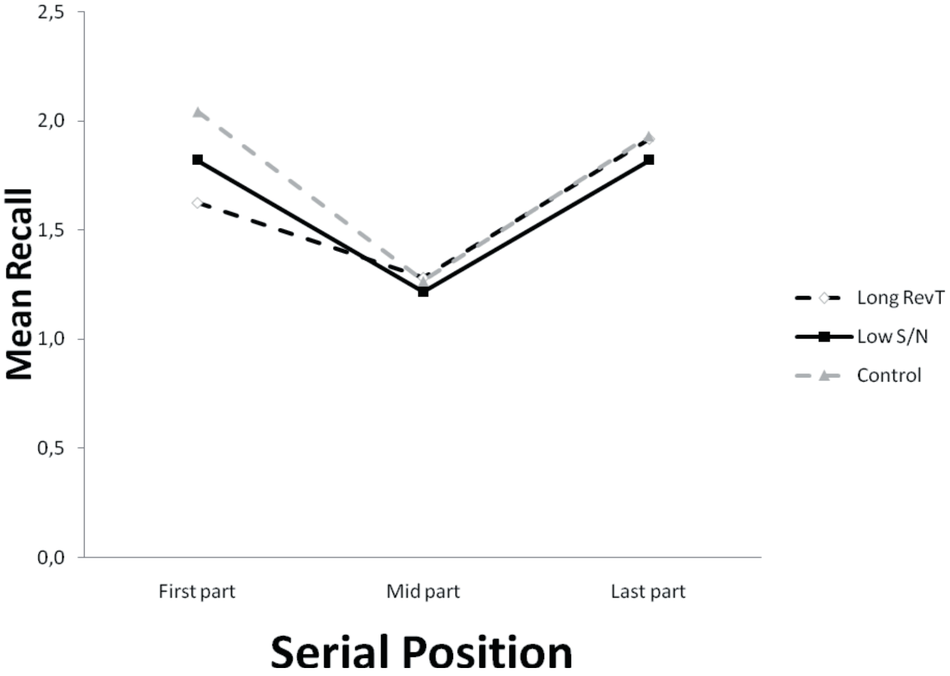

2,5
2,0
1,5
1,0
0,5
0,0
Mean Recall
First part
Mid part
Last part
Serial Position
Long RevT
Low S/N
Control

Figure 4

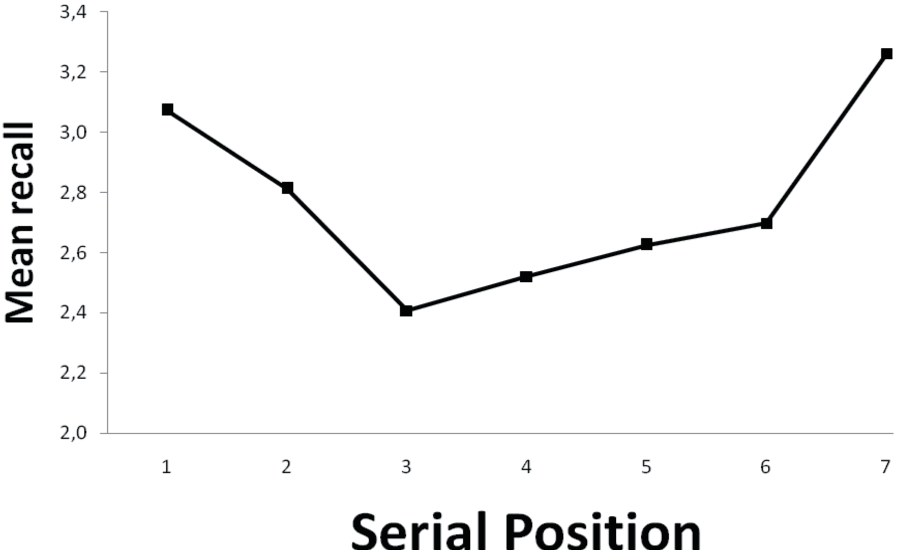
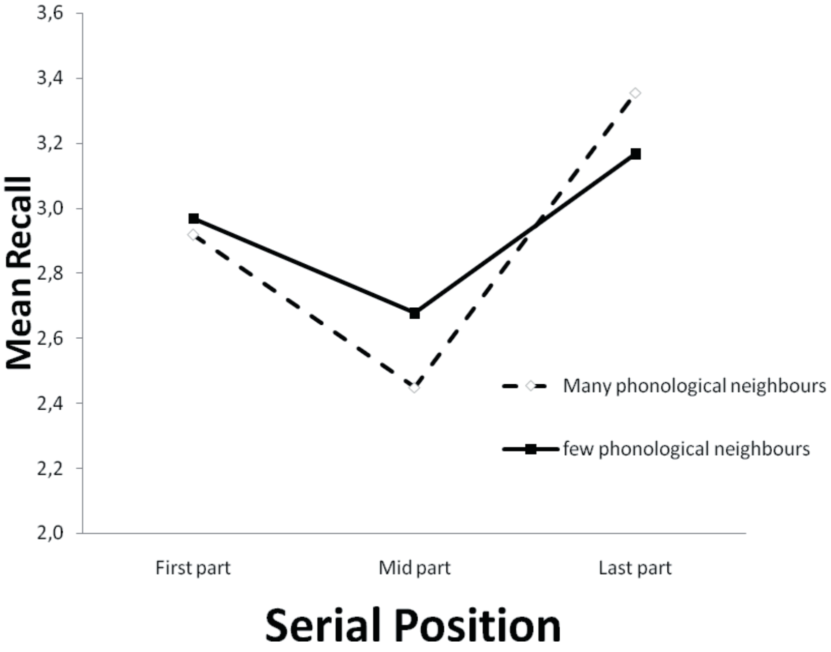


Figure 5



3,6
3,4
3,2
3,0
2,8
2,6
2,4
2,2
2,0
Mean Recall
Many phonological neighbours
few phonological neighbours
First part
Mid part
Last part
Serial Position

Table 1. Product-moment correlations between word list recall and the four measures of operation span performance. The data are divided in three parts of the word lists (primacy, mid part, recency) and the three experimental conditions (Long RevT, Low S/N, Control).

		Operation span recalled words M = 49.13 SD = 5.57	Operation span prior-list intrusions M = 2.25 SD = 2.27	Operation span inventions M = 0.17 SD = 0.48	Operation span response time M = 2.49 SD = 8.40
Primacy Long RevT	M = 1.63, SD = 0.66	0.39	-0.48*	-0.20	0.14
Primacy Noise	M = 1.82, SD = 0.85	0.61**	-0.68**	-0.17	0.22
Primacy Control	M = 2.04, SD = 0.73	0.47*	-0.37	-0.02	-0.01
Mid part Long RevT	M = 1.28, SD = 0.48	0.01	-0.22	0.09	0.34
Mid part Noise	M = 1.22, SD = 0.65	0.38	-0.46*	-0.04	0.14
Mid part Control	M = 1.27, SD = 0.60	0.37	-0.33	-0.16	0.17
Recency Long RevT	M = 1.92, SD = 0.57	0.32	-0.12	0.00	0.00
Recency Noise	M = 1.82, SD = 0.59	0.17	-0.05	0.26	-0.19
Recency Control	M = 1.93, SD = 0.51	0.47*	-0.16	-0.25	0.17

** Correlation is significant at the 0.01 level.

* Correlation is significant at the 0.05 level.

Table 2. Product-moment correlations between word list recall and the four measures of operation span performance. The data are devided in three parts of the word lists (primacy, mid part, recency) and the two list types (phonological neighbours and No phonological neighbours) and the two experimental conditions (Long RevT, short RevT).

		Operation span recalled words M = 48.25, SD = 7.01	Operation span prior-list intrusions M = 1.88, SD = 2.14	Operation span inventions M = 0.23, SD = .56	Operation span response time M = 2.79, SD = 10.01
Short RevT Neighbours Primacy	M = 3.02, SD = 1.08	0.27	-0.17	-0.46*	0.10
Short RevT, No neighbours Primacy	M = 3.29, SD = 0.72	0.34	-0.18	-0.15	0.48
Long RevT Neighbours Primacy	M = 2.81, SD = 0.87	0.63**	-0.42*	-0.30	-0.04
Long RevT , No neighbours Primacy	M = 2.65, SD = 0.94	0.35	-0.32	-0.11	0.30
Short RevT Neighbours Midpart	M = 2.59, SD = 0.73	0.41*	-0.32	-0.16	0.06
Short RevT, No neighbours Midpart	M = 2.73, SD = 0.65	0.67**	-0.34	-0.20	0.16
Long RevT Neighbours Midpart	M = 2.3, SD = 0.68	0.58*	-0.30	-0.27	0.16
Long RevT , No neighbours Midpart	M = 2.63, SD = 0.75	0.73**	-0.57**	0.01	0.08
Short RevT Neighbours Recency	M = 3.33, SD = 0.82	-0.21	0.19	-0.28	-0.01
Short RevT , No neighbours Recency	M = 3.42, SD = 0.58	0.03	0.11	-0.23	0.29
Long RevT Neighbours Recency	M = 3.38, SD = 0.65	0.04	0.09	0.04	-0.09
Long RevT , No neighbours Recency	M = 2.92, SD = 0.83	0.05	-0.34	0.13	-0.13

** Correlation is significant at the 0.01 level.

* Correlation is significant at the 0.05 level.