Does distortion of digitally coded speech decrease recall of spoken words?

Louisa Wruck
2014

Bachelor of Science
Engineering Physics and Electrical Engineering

Luleå University of Technology
Department of Computer science, Electrical and Space engineering
Bachelor Thesis
Does distortion of digitally coded speech decrease recall of spoken words?

Louisa Wruck

2013

Supervised by
Arne Nykänen (Luleå Technical University)
Sebastian Möller (Technische Universität Berlin)
Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Berlin, den 10.07.2013

Louisa Wruck
Does distortion of digitally coded speech decrease recall of spoken words?

Contents

1. Introduction .................................................. 5
   1.1. Working Memory ...................................... 7
   1.2. Speech and Intelligibility .......................... 9
       1.2.1. Intelligibility ................................ 9
       1.2.2. Long reverberation times and noise impairing recall of words 9
   1.3. Digital Signals ...................................... 12
       1.3.1. General System ............................... 12
       1.3.2. Codecs and distortion ....................... 13
       1.3.3. Quality assessment of mobile phones and the role of intelligibility 14
   1.4. Small introduction to statistics .................... 15

2. Experiment .................................................. 16
   2.1. Participants .......................................... 17
   2.2. Apparatus ............................................ 17
   2.3. Speech Signals ....................................... 17
       2.3.1. Choice and Creation of Signals ............. 18
   2.4. Method ............................................... 20
   2.5. Evaluation .......................................... 21
   2.6. Results .............................................. 21

3. Discussion .................................................. 23

4. Conclusion .................................................. 26

5. Summary ..................................................... 27
   5.1. Zusammenfassung ................................... 27
   5.2. Sammanfattning ...................................... 27

A. Word lists ................................................. 31

B. Programming ................................................. 31
   B.1. Python Script ...................................... 31
   B.2. Matlab Code ........................................ 35

C. List of Materials ........................................... 36

D. Equations to complete statistics .......................... 36
Abstract

This study is examining if digital distortions of GSM coded telephone signals impair one's capability of remembering speech. Earlier studies have shown an impairing effect on memory of spoken words when being presented in noise or long reverberation. Digital distortions, e.g., packet loss, echoes or jitter are assumed to have similar effects on working memory. Digital distortions are often caused by packet loss. In this experiment, effects on working memory due to packet loss in GSM-coded speech were investigated and compared to a noise condition. Word lists with 10 words in each list were presented via headphones. To ensure that the subjects perceived all words correctly the subjects were asked to repeat all words directly after listening to them. After all ten words had been presented they were asked to write down as many words as they could remember. Four conditions were compared: one undistorted signal, one signal distorted by noise (signal to noise ratio S/N 4 dB), one signal which was distorted by light and one with severe packet loss. The order of conditions was randomized.

In consistence with previous studies, an impairing effect of noise on the participants' ability to recall the spoken words was shown, while packet loss did not affect recall of spoken words.

It is suggested that further studies with other kinds of distortions should be carried out. Future studies should also include recall of spoken lectures or stories, as it is assumed that different memory processes are used under such conditions.
1. Introduction

One cannot deny that digital telephony, i.e. mobile phones or internet telephony are the most commonly used communication systems in today’s speech communication. Speaking with friends or business partners on a mobile phone or via internet telephony is not the exception but the rule. Moreover, lectures being streamed via internet is a technical tool commonly used by universities.

Our relations and cooperations are internationally spread so that internet telephony is becoming immensely important in economy, businesses and networking. Quality standards as well as low costs are essential demands of customers and many services already try to meet these requirements. Nevertheless, distortions are often present during telephony and sometimes problems understanding your interlocutor occur. If the understanding of spoken words or phrases is already challenging, does the extra effort of comprehension limit our abilities to remember the presented information?

Distortions originate from errors in signal transmission resulting in various kinds of degradations. In GSM, digitally coded speech is organized in packets, which are transmitted separately. Due to faulty packet transmission, effects, i.e. delays, echoes, jittering, and missing syllables or words, interrupt communication. Moreover, background noise is often introduced during packet transmission causing a signal of minor quality in comparison with the original.

Noise and long reverberation times in rooms are acoustical phenomena which have similar sound characteristics as some digitally distorted signals mentioned above. Several studies concerning noise and reverberation times affecting working memory have been carried out. Nevertheless, distorted digital signals, and how they occur in digital telephony, have not been investigated thoroughly.

Therefore, this thesis is a first step in examining the effects of digital coding of speech on working memory. The topic is complex and this thesis was limited to the examination of how packet loss affects recall of spoken information. The following research question was raised: Does distortion of digitally coded speech decrease recall of spoken words?

The research question aims to prove an impairing effect of distortions of digitally coded speech on recall of spoken words. The study was made using free recall of spoken words. Contextual prose, serial recall, the relation to long term memory etc. are not investigated. The digital distortions used in this experiment were based on distortions due to packet loss in GSM. Words with distorted time frames are still intelligible. However, other kinds of distortions, e.g. echoes, jitter etc. are not being considered and could reveal different results.
If it is shown that digital distortions are affecting working memory, many services based on digital coding should be investigated or thought through in detail. Streamed lectures, conferences held online or emergency calls could be comparably less effective than other forms of communication.

From an economical point of view, the impacts on working memory by digital distortions could help to judge which technologies should be further developed to improve telephony and quality assessments.
1.1. Working Memory

In psychology, numerous different models have been developed to explain several cognitive tasks, i.e. learning, reflexes, memory, emotions and other processes. One widely accepted model concerning memory divides the entire process into two parts, i.e. short-term/working memory and long-term memory. Stimuli perceived by the eye, ear or any other perceptual organ are further processed in working memory before a conscious task can be executed. In the following section, the process of working memory will be explained briefly.

After perception of a stimulus, it can be used in complex processes, i.e. comprehension, learning or reasoning. Regardless, the information needs to be first stored in working memory where it is temporarily stored and filtered so that only relevant information is further processed. Working memory is often seen as a primary but limited part of a conscious process [1].

Baddeley’s cognitive model is widely accepted in current psychology. The first model of 1974 [2] splits working memory in three parts; the visuospatial sketchpad, the phonological loop, also referred to as slave systems which are controlled by the third part - the central executive. The latter is again controlled attentively.

The visuospatial sketchpad and the phonological loop have similar functions, i.e. to combine sensory inputs and information [1]. While the visuospatial sketchpad is responsible for visual information, the phonological loop is in charge of auditory stimuli including the perception and production of speech. Therefore, vocabulary learning, for example, is directly connected to the phonological loop. The episodic buffer differs from the slave systems in its ability to store and retrieve information of the latter in the form of conscious awareness, of reflecting on that information and, where necessary, manipulating and modifying it [1].

After inconsistencies occurred in Baddeley’s 1974 model, he extended his model in 2002 with the episodic buffer [1]. The episodic buffer, like the phonological loop and the visuospatial sketchpad, is a limited capacity controlled by the central executive. It is capable of holding slave systems’ information in episodic frames, which can then be recalled by the central executive. The central executive is able to access the episodic buffer through conscious awareness and can influence the content stored by paying special attention to the source, which can be perceptual, a slave system or long term memory [1].

However, for this thesis, it is also of interest how many words we are able to remember. Miller [3] investigated this matter of interest in 1956 and came to the following result: We are able to remember seven plus/minus two chunks. Whereas chunks refer to blocks of information, e.g. a letter, a word, a sentence etc., with other words, initial information is grouped and recoded to increase performance, e.g. various letters are stored as a word so that in total more information can be retrieved. Miller compared the limitation of working memory with a bottleneck,
Does distortion of digitally coded speech decrease recall of spoken words?

so that only a small amount of information can pass through. Still, we are able to "widen" this bottleneck by regrouping or recoding information into chunks.

Miller’s model [3] of chunking can explain Baddeley’s [1] following observation: Subjects, when asked to remember a sequence of words, start to make mistakes if the number of words exceed five or six. In meaningful sentences more than 16 words can be remembered.

Coming back to Baddeley’s model of working memory as well as to the limited capacity of recall of words, there are several known effects connected to the phonological loop and therefore working memory. Lists consisting of long or similar words are more difficult to recall. These effects are known respectively as the word length effect and the similarity effect. Moreover, recall of words is inhibited if words cannot be repeated. If this is the case, the word length effect does no longer applies. Auditory information is blotted out quickly unless being repeated [1], [4].

The word length effect and others lead Baddeley to doubt Miller’s magic number seven. Nevertheless, he concludes that chunking explains many processes of learning and memory [5].
1.2. Speech and Intelligibility

Speaking produces sound waves which are transmitted via the air to an interlocutor. Depending on the surrounding, sound waves are manipulated introducing noise, echos or other acoustical effects.

When communication is taking place via telephony, sound waves are not transmitted through the air directly, rather they are encoded on the speaker’s side and decoded by the phone on the listener’s side. It is possible that information from signals can be lost during transmission. The latter can be compared to poor acoustical surroundings, since the effects created by digital distortion can be similar. Noise and room acoustics have been more thoroughly investigated and an impairing effect has been shown.

1.2.1. Intelligibility

In acoustics speech intelligibility is a measure which describes the ability of the acoustic environment to transmit speech intelligibility, usually expressed relative to perfect listening conditions [6]. Syllable intelligibility is one kind of intelligibility and often used in room acoustics; it is defined as the percent of incoherent syllables a listener understood, whereas more than 50% define good and more then 70% a very good intelligibility, see [7]. However, intelligibility can be measured by the percentage of correctly understood words.

When looking at noisy conditions, a signal to noise ratio (S/N) of 0 dB is needed to ensure good speech intelligibility, see [8].

1.2.2. Long reverberation times and noise impairing recall of words

In the past years, several studies concerning noise, long reverberation times in rooms and masking effects of the latter on working memory have been carried out. In terms of free recall of spoken words, an impairing effect of the above mentioned effects has been shown.

Noise is one of the most commonly studied distortions. Rabbitt [9] performed a series of experiments investigating the effect of noise on immediate memory, serial recall and recall of prose. The remembering of list of digits was inhibited when played with noise. Moreover, early list positions were more affected than recent list positions and the first part of prose was generally not remembered as well as the second part. However, when the second part of a prose text was played with noise and the first part as a clear signal, the clear signal was not remembered as well as when both parts were played as a clear signal. He suggested that special efforts need to be made to understand signals presented with a high S/N; extra
Does distortion of digitally coded speech decrease recall of spoken words?

capacity in working memory would be needed, replacing words the participants had already listened to.

However, he criticized his experiment, stating that it was not clear whether the participants could not remember the data or if the data was inaudible due to the high noise level.

Another study carried out by R. Ljung [4] examines the effects of noise on intelligible word lists and prose. In his experiments, he used word lists consisting of 50 one-syllable words. Effects on the first and last words of the lists were shown, but he suggested that a floor effect occurred in the middle part of the word lists. Since intermediate-term and short-term memory were impaired, Ljung suggested that more resources would be needed for word understanding so that less storage capacity or time was left for encoding and processing, which is consistent with Rabbitt’s findings [9].

A. M. Suprenant [10] analysed the effect of different S/N, S/N 10 dB and 5 dB, respectively. Even though more than 90% of the spoken nonsense syllables were intelligible in all conditions, a decline in performance in the noise conditions was recorded. A comparison of the different S/N results reveals that, in general, subjects performed equally on early lists positions. At recent list positions the S/N 5 dB affected recall of syllables greater than the condition with a S/N 10 dB. Suprenant interpreted her findings in terms of working memory as not corresponding to any of the current models. Her findings and reasoning are mostly consistent with Ljung’s experiment, so that they suggest the same reasons for a declining working memory performance. Moreover, she stated that intelligibility is not always a reliable factor when evaluating noise and distortions. Even though a signal could be intelligible, it might still be processed and remembered differently:

Regardless of the final theoretical interpretation, these data show that intelligibility tests provide a crude and sometimes inaccurate index of the extent to which noise may limit the efficiency of the encoding of information. The effect of noise cannot be disregarded until a point is reached at which recognition errors occur. Even levels of noise that do not have measurable effects on intelligibility may cause measurable decrements in the ability of listeners to remember spoken discourse. Noise may, in effect, impose an additional, ”secondary task” that must be carried out whenever speech has to be understood. [10], page 332.

**Reverberation Time** In room acoustics reverberation time describes the time which the sound level of an initial sound source decreases by 60 dB, see [11]. Describing reverberation time in nontechnical terms one could say it is how long a sound is still present in a room. The main sound characteristics in rooms are often described by this measure.

R. Ljung, see [12] and [8], investigated if long reverberation times impair working memory. In several experiments, non-contextual words and short lectures showed
that, when presented in long reverberation times, working memory was affected. At the same time all signals were intelligible. He suggested that the long delays in speech caused an additional effort in speech perception so that less cognitive resources were left for memorization. Similar interpretations are used to explain declining performance of working memory for recall of words presented with noise. Long reverberation times did not affect serial recall but rather free recall of spoken words. Moreover, the impairing effect of words on early list positions was greater than in recent list positions, see [4]. However, long reverberation times had a greater impact on recall of spoken words than on sentences, see [12]. R. Ljung’s experiment analyzing long reverberation times on lectures revealed an impairing effect on working memory, underlining his earlier findings. He criticized today’s standard of building acoustics which do not consider the effects of long reverberation times and noise on working memory. So far, speech intelligibility, usually measured with the objective speech transmission index (STI), is one of the most commonly used measures. Nevertheless, appropriate reverberation times and S/N in rooms would ensure a good learning environment.

Another study investigated the effects of reverberation time on distraction. P. Beaman and N. Holt [13] studied if reverberant offices supported distraction of irrelevant speech, investigating effects on serial recall. A graph showing the list position versus the number of participants remembering the word was used for explanations. When comparing the curve progression of the reverberant and quiet conditions a general decline in performance was recorded, whereas the curve progressions themselves were similar. Nevertheless, a reverberation time difference of 0.7 s and 0.9 s was not detectable. These findings are generally consistent with the findings of R. Ljung.

According to R. Ljung [4], recall of spoken words was inhibited by long reverberation times and noise. Now, digitally distorted signals can have characteristics similar to the latter. An echo, missing words or delays in transmission might result in effects inhibiting working memory due to an additional effort which needs to happen to facilitate comprehension.
1.3. Digital Signals

Modern communication methods are no longer restricted to physical locations and the limitations of wired connections. Wireless applications are based on digital codecs, and are replacing wired and analog systems rapidly if they not have done so already. Since present communication methods are using digital coding for transmission, a brief overview of the current transmission systems and codecs relevant to the understanding of the thesis is provided below.

1.3.1. General System

A digital signal consists of distinct elements, transmitting the signal in combinations of the digits 1 and 0. Most communication devices follow a general series of gadgets when being transmitted from their source to their destination, see Figure 1.

![Transmission in mobile phones](image)

Figure 1: Transmission in mobile phones

Speech is transmitted by sound waves, forming an analog signal which must first be transformed into digital signals in order for communication devices, i.e. mobile phones, to be able to process and transmit them. An A/D (analog to digital) converter is used for this purpose. The original wave forms are approximated by discrete values of the original signals and a bandwidth of up to 4 kHz is created. These values are then transmitted via a codec, see 1.3.2, from the source to the destination where it is decoded and played by a loudspeaker, see [14].

When the signal is transmitted via the transmission channel from the input to the output transducer, noise, distortion and interference will be added to the initial signal, see Figure 1. This reduces the quality of the latter at its destination, and therefore a hard to hear or incomprehensible signal reaches the listener’s end.

In the experiment, conditions with noise and distortion will be analyzed. Distortion is created by faults in the transmission channel when "imperfect response of the system to the desired signal itself" [14] creates a perturbation of the original waveform.

In digital telephony, the signal is equalized to a bandwidth of 200-3400 Hz. To minimize the load of the transmission channel, the digital signal passes through an encoder, which compresses the signal a second time. Encoders applied in telephony
are optimized for voices, so that all necessary information is obtained. Nevertheless, coding of signals often results in a perceivable degradation of speech quality [15].

1.3.2. Codecs and distortion

As already mentioned in the previous paragraph, the discrete signals of the original signal can easily be transferred into a digital signal. Those signals are encoded into a binary code for easier transmission. Due to the above mentioned distortion (noise and interference), the original signal might differ from the heard signal at the destination.

A number of codecs have been developed for transmitting data; different limitations of the first codecs have been continuously improved to today’s standards. One of those codecs is the Adaptive Multiple Rate Codec (AMR) and the Adaptive Multiple Rate Codec Narrowband (AMR-NB) which is used in the experiment. AMR and AMR-NB adapts to the performance of the operating network and can be used in Groupe Special Mobile/Global System for Mobile Communication (GSM), Universal Mobile Telecommunications Services (UMTS), and Enhanced Data rates for GSM Evolution (EDGE). The above mentioned codecs are often used respectively in modern mobile communication or in digital telephony.

The signals are saved in speech frames which each consist of four packages. One speech frame represents 20 ms or, in other words, 160 samples when a sampling frequency of 8000 samples per second is applied.

However, not all systems work with four frames per packet and sometimes speech frames are transmitted. For fast transmission, a speech frame is only transmitted once; in bad transmission environments, some services transmit frames twice in two different packets to ensure correct transmission. A lost speech frame can be heard when the second packet arrives. This creates some delays, but at least all information arrives at the listener.

Voice encoding rates and error control are also included in many systems, providing a good voice quality at the destination. However, as we know from our daily life, errors and incomprehensible signals impairing communication occur.

When packet loss occurs, different kinds of digital distortions inhibit communication. Jitter occurs when packets arrive late. However, when packets are only partly lost, speech frames are missing. This can cause a loss of syllables or entire words. Packet loss occurs mostly due to network congestions.
1.3.3. Quality assessment of mobile phones and the role of intelligibility

Speech intelligibility is not only used in room acoustics, but is also a measure in speech transmission services. There exist different standards to evaluate the quality of transmission in telephony. Sebatian Möller [16] distinguishes in between older and recent quality assessment approaches of speech quality in telephones, which will be shortly discussed in the following paragraph.

Older studies on voice transmission quality found different measures, e.g. speech intelligibility or clarity, naturalness, loudness, sound coloration and the differentiation of background noise and system distortions. The demands of the communication systems changed with the times, and recent studies focus on other measurements, i.e. sound coloration, frequency content, directness, continuity, and noise are referred to as the mostly used attributes. Nevertheless, the voice transmission quality is not only influencing the subjective experienced quality, but also the ease of communication and the conversation effectiveness. Here, delays, interactivity of conversational partners and motivation, including the speed of the conversation, are more relevant. Moreover, the service efficiency and the so-called advantage of access has an impact on the approach towards the quality of the service, e.g. the advantage of availability can outrun the disadvantage of a worse speech quality when comparing mobile telephony (digital system) with a stationary telephone (analog system).

Still, how important is the role of intelligibility in quality assessments? Most network providers try to achieve a speech intelligibility of 100%. Nevertheless, the subjective impression of the quality might still be poor; therefore, new quality assessments need to be developed. Possible measurements could be an impairing effect on working memory or an increasing fatigue [17] during long conversations, however, they have not been thoroughly investigated.
1.4. Small introduction to statistics

In psychology, two kinds of experiment set-ups are widely used: the between and the within subject design. The latter is also known as the repeated measures design. Moreover, several experiment set-ups use or a mix of the latter. Whereas the between subject design compares different conditions for different groups, the within subject design compares different conditions for the same participants.

However, this experiment set-up as described in section 2, page 16, is using a repeated measures design. Therefore, a rough explanation of the evaluation method of this experiment setup follows.

For the comparison of two conditions, the means of two conditions are compared. If only two sets of data need to be evaluated, a t-test is used. A difference in between the two conditions is described as significant when it would occur in 95% of the observed cases. Obviously, there is still a 5% chance that a mistake in prediction occurs, which is referred to as a mistake of the first kind.

When comparing more than two conditions, an analysis of variance (ANOVA) is often used. The principle of t-tests is still applied, but a slightly different method of evaluation is used.

Many t-tests comparing each condition separately should not be carried out, because the chances of making an error of the first kind increase with every additional condition. This kind of decline of the significance level is referred to as familywise errors and can be seen in equation 1, where \( n \) refers to the number of conditions which are going to be compared.

\[
\text{familywise error} = 1 - (0.95)^n
\]  

(1)

In order to avoid familywise errors, the Bonferroni correction, is applied to adjust the confidence intervals so that the chance of making an error of the first kind remains at 5%. Therefore, see equation 2 where \( \alpha \) refers to the error rate of making mistakes of the first kind and \( k \) is the number of comparisons.

\[
\text{new p-value} = \frac{\text{confidence interval}}{\text{no. of comparisons}} = \frac{\alpha}{k}
\]  

(2)

In t-tests and ANOVA, the effect size (\( r \) or \( \omega^2 \)) is often also calculated. The two measures help to estimate how important an effect is. Even though the effect size \( r \) and \( \omega^2 \) represent the same, their calculation is different from each other and can be seen in appendix D.

For further information, see Andy Field’s book of statistics [18], which this section is based on.
Does distortion of digitally coded speech decrease recall of spoken words?

2. Experiment

Various kinds of digital distortions exist with different characteristics and subjective impressions for the listener. We experience these effects in our daily life when communicating via mobile phone or internet based telephony. Even though a lot of improvement in digital telephony have been accomplished in the past years; who has not experienced conversations where the interlocutor was not understandable or where the conversation needed to be interrupted due to bad listening conditions? It is obvious that those kinds of conversations are not in favor of communication, but what can be said about those held in good transmission conditions? When all words are understandable, a good speech intelligibility of almost 100% is reached. Nevertheless, this still includes minor distortions. In comparison with ideal listening conditions is this affecting our ability to remember? This experiment is a first attempt to examine this question in more detail.

This thesis is investigating the effects of digital distortions on working memory, or, more precisely, recall of spoken words. However, other experiments have used different approaches to investigate the effects of distortions or bandwidth limitations. In quality assessments in telephony, subjective measurements are often used; J. Antons [17] analyzed the effect of bandwidth limitations on fatigue. To be able to have an objective measurement, electroencephalography was used. When listening to a 20 minute presentation, participants were shown to be more tired than when no bandwidth limitations where applied. This was in line with the subjective impressions of the participants.

Studies on noise and reverberation times on working memory suggested that more cognitive resources need to be dedicated for comprehension so that less resources are left for memorization, see section 1.2.

Additionally, J. Antons’ findings show an impairing effect of bandwidth limitations on human performance in terms of working memory; studies on noise and long reverberation times indicate similar results. These conclusions suggest that it is worthwhile to investigate the effects of different digital distortions on working memory.

Before the experiments were realized, a pre-study with 16 participants was carried out: a reference signal was compared with three different signals representing analog and digital telephony. The experiment set-up was mostly identical to the one described below, including the same word lists and reference signal. The four conditions were the reference signal, a simulated analog signal (Bandwidth: 300-3400 Hz and a S/N 35 dB) and two kinds of digital signals with different bandwidths (12.2 kbit/s and 4.75 kbit/s) and light digital distortion with 1% packet loss. The packet loss was not distributed in a controlled way so that words were not necessarily affected by the 1% packet loss; this was the main point of critique which was changed in this experiment.
A randomization of the word lists was not possible due to software limitations, therefore, a controlled sequence of the word lists was used. However, the evaluation of the repeated measures did not reveal an effect of any of the conditions on recall of spoken words.

It was suggested that the analog signal and the digital distortions used were not severe enough to show an impairment on free recall of spoken words. In terms of bandwidth limitations, it is clear that the common bandwidths in analog and digital distortions do not have an impairing effect. Therefore, this study focuses on the effects of packet loss using noise as a controlling condition, since effects have already been shown in the past.

2.1. Participants

All 20 participants had self-reported normal hearing. Visual impairments were corrected by glasses or contact lenses and all participants were reported to be non-dyslexic.

The 12 female and 8 male participants were students or employees at the Luleå University of Technology and were native Swedish speakers within an age range of 18-35 (a mean age of 25.4) and received a movie ticket as compensation.

2.2. Apparatus

The experiments were carried out in a well tempered and quiet control room of a studio; no external noises distracted the participants.

All participants were seated in front of a screen, and the computer was placed outside the control room to prevent disturbances by noise. Headphones were worn during the entire experiment, with a duration of approximately 30 minutes. The tests were carried out in between 9 a.m. and 6 p.m.

A non-native Swedish speaker acted as an experiment leader and was present during the entire experiment to clarify questions. A list of the relevant material can be found in the appendix C.

2.3. Speech Signals

To determine if digital distortions impair recall of spoken words, four different kind of signals were created: word lists with a good quality (reference signal), with a low S/N (Noise) and two kind of digitally distorted word lists (light and severe distortion) were used. The reference signal was used to create the conditions with a low S/N and digital distortions.
2.3.1. Choice and Creation of Signals

The word lists were recorded by a male voice and consisted of short, contextless, one or two syllable words. In between the words, a pause of approximately 5 seconds was introduced.

In order to be able to compare the results of the randomized signals, different requirements needed to be met. Therefore, all words needed to be distorted in a randomized order to an equal extent. Signals needed to be distorted in a similar randomized but controlled way.

Reference Signal  The word lists were recorded with a S/N 30 dB. Originally, the word lists consisted of 40 phonetically balanced words, but for the experiment, set-up lists of 10 words were needed. Therefore the lists were divided using Adobe Audition into word lists, each being approximately 60 seconds long.

Noise  To be able to compare the experiment setup with Robert Ljung’s in [4] a S/N 4 dB was introduced. A speech spectrum, see Figure 2, was used to create the noise condition with the desired S/N. The original lists, used in the experiment as a reference signal, were dubbed with white noise in correspondence with the speech spectrum to create a signal close to real life situations. The sound level of white noise was adjusted in octave bands to ensure an equal distortion of all frequencies. A Matlab script, see appendix B.2, was used to create the sound files with the above mentioned characteristics.

Several studies showed that word lists with a high S/N inhibit working memory, see section 1.2.2 on page 9. Robert Ljung [4] has shown an impairing effect on recall of spoken words if presented with noise; therefore similar results are expected for this condition. If no significant effect on noise affecting recall of spoken words can be shown the experiment setup might include fatal errors leading to incorrect results.
Does distortion of digitally coded speech decrease recall of spoken words?

Digitally Distorted Signals  As already described in section 1.3.2 on page 13, most digital distortions are caused by packet loss. If severe packet loss occurs, syllables or entire words in a conversation can be missing. However, free recall of spoken word lists were used as a task in this experiment. Therefore every word of each list needed to be similarly distorted.

Speech transmission in AMR-NB is organized in packets which consist of four speech frames respectively, see section 1.3.2 on page 13. One speech frame represents 20 ms; with other words, if one packet is lost, four speech frames or 80 ms of a word or sentence would be lost. Nevertheless, AMR-NB adjusts the number of frames per package according to the quality of the transmission channel. For the experiment it is important that the artificially introduced packet loss does not only occur during pauses. The AMR-NB used in the experiment uses the method of fast transmission, where speech frames are not transmitted twice. Moreover, signals were limited to a bandwidth of 4.75 kbit/s. From the original records, see reference signal 2.3.1 on page 18, a S/N 30 dB was transferred.
Does distortion of digitally coded speech decrease recall of spoken words?

To be able to create a randomized distortion suitable for the purpose of the experiment, a packetization of one frame per packet was assumed; AMR-NB uses a packetization of one or more speech frames per packet [19]. Packages were lost in a randomized order but only two to three consecutive packets were lost at once so that it may look like the following example: 001001111000, where 0 are lost and 1 correctly received speech frame.

In order to have no effects due to the order of frame losses, the order of losses was, as already mentioned, randomized. In Figure 3, a visualization of the speech signals with the packet loss of the above example can be seen.

2.4. Method

All participants listened to 16 word lists, each consisting of ten words. Four lists were used for each condition - the reference signal, noise, light, and severe distortion. The way the signals were created is described in section 2.3.1. The order of the word lists and the conditions were randomized so that first the number of the word lists (1-16) were randomized and afterwards the kind of distortions were randomly assigned to the word lists. However, there were always four lists of the same condition.
Does distortion of digitally coded speech decrease recall of spoken words?

After listening to a word the participants were asked to repeat it aloud during a 5 second break. This way, correct hearing was ensured and intelligibility measured. When one word list was finished, the participants were asked to write down all remembered words in an input box. There was no time limitation in which the participants had to accomplish the test and they could decide when they wanted to continue with the next word list. An experiment leader was present in the room during the experiment to control the correct understanding of words and to answer questions.

2.5. Evaluation

Instructions, the playing of sound files, and saving of data was carried out with a self programmed python script, which can be reviewed in the appendix B.1. After the experiments were carried out the collected data was loaded and evaluated again using a python script. Correctly remembered words were scored with a one; incorrectly remembered words with a zero. The sum of scored points for each list was divided by the sum of the correctly understood words for the same list, and then written into a file which was loaded into the program for statistical evaluation SPSS. Repeated measures using the statistical method analysis of variance (ANOVA) for evaluation compares each condition for each participant separately. As already mentioned, the experiment set-up uses a within-subject design.

2.6. Results

The data was evaluated by a one-way ANOVA using repeated measures. Explanations on the methods and values can be found in section 1.4. In the following paragraphs, the evaluated data will be presented:

Maulchy’s test of Sphericity was not significant ($X^2(5) = 4.207; p = 0.521 > 0.05$) so that the assumption of sphericity was not violated. Therefore the degrees of freedom were not corrected $F(3, 57)= 3.85; p< 0.05; \omega^2 = 0.997$. This shows that the distortion, even if not stated which one, has an impact on the performance on free recall of spoken words.

To identify which kind of distortions had an impairing effect on recall of words, the pairwise comparison is considered. The relevant ANOVA output can be seen in table 1. SPSS automatically compares all conditions with each other using the Bonferroni correction to avoid familywise errors. However, in the experiment only the first condition (reference signal) is compared with the other three conditions (noise, light and severe conditions); no comparisons in between the other conditions, e.g. noise and light distortion, are carried out. Therefore the Bonferroni correction taking all possible pairwise comparisons into account is a too strong requirement for this evaluation.
Does distortion of digitally coded speech decrease recall of spoken words?

**SPSS Results Pairwise Comparison**

<table>
<thead>
<tr>
<th>factor1</th>
<th>factor1</th>
<th>Mean Difference</th>
<th>Std Error</th>
<th>Sig.</th>
<th>95% confidence intervall</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>J</td>
<td>I-J</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference Noise</td>
<td>0.64</td>
<td>0.25</td>
<td>2.61</td>
<td>-0.008</td>
<td>0.137</td>
</tr>
<tr>
<td>Reference Light Distortion</td>
<td>0.006</td>
<td>0.019</td>
<td>0.306</td>
<td>-0.051</td>
<td>0.063</td>
</tr>
<tr>
<td>Reference Severe Distortion</td>
<td>-0.010</td>
<td>0.028</td>
<td>-0.343</td>
<td>-0.092</td>
<td>0.073</td>
</tr>
</tbody>
</table>

*Table 1: Pairwise Comparison: Significance of noise, light and severe distortion in comparison with the reference signal*

**Results T-tests**

<table>
<thead>
<tr>
<th>Reference - Noise</th>
<th>M</th>
<th>SE</th>
<th>t(19)</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.803/0.738</td>
<td>0.029/0.031</td>
<td>2.61</td>
<td>0.017</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Reference - Light Distortion</td>
<td>0.803/0.797</td>
<td>0.029/0.134</td>
<td>0.306</td>
<td>0.763</td>
<td>0.07</td>
</tr>
<tr>
<td>Reference - Severe Distortion</td>
<td>0.803/0.812</td>
<td>0.029/0.162</td>
<td>-0.343</td>
<td>0.763</td>
<td>0.08</td>
</tr>
</tbody>
</table>

*Table 2: T-test: Significance of noise, light and severe distortion in comparison with the reference signal. A significant difference occurs when p<0.05*

Instead, only three pairwise comparisons were made, reference to noise, reference to light distortion and reference to severe distortion. Bonferroni-corrected t-tests showed the following results, which can also be seen in table 2.

When comparing the data of the reference signal (M=0.803, SE= 0.029) with the data of the noise condition, the high S/N affected recall of spoken words (M=0.738, SE= 0.031, t(19)=2.61, p <0.05, r=0.51). The t-tests which compare the reference signal (M=0.803, SE= 0.029) with the light (M=0.797, SE=0.134, t(19)=0.306, p>0.05, r=0.07) and severe distortion (M=0.812, SE=0.162, t(19)=-0.343, p>0.05, r=0.08) did not show a significant difference in the performance of recall of spoken word tasks. The formula used to calculate r and $\omega^2$ can be seen in the appendix D.

In summary, the t-tests showed a significant difference in between the reference - noise signal. Simultaneously the comparison of the reference signal with light or severe distorted word lists did not show a significant effect on recall of spoken words.

Intelligibility of the word lists were measured and can be seen in table 3. All lists for the different conditions had an intelligibility above 90%.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Noise</th>
<th>Light Distortion</th>
<th>Severe Distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligibility</td>
<td>99.5</td>
<td>95</td>
<td>95.25</td>
</tr>
</tbody>
</table>

*Table 3: Intelligibility in %, summarized for all lists of one condition*
3. Discussion

Does distortion of digitally coded speech decrease recall of spoken words? Intelligible words which are distorted by packet loss are not having an impacting effect on free recall of spoken words. Why did we not measure the expected results?

One reason could be the relatively low quality of the reference signal. However, looking at the reference condition, a S/N 30 dB is relatively high compared to a S/N -4.68 dB. This S/N was obtained by L. Wong, E. Ng and S. Soli [20] to describe the S/N when the Sentence Recognition Threshold (SRT) with a speech intelligibility of 50% was reached. Even though unrelated words were used in the experiment, the difference in between the reference and the noise condition from S/N 30 dB and S/N 4 dB should have been great enough, so that the relatively bad reference signal should not have affected the results. Furthermore, the S/N ratio of the reference signal and the digitally distorted signals were exactly the same so that only effects of packet loss were measured.

The experiment leader was not a Swedish native speaker, which might have affected the collection of data when measuring intelligibility. However, this should not have had a significant impact on the results. Moreover, incorrectly heard words were excluded from the analysis. Since this is mostly important for the distorted signals and would only have a positive influence, it should not change the results obtained.

A main difference in between the conditions of noise and digital distortion was that the pause in which the participants were asked to repeat the words heard aloud was once with noise (S/N 4 dB) and in the other conditions with silence (S/N 30 dB). In one of his experiments, P. Rabbitt [9] used lists of digits which the participants were supposed to remember. He varied the moment when noise was presented additionally to the speech signal so that in total he examined four conditions: two undistorted signals, first list undistorted second list distorted, first list distorted and second list undistorted, and both lists distorted. His findings revealed that when the second list was presented with noise, the first list was less well remembered than the condition which used the opposite order. P. Rabbitt suggests that masking effects, "overwriting" the digits which were supposed to be remembered, were responsible for this results. When participants in the experiment listened to the noise which was played during the pauses, this effect could have caused an impairing effect on memorization.

Recent experiments (Posner and Rossman, 1965) suggest that items are maintained in immediate memory by some activity, conveniently termed “rehearsal,” which requires part of the subject’s channel-capacity. Any activity intervening between the presentation of items to the subject and his subsequent reproduction of them is
assumed to occupy channel capacity, inhibiting rehearsal and reducing the probability that items will be correctly recalled. - P. Rabbitt [9]

However, R. Ljung [12] studied the effects of long reverberation times on working memory. In his experiment, the pause, in which the participants were asked to repeat the words heard aloud, were similar to the experiment conditions of light and severe digital distortion; namely without additional perturbation of any kind. At the same time, he used word lists with 50 words instead of 10. Long reverberation time (R. Ljung [12]) affected early list positions whereas noise (R. Ljung [21]) impaired early and recent ones. Since reverberation times only showed an effect on one part of the long word lists it is assumed that the task used in this experiment might show different results for longer word lists.

P. Bhargava and D. Baskent [22] studied if periodic interruptions and low-pass filtering of speech influences speech intelligibility. The highest low pass filter was applied at 3 kHz, which is close to the generally applied cut-off frequency of 3.4 kHz, in digital telephony. At this point, slow interruption rates of 1.5 Hz had an impact on speech intelligibility, whereas fast interruption rates, i.e. 10 Hz did not show an impairing effect. The periodical interruptions can be compared to the experimental condition of packet loss. Still, the difference of a non-periodical and periodical interruptions persists. However, P. Bhargava and D. Baskent used contextual dutch sentences investigating the effects on speech intelligibility, whereas this thesis studied unrelated words and impairments of working memory. Nevertheless, memory tasks considering contextual phrases might show different results. Moreover, their findings concerning speech intelligibility, as not being an ideal measure for effects on working memory, performance is consistent with the above mentioned results of R. Ljung’s researches.

When a bilingual person listens to speech in noise, the reception threshold differs for the first and second language. It is even worse for the first language when being compared with monolingual speakers, see [23]. Moreover A. Stuart, J. Zhang and S. Swink showed in their investigation that people speaking a foreign language are more vulnerable to noise conditions. So it would be possible that packet loss might effect this group of people more since a higher memory capacity would be needed for the understanding of words when being spoken in a foreign language.

At the same time, another group which might be more vulnerable to noise and long reverberation times are people with hearing impairments. Hearing aids are using digital codings, being an additional source of digital distortion. Would the combination of the two perturbations increase the demand of processing and thereby influence working memory?
Several studies, e.g. [24], have shown that age influences working memory so that speech processing slows down. Moreover, it has been shown that noise increases the listening effort of elderly people greater than for younger people, see [25]. Therefore, it is suggested to investigate if digital distortions are having an impairing effect on intelligibility of digital signals as well as memory of speech for bilingual, hearing impaired or elderly people.

Impairing effects, due to reverberations and noise, see section 1.2.2, have been proven; digital distortions can have similar effects, i.e. echos or jitter and background or signal correlated noise. The different acoustic effect could affect one’s ability to remember and recall speech.

Since effects of packet loss on working memory could not have been shown, it is assumed that none or only a small amount of cognitive load is necessary for understanding. Therefore, a high amount of cognitive load is available for further processing. Other digital distortions need to be further investigated, as they sound different and can therefore have other effects on working memory. Moreover, intelligibility does not seem to be able to represent effects on memory. The reference signal was more intelligible (99.5% intelligibility) than the signal distorted by "severe" distortion (90.38% intelligibility); the signal for noise and light distortions had approximately the same intelligibility (95% intelligibility). However, effects on working memory could not be shown for the digitally distorted conditions.
4. Conclusion

This thesis is a first attempt to investigate the influences of digital distortions on working memory. Experiments have been carried out using a free recall task with differently distorted signals. An undistorted signal and three distorted signals with noise and two digital distortions have been used. From previous studies it is known that noise and long reverberation times in rooms have a negative effect on working memory. Since digital distortions can have similar characteristics as the above mentioned conditions, similar effects were expected.

Nevertheless, with packet loss distorted digital signals used in this experiment did not reveal the expected results; no impairing effect could be shown. At the same time, background noise inhibited the participants capability of memorization.

Even though, packet loss does not influence working memory, other kinds of digital distortions have not been investigated. Jitter, echoes or other errors may have different sound spectra compared to the ones of packet loss. It is possible that the different kinds of distortions have different effects on listeners. Moreover, only free recall has been investigated. Tasks using contextual questions, could show different results, since contextual texts are differently processed in working memory. This thesis is a first step in investigating the effects of digital distortions; the experiments used are hold simplistic. Therefore, it is suggested that further studies are carried out, which investigate the effects thoroughly and are not only scratching on the surface. Different kind of distortions, combinations of distortions, and more complex tasks should be considered.
5. Summary

5.1. Zusammenfassung


5.2. Sammanfattning

I denna studie undersöktes om digitala störningar av GSM-kodade telefonsignaler påverkar minne av tal.


I överensstämmelse med tidigare undersökningar visades en försämring av försökspersonernas förmåga att komma ihåg tal när signalen stördes med brus, medan paketförlusterna inte påverkade försökspersonernas minne.

Förslag på fortsatta undersökningarna är andra sorts distorsioner och längre berättelser eftersom man kan anta att olika processer i arbetsminne påverkas beroende på typ av störning och typ av uppgift.
List of Tables

1. Pairwise Comparison: Significance of noise, light and severe distortion in comparison with the reference signal ........................................... 22
2. T-test: Significance of noise, light and severe distortion in comparison with the reference signal. A significant difference occurs when \( p<0.05 \) .............................................................................. 22
3. Intelligibility in \( \% \), summarized for all lists of one condition ...... 22
4. Word lists .................................................................................... 31
5. Materials used in the experiment. .................................................. 36

List of Figures

1. Transmission in mobile phones ...................................................... 12
2. Speech Spectrum ............................................................................ 19
3. Speech signal when a pattern 001001111000 packet loss occurs. .. 20
4. First instructions and information for the experiment .................... 32
5. To make sure that the participants had no hearing impairments etc. people with impairments were asked to quit the test. .................. 32
6. Instructions about the task of the participants. ................................. 33
7. When pressing the button "Spela ordlista" the word list was played and no words could be entered in the entry boxes. Once the word list finished, participants could write down all the words they remembered. 33
8. Collection of personal information was carried out after the experiment was finished. ................................................................. 34
9. The experiment leader marked the correctly understood words using the check boxes. ............................................................... 34
References


[16] Sebastian Möller. *Quality Engineering Qualitätsgestaltung kommunikation-
Does distortion of digitally coded speech decrease recall of spoken words?


A. Word lists

<table>
<thead>
<tr>
<th>1</th>
<th>vas</th>
<th>fisk</th>
<th>frukt</th>
<th>hål</th>
<th>nerv</th>
<th>kalv</th>
<th>näbb</th>
<th>snår</th>
<th>blad</th>
<th>hiss</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>sill</td>
<td>rad</td>
<td>blixt</td>
<td>sol</td>
<td>plats</td>
<td>torsk</td>
<td>dans</td>
<td>stork</td>
<td>kök</td>
<td>spalt</td>
</tr>
<tr>
<td>3</td>
<td>doft</td>
<td>skärp</td>
<td>lik</td>
<td>grabb</td>
<td>stund</td>
<td>park</td>
<td>slips</td>
<td>bly</td>
<td>mjölk</td>
<td>frö</td>
</tr>
<tr>
<td>4</td>
<td>puss</td>
<td>skydd</td>
<td>grav</td>
<td>damm</td>
<td>dräng</td>
<td>mur</td>
<td>glass</td>
<td>kust</td>
<td>dvärg</td>
<td>sits</td>
</tr>
<tr>
<td>5</td>
<td>eld</td>
<td>gift</td>
<td>köp</td>
<td>slant</td>
<td>skämt</td>
<td>lån</td>
<td>chef</td>
<td>bär</td>
<td>häl</td>
<td>pil</td>
</tr>
<tr>
<td>6</td>
<td>släkt</td>
<td>ring</td>
<td>spis</td>
<td>sko</td>
<td>fru</td>
<td>kam</td>
<td>kniv</td>
<td>haj</td>
<td>brev</td>
<td>skur</td>
</tr>
<tr>
<td>7</td>
<td>stänk</td>
<td>kött</td>
<td>fjäll</td>
<td>spår</td>
<td>blink</td>
<td>tand</td>
<td>död</td>
<td>knä</td>
<td>tjur</td>
<td>skratt</td>
</tr>
<tr>
<td>8</td>
<td>fat</td>
<td>folk</td>
<td>glas</td>
<td>hatt</td>
<td>barr</td>
<td>broms</td>
<td>rum</td>
<td>knapp</td>
<td>svan</td>
<td>bänk</td>
</tr>
<tr>
<td>9</td>
<td>storm</td>
<td>djur</td>
<td>burk</td>
<td>fluga</td>
<td>zoo</td>
<td>yta</td>
<td>hud</td>
<td>kjol</td>
<td>vete</td>
<td>groda</td>
</tr>
<tr>
<td>10</td>
<td>yxa</td>
<td>vinge</td>
<td>palm</td>
<td>byxa</td>
<td>kyrka</td>
<td>mus</td>
<td>bäck</td>
<td>lax</td>
<td>post</td>
<td>nymf</td>
</tr>
<tr>
<td>11</td>
<td>regn</td>
<td>bonde</td>
<td>gräs</td>
<td>skrik</td>
<td>löv</td>
<td>stuga</td>
<td>ben</td>
<td>skrift</td>
<td>land</td>
<td>ägg</td>
</tr>
<tr>
<td>12</td>
<td>virus</td>
<td>matta</td>
<td>eka</td>
<td>blus</td>
<td>polis</td>
<td>dass</td>
<td>hals</td>
<td>mössa</td>
<td>fält</td>
<td>liv</td>
</tr>
<tr>
<td>13</td>
<td>kung</td>
<td>gran</td>
<td>mynt</td>
<td>vante</td>
<td>natt</td>
<td>dam</td>
<td>bild</td>
<td>tumme</td>
<td>gevär</td>
<td>kropp</td>
</tr>
<tr>
<td>14</td>
<td>strand</td>
<td>frack</td>
<td>tak</td>
<td>duk</td>
<td>plåt</td>
<td>sax</td>
<td>gös</td>
<td>stock</td>
<td>tall</td>
<td>hjul</td>
</tr>
<tr>
<td>15</td>
<td>keps</td>
<td>päron</td>
<td>berg</td>
<td>vev</td>
<td>dröm</td>
<td>häst</td>
<td>sked</td>
<td>bil</td>
<td>hugg</td>
<td>svala</td>
</tr>
<tr>
<td>16</td>
<td>puls</td>
<td>klass</td>
<td>ask</td>
<td>lön</td>
<td>lärka</td>
<td>pedal</td>
<td>hall</td>
<td>skåp</td>
<td>kliv</td>
<td>vik</td>
</tr>
</tbody>
</table>

Table 4: Word lists

B. Programming

B.1. Python Script

The software used in the experiments to present the signals to the participants and to collect and evaluate data was done using a python script. Pictures of the output of the programs used for collecting the data can be seen below.

Software requirements: The python scripts need a Python version of Python 2.7 or higher. If Python 3.3 is used the module "Tkinter" might need to be installed separately. However, the external library "Pyglet" needs to be installed individually. With python versions higher than Python 2.7 the installation of Pyglet might fail. In this case it is recommended to use "easy install" via the command window (Windows), terminal (Mac). The programmes should be compatible with Mac Os X 10.5.8, Windows XP, and Windows 7. Linux should be compatible as well but has not been tested.
Does distortion of digitally coded speech decrease recall of spoken words?

Figure 4: First instructions and information for the experiment

Figure 5: To make sure that the participants had no hearing impairments etc. people with impairments were asked to quit the test.
Does distortion of digitally coded speech decrease recall of spoken words?

Figure 6: Instructions about the task of the participants.

Figure 7: When pressing the button "Spela ordlista" the word list was played and no words could be entered in the entry boxes. Once the word list finished, participants could write down all the words they remembered.
Does distortion of digitally coded speech decrease recall of spoken words?

Figure 8: Collection of personal information was carried out after the experiment was finished.

Figure 9: The experiment leader marked the correctly understood words using the check boxes.
B.2. Matlab Code

Matlab Code for the generation of signals with noise impairment.

```matlab
fs = 44100;
p0 = 2e-5;
LAtarget = 56;

%white noise 10 s
wn = randn(10*fs,1);

%speech spectrum noise

%bandpass filter
[B,A] = butter(1,[125 500]./(fs/2));

spn = filter(B,A,wn);

%A-weight filter
h = fdesign.audioweighting('WT,Class','A',1,fs);
Ha = design(h,'ansis142');

LA = 10*log10(var(filter(Ha,spn))./p0^2);
gain = 10.^((LAtarget-LA)/20);
spn = gain.*spn;

fprintf(1,'Level verification spn level = %0.1f
',
       10*log10(var(filter(Ha,spn))./p0^2));

% 1/3 octave filter levels
%1/3-octave filter
f = fdesign.octave(3,'Class 1','N,F0',6,1000,fs);

%base 10 center frequencies 25-20k Hz
fc = validfrequencies(f);

L = zeros(size(fc));

for k = 1:length(fc)
    f.F0 = fc(k);
    H = design(f,'butter');
    L(k) = 10*log10(var(filter(H,spn))./(2e-5)^2);
end

plot(L);
set(gca,'xtick',5:3:28,'xticklabel',round(fc(5:3:28)),'xlim',[5 28])
xlabel('Frequencies (Hz)');
ylabel('Level (dB)');
```
C. List of Materials

The speech signals have been recorded in an anechoic chamber at Luleå Technical University (LTU) and have been taken from B. Hagerman’s list for testing speech intelligibility in noise [26]. The original lists were phonologically balanced. Since the lists were divided so that ten words per list were always played, the lists used in the experiment were not phonetically balanced anymore but all lists together were. However, the same recordings, but undivided, have been used in the experiment of R. Ljung [12].

The following table 5 shows the hardware used in the experiment.

<table>
<thead>
<tr>
<th>Headphones</th>
<th>Head Acoustics HPS IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound Card</td>
<td>M-audio Delta 1010lt sound card</td>
</tr>
</tbody>
</table>

Table 5: Materials used in the experiment.

D. Equations to complete statistics

In this section, the statistical calculations shown in the results of the experiment 2.6 can be seen. All formulas were extracted from Andy Field’s book on statistics [18].

The measure $\omega^2$ used for the evaluation in the analysis of variance (ANOVA) can be seen below.

\[
\omega^2 = \frac{SS_M - (df_M)MS_R}{SS_T + MS_R}
\]
\[
= \frac{SS_T + MS_R}{49.6 - 3(0.022)}
\]
\[
= \frac{0.067 + 49.6 + 0.022}{0.997}
\]

The effect size $r$ for t-tests is calculated as follows. In the equations 7 to 12 the r values for the different comparisons of means can be seen.

\[
r = \sqrt{\frac{t^2}{t^2 + df}}
\]

Reference-Noise:

\[
r = \sqrt{\frac{2.61^2}{(2.61)^2 + 19}}
\]
\[
= 0.51
\]
Does distortion of digitally coded speech decrease recall of spoken words?

Reference - Light Distortion:

\[
\begin{align*}
 r &= \sqrt{\frac{0.306^2}{0.306^2 + 19}} \\
 &= 0.07 \\
\end{align*}
\]  

(9) \hspace{1cm} (10)

Reference - Severe Distortion:

\[
\begin{align*}
 r &= \sqrt{\frac{0.343^2}{0.343^2 + 19}} \\
 &= 0.08 \\
\end{align*}
\]  

(11) \hspace{1cm} (12)