Recall of spoken words presented with a prolonged reverberation time

Robert Ljung*, Anders Kjellberg

Centre for Built Environment, University of Gävle, Sweden

*corresponding author: e-mail: robert.ljung@hig.se

ABSTRACT

The aim of the study was to explore if a long reverberation 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 role of working memory capacity for 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. The reading span score was unrelated to recall performance.

INTRODUCTION

One common indicator of speech intelligibility is the signal-to-noise ratio. When the difference between signal and noise decreases, the listeners lose information and have to rely more on redundancy and contextual cues to understand the message. Reverberation time (RT) is another parameter that may have the same effect. RT 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 the surfaces in the room. 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 (Hodgson & Nosal 2002). 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.

A previous study of Kjellberg et al. (2008) showed that a background noise impairs 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 loudly repeated each word 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 primacy and recency parts of the list. Their interpretation of the noise effect was that that the noise made word identification more difficult, which left less 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. However, the noise was continuous and an alternative interpretation therefore would be that the background noise disturbed rehearsal and encoding processing between the presented words. In the present experiment identification of the spoken words was made more difficult by presenting the words in a virtual room with a long RT. Since RT distorts the signals without affecting the pauses between them, an effect of RT on recall of words is not open to this alternative interpretation.

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, there are a few exceptions. Beaman & Holt (2007) presented distracting irrelevant sounds with different RT during memory tasks with visual stimuli. They used an extremely long RT (5 s), which smoothened the distracting sound and therefore should decrease the “irrelevant sound effect” on serial recall as predicted by the changing-state hypothesis (Jones & Macken 1993). Their results confirmed this hypothesis, this is interesting as a test of the changing state theory, but have small practical importance since five seconds RT is unrealistic. Perham et al. (2007) performed a similar study but with more realistic RT values (0,7 and 0,9 seconds) for the distracting sound, and found no effect on serial recall for the visual presented stimuli.

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 have a very poor acoustic quality (Seidel et al. 2005). 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 (Helfer & Wilber 1990). The situation is even worse if it turns out that understanding and memory of what is said may be impaired also with RT’s that only make it more effortful to listen.

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 this effect. Kjellberg et al. (2008) found such a relationship for the noise effect but only on the recency part of the serial position curve.

The present study is a near replication of Kjellberg et al. (2008) with the background noise substituted for a long RT. Words and sentences were presented with a long or a short RT, and working memory capacity was tested with a test of reading span. The objective was to test the following three hypotheses:

- Recall of words is better when they are presented with a short RT.
- Recognition of sentences is less sensitive to the RT than the recall of words, but prolongs response times. Kjellberg et al. (2008) found no effects on recognition, but they used an easier task than in the present study.
- The better the working memory capacity, the less effect does the long RT have on recall and recognition.

A further aim was to analyze how the RT affected the recall of different parts of the list of items-to-be-remembered by comparing the serial position curves for the long and short RT conditions.

METHOD

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 or short RT. The order between conditions was counterbalanced.

Apparatus

The experiment was conducted in an anechoic chamber with the subject 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. For visual presentation of the reading span and recognition tests a laptop was used.

The speech and acoustic conditions

The speech stimuli were a part of a package of standardized tests for speech audiometry (Hagerman 1982). The speech had an equivalent sound level of 64 dB(A), and was mixed with broadband noise to get a S/N-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 had the same size (length 10 m, width 6 m, height 3 m) and were furnished with 30 desks. The sound source was placed one meter ahead the blackboard in the center of the classroom at a height of 1.7 meters, and the receiver was placed 6.6 meters in front of the source at a height of 1 meter. The classroom with short RT had various absorbing panel on the walls and the ceiling, and 30 pupils where seated in the desks. In the classroom with long RT only 15 pupils where seated and some absorbing panels where substituted with concrete walls. In the short RT condition mean RT 0.25-4 kHz was 0.53 s (with max RT 0.58 at 0.25 kHz) and in the long RT condition it was 1.17 s (with max RT 1.41 at 0.125 kHz). The STI values indicate that the two conditions stood for very good (73.5) and fair (56.1) intelligibility, respectively.

Performance tests

Reading Span test. Working memory capacity was assessed with the reading span test, which was taken from the cognitive test battery TIPS (Hällgren et al. 2001). 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 in a word-by-word fashion. Each word was shown on the screen for 0.8 s. The inter-word interval was 0.075 s. Half of the sentences were absurd (e.g., "the house read a newspaper"), and half normal (e.g., "the pupil came too late"). The subjects' task was to indicate, during a 1.75 s 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.

Hagerman's sentences (recognition test). The Hagerman test 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 (Hagerman 1982). The subjects' task was to memorize the orally presented sentences for later recognition. There were approximately seven seconds of silence between sentences; the subjects repeated each sentence aloud to check that they had identified it. 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, ten 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 reaction times were measured.

Phonetically balanced word list (recall test). 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 tests for speech audiometry (Hagerman 1982). The task was to memorize the words for later recall. There were approximately three seconds of silence between each presented word when the subject was asked to repeat the word aloud in order to check whether 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 a 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 words recalled of those that had been correctly identified (correct words). The lists of words and sentences have previously been shown to be equally intelligible (Hagerman 1982; Magnusson 1995).

Rated effort

To validate the assumption that word identification became more effortful by the long RT 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 performed the Reading Span Test in silence followed by the auditory recall (word list) and recognition (sentences) tests, with short RT and long RT (the order between the two conditions was counterbalanced).

Altogether, 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 RT 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 (RT conditions X order of conditions) showed that effort and the number of incorrectly repeated words and sentences were significantly higher in the Long RT condition (Table 1).

Table1: 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 number recalled of correct and stated words were almost perfectly correlated (.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 of recall performance. Recall of stated words was significant lower in the Long RT than in the Short RT condition (mean=12.97 and 10.78 respectively ($F(1,30) = 7.67$, $p=0.01$, $eta^2=.20$). The 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 whereas no such difference was seen in the long RT condition. This was shown as an interaction between condition and presentation order ($F(1,30) = 4.32$, $p<0.046$, $eta^2=.25$).

The interaction between order and RT was primarily the result of four subjects that performed extremely 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. An interview directly 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 with these four subjects excluded showed that the main effect of RT conditions remained significant ($F(1,25) = 6.31$, $p =<0.019$, $eta^2=.20$) despite the loss in mean value difference (2.19 in the whole group and 1.58 in the reduced group). Furthermore, in this 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 by Figure 1, the RT effect had only an influence on the recall of the first two parts of the list. This was reflected in an interaction between condition and parts in the linear trend ($F(1, 30) = 10.16$; $p = 0.003$, $eta^2=.25$). A test of the difference between RT conditions in the first two parts of the list showed that this effect was significant also after exclusion of the four subjects with an extreme order effect ($F(1,26) = 16.31$, $p<0.001$, $eta^2=.39$) 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).

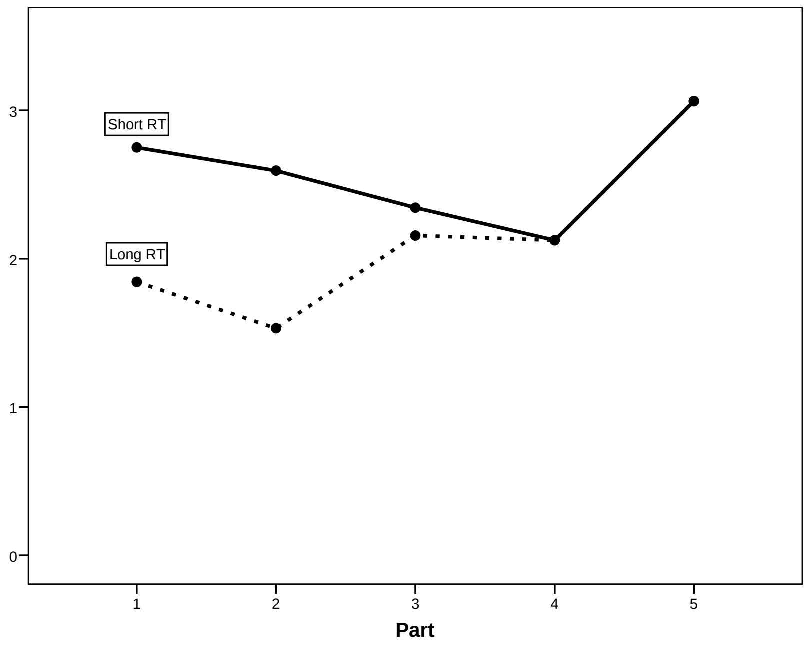

Figure 1: Correct recalled words in the five parts of recall test in the Long RT and Short RT conditions.

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. Neither was there any effect of presentation order or interaction between order and condition for any of the accuracy measures. However, an analysis of response times revealed that when sentences were heard in the short RT condition subjects were faster to identify that a sentence had not been presented previously (Short RT M = 3.03 s, Long RT M = 3.42 s, $F(1,31) = 4.62$, p =<0.04, eta^2=.13).

Relation between working memory and RT effect. Correlations were calculated for the relation between reading span score and recall score overall and for the five parts of the word lists. No significant correlations were found. Neither was there any correlation between reading span score and the effect of RT on recall. A corresponding analysis was made for the sentence recognition test, with the same result.

DISCUSSION

In line with the hypothesis subjects recalled fewer words when the word list had been presented with a long RT. The RT effect was most pronounced at the beginning of the word list. Recognition of sentences was expected to be less sensitive to long RT, which was true, but measurements of response time revealed faster responses then sorting out irrelevant sentences. Contrary to the hypothesis, reading span performance was unrelated to both recall and recognition as well as to the effect of RT in these tests.

The effect of the long RT was apparent on the recall of the primacy part of the word list; which indicates that the long RT impaired the encoding and transfer to long-term storage, alternatively early consolidation in the long-term memory.

In line with the noise effect found in the previous study (Kjellberg et al. 2008) recall of words was impaired in the deteriorated listening condition. They assumed that the critical effect of the noise was to make word identification more cumbersome. The alternative interpretation was that the noise between the words disturbed the short- or long-term storing of the word. In the present study this interpretation was excluded since only the speech signal was affected by the RT, making the two conditions identical in the pauses between the words. In the previous noise experiment there were both recency and primacy effects, but in the present study there was only a primacy effect. This indicates that the effect of the noise on the recency part in the former study was a result of interference with the rehearsal process in working memory by the noise in the pauses between the words. The primacy effect obtained in both studies thus probably is an effect of the degraded signal, which is more resource demanding to listen to and to understand. This leaves fewer resources and less time for the transfer to long-term storage, alternatively early consolidation in the long-term memory. In the previous study (Kjellberg et al. 2008) the results from the working memory test lent some support to this interpretation. They found a significant correlation between the subject's working memory capacity and the noise effect on recall, but this was just true for the mid and last part of the list (recency part), not for the first part (primacy). In the absence of a recency effect in the present study we obtained no significant correlation between working memory capacity and the effect of RT on recall. A strong order effect might conceal such a correlation, but this seems unlikely in this case since the exclusion of the four subjects with the strongest order effect did not change the result.

As predicted the RT effect on recognition performance was restricted to the reaction time measures. This effect was shown as a shorter 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-remember list of sentences before you are able to determine that it was not presented. That task therefore is more demanding than the recognition of previously presented sentences and should be more vulnerable to bad listening conditions.

The result supports the hypothesis that a degraded signal impairs recall. A further prediction was that the noise effect could be less severe for persons with a better working memory capacity (Pichora-Fuller et al. 1995; Pichora-Fuller 2003). The present study gave no support for that hypothesis. Maybe, the reading span test used in the present study was not sensitive enough because only number of correct responses could be used as dependent variable. It would be interesting to expand the reading span test and measure reaction time, or add tests taping other executive functions in a future study.

The present study demonstrated that a long RT may disrupt memory of spoken information, also for words that have been correct 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.

REFERENCES

Beaman CP, Holt NJ (2007). Reverberant auditory environments: the effects of multiple echoes on distraction by irrelevant speech. Appl Cogn Psychol 21: 1077-1090.

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

Hällgren M, Larsby B, Lyxell B, Arlinger S (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. J Am Acad Audiol 12: 357-370.

Hagerman B (1982). Sentences for testing speech intelligibility in noise. Scand Audiol 11: 79-87.

Helfer SK, Wilber AL (1990). Hearing loss, ageing, and speech perception in reverberation and noise. J Speech Hear Res 33: 149-155.

Hodgson M, Nosal E (2002). Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms. J Acoust Soc Am 111: 931-939.

Jones DM, Macken WJ (1993). Irrelevant tones produce an ”irrelevant speech effect”: Implications for phonological coding in working memory. J Exp Psychol: LMC 19: 1-13.

Kjellberg A, Ljung R, Hallman D (2008). Recall of words heard in noise (p n/a). Appl Cogn Psychol (in press). DOI: 10.1002/acp.1422.

Magnusson L (1995). Reliable clinical determination of speech recognition scores using Swedish PB words in speech-weighted noise. Scand Audiol 24: 217-223.

Perham N, Banbury S, Jones DM (2007). Do realistic reverberation levels reduce auditory distraction? Appl Cogn Psychol 21: 839-847.

Pichora-Fuller MK (2003). Cognitive aging and auditory information processing. Int J Audiol 42 (Suppl 2): S26-S32.

Pichora-Fuller MK, Schneider BA, Daneman M (1995). How young and old adults listen to and remember speech in noise. J Acoust Soc Am 97: 593–608.

Seidel J, Weber L, Leistner P, Laschczok S (2005). Acoustic properties in German class rooms and their effect on the cognitive performance of primary school pupils. Paper presented at the Forum Acusticum, Budapest.