Effects of background speech on reading performance in adults: Results using a new test procedure

Helga Sukowski\(^{(a)}\), Erik Romanus\(^{(b)}\)

Federal Institute for Occupational Safety and Health (BAuA), Dortmund, Germany
\(^{(a)}\) Sukowski.Helga@baua.bund.de
\(^{(b)}\) Romanus.Erik@baua.bund.de

Abstract

Among the various factors that determine the working environment noise is often rated as stressful by many employees. This holds not only for especially noisy workplaces, with a potential of hearing damage, but also for many other settings with lower noise levels like, e.g., offices, hospitals or schools. It is known that cognitive performance can be affected by background noise, even at moderate levels. This has frequently been reported for effects of spoken language on working memory. There is less experience with the effects of moderate noise on other cognitive tasks. As reading is an essential requirement at work, a new reading task was developed and applied to investigate effects of noise. The development of the procedure is based on results from previous studies on noise effects on reading carried out by the first author. The procedure is designed as a computer-based task for adult workers with normal reading ability. The participants have to find and mark mistakes in written sentences, under a moderate time pressure. In a pilot study 12 participants worked on this task twice, once in a silence condition and once during the presentation of background speech at a moderate level. Even the data from this small group revealed that the test procedure is of high practicability. With respect to the two different experimental conditions, it was found that there were significantly more correctly finished items and less reported effort in the silence condition than under noise. The results indicate that the new procedure may have the potential to serve as an instrument for the quantification of effects of noise or specific noise characteristics on one important aspect of cognitive performance at work. Further measurements with more participants and further sound conditions will be carried out in near future.

Keywords: effects of noise, reading performance, occupational
1 Introduction

There are many factors that determine the environment at a particular workplace, for example: room size, space at workplace, furnishings and equipment, climate, illumination and also the acoustical conditions. Noise is often rated as stressful by many employees, among the factors constituting the work environment, and in some professions noise is often named as a reason for calling consultation. This does not only hold for work places with especially high noise levels, like in certain industries. This even holds for work places with typically lower noise levels, e.g. noise in hospitals, schools, or offices.

It is well known that noise at high levels causes hearing damage. Therefore, occupational safety and health regulations exist stating that employers have to protect their workers from this possible impairment, by noise reduction, instruction / training and finally by using personal protective equipment. The prevention of noise induced hearing loss as a result of long term exposure to high sound levels is still an important concern in the frame of the efforts in occupational safety and health. In addition to this concern the question whether and to what amount noise at high and also at lower levels may also cause other detrimental effects on employees, like effects on the cardiovascular system, performance or mental stress gained more and more interest during the last years.

This contribution deals with the effects of noise on one aspect of cognitive performance represented by effects on reading. For many years it has been well investigated and documented in numerous studies that spoken language as an unwanted background sound affects working memory [e.g. 1; 2]. Working memory is of course involved in many cognitive tasks, but the tasks typically used to test the working memory ("serial recall", i.e. repeating numbers or words in the same sequence as presented) are to some extent artificial tasks that are only rarely required in this isolated form in working life.

Since the main interest was on possible effects on employees and typical tasks at the workplace reading was chosen for this study as one representative task essential for many workplaces. This of course holds for office jobs but this also holds for jobs in many other settings, where for example, instructions on how to operate a machine have to be read accurately in order to work safely or a report in a hospital setting has to be read correctly to choose the right medication for a patient.

Reading ability has different components, essentially reading fluency, reading accuracy, reading comprehension. Therefore reading performance in general can be measured in different ways and described along different parameters. Studies investigating effects of noise on reading have existed for many years, but the results whether noise disturbs reading, or, more specific, to what amount which characteristic of a sound disturbs which component of the reading performance is not clear. A comparison between different studies or the results, respectively, is often difficult. On one hand the comparison is difficult, because different principles are used to measure the performance (different tasks that focus on different components), and on the other hand the comparison is difficult, because of differences in the applied noise conditions and test settings.

Therefore it is no surprise that over the years varying results concerning noise effects have been published. For example: Weinstein [3] revealed a detrimental effect of moderate noise of a
teletype machine on contextual errors in a proof reading task while the noncontextual errors were not affected; in one of their experiments Martin et al. [4] demonstrated that a task on reading comprehension was more affected by a comprehensible background speech than by an incoherent background speech, random noise or in a silent condition; Schwabe [5] did not find any detrimental effect of traffic noise on a 3 hours lasting proof reading task that merely contained typing errors; in a more recent study Haka et al. [6] did not find a significant effect on performance in a proofreading task conducted in three acoustical background conditions with different speech transmission indices.

Because reading is such an essential part of everyday work, and effects of noise on reading even in safety critical situations might have considerable consequences, it seems to be worthwhile to deal with this topic in a more systematic way in future. The pilot study presented here is one step on this way.

The aims of the pilot study were:

1) To test whether the in-house developed reading task is of practical usability for studies on noise effects on reading on employees.
2) To get a first impression, whether there are differences in the reading accuracy between a silent condition and a condition with background speech.

2 Method

2.1 Reading test procedure

The preliminary design of the procedure has two different test versions (A and B). Both versions consist of 52 items each. The items themselves are made up by one or two sentences of different length and content. The task is computer based, i.e. participants read the sentences on a computer screen. The items are either completely correct or they contain one mistake. The task for the participants is to read each item separately and to decide whether the item is correct or not. If the item contains a mistake, participants need to mark the word that is not correct or that does not fit into the context of the sentence. If the item contains no mistake, participants should mark a field with the label “everything correct” (original: alles richtig).

After a decision for either a mistake or the field “everything correct” participants click on a field labelled “finished” (original: fertig). The mistakes implemented in this procedure are of different quality. There are e.g. orthographic mistakes, grammatical mistakes but also words that are not reasonable in the given context. The participants have to work on the sentences with a moderate time pressure. They know that the sentences disappear after a predefined time. Based on several pre-tests the presentation time was set in a way to make the task challenging.

Finally, with this procedure that can be titled as a proof reading task, information can be gathered about reading accuracy (number of correctly finished items), working speed (time for editing the items) and in part, about reading comprehension. The latter aspect only holds for
those items with mistakes, where a comprehension of relevant parts of the sentence is necessary in order to realize that one word does not fit with the rest of the statement.

The idea for this procedure is based on prior experience with a procedure the first author had developed for a study on noise effects in children several years ago. The original procedure was shown to be sensitive to reveal effects on reading accuracy between different road traffic sounds [7; 8; 9]. Later the procedure was also applied in a study with adults, using considerably harder conditions concerning the time pressure and a speech noise condition. Again the procedure could uncover detrimental effects of noise on reading performance [10].

Based on these experiences a new, separate procedure for adults was now constructed as a next step. The main differences between the procedure for children and the new one for adults are: The contents of the sentences are rather adjusted to everyday life situations of adults, many of them with reference to working life; orthographic mistakes have been inserted; the procedure is computer-based now. The computer program used to implement the described application is an in-house development, especially created for this study. In the current form the procedure is - concerning difficulty and time pressure - only appropriate for participants with normal reading ability.

2.2 Effort

After each test run the participants were asked to assess the effort they experienced while they carried out the proof reading task. On the screen they got a scale designed like a ruler from 0 (no effort at all) to 100 (very much effort). The personal judgment was marked as a stroke on this ruler by a mouse click.

2.3 Participants and study design

In this contribution data from the first part of a pilot study are presented. Twelve volunteers participated (age: from 26 to 62 years; male: 6; female: 6). All participants were staff members of the German Federal Institute for Occupational Safety and Health (BAuA). The participants differed in their particular jobs in the institution and the amount of prior experience with scientific studies. None of the participants knew the sentences and mistakes in advance. All volunteers carried out the test once in a silent condition and once in a condition with background speech. The sequence of the acoustical test conditions was balanced as well as the sequence of the test versions. In most cases the two runs were carried out with an interval of one or two days between the different conditions.

2.4 Acoustical conditions

As the speech sound condition several sections from an audio drama were presented stereophonically via closed circumaural headphones (Sennheiser HD 380 pro). The mean sound pressure level in the right ear was 62 dB(A) and in the left ear 57 dB(A). This slight level difference was part of the setting of the audio drama. The recording contained scenes with different speakers sometimes in different rooms. The sound was played back with a notebook, the same computer that was also used to run the experimental task. Participants wore the
headphones also in the silent condition, in order to give some shielding from unwanted sound events, but mainly in order to have a largely comparable situation as during the sound condition.

2.5 Room
The pilot study was carried out in a quiet single office room during typical daytime office working hours. All unwanted potential sound disturbances were precluded before starting the reading test. There was no disruption by unwanted, non-experimental sounds during the measurements.

3 Results

3.1 Range of the number of correctly finished items
Items were rated as finished correctly when they were completely edited correctly. That means the mistake was correctly marked for the sentences with a mistake and the decision was finalised with the "finished-button" in time. For sentences with no mistake the "everything-correct-button" had to be marked and this decision also had to be finalised with the “finished-button”. The range of items finished correctly in the silent condition was from 27 to 45 items ($\Delta$ 18). The range in the speech sound condition was from 19 to 43 items ($\Delta$ 24).

3.2 Effect of acoustical condition
A paired $t$-test revealed that the number of correctly finished items was significantly lower in the speech sound condition than in the silent condition ($t(11) = 4.509$, $p = 0.001$; Mean ($M$)$_{Silence} = 36.17$; $M_{Speech} = 31.33$). In this contribution differences are called “significant” when $p < 0.05$. The correlation coefficient (Pearson) between both conditions was $r = 0.88$.

3.3 Differences between test version A und test version B
The mean values for correctly finished items in versions A and B were generated by a calculation over all participants and both acoustical conditions. The mean value for version A was 34.17 items and for version B 33.33 items. This small difference did not show up as statistically significant in a paired $t$-test ($t(11) = 0.465$, $p = 0.651$).

3.4 Effort
A paired $t$-test on the assessments of the effort during editing the items resulted in a statistically significant difference between both acoustical conditions. The ratings were significantly lower in the silent condition than in the speech sound condition ($t(11) = -5.003$, $p < 0.01$; $M_{Silence} = 42.17$; $M_{Speech} = 69.83$).
Figure 1: Number of correctly finished items in both acoustical conditions “Silence” and “Speech”, maximum number: 52 items, * = the difference is significant

Figure 2: Rated effort during the both acoustical conditions “Silence” and “Speech”, * = the difference is significant
4 Summary and discussion

4.1 General findings

The design of the new reading task turned out to be suitable for the application in studies on noise effects on reading with normal reading participants. All volunteers got the instructions of the task quickly, and during a short conversation after the test all test persons confirmed that the handling was easy. Thus, the procedure seems to be of high practicability. The range of the results shows that the procedure all in all was an adequate challenge for the tested group. There was a large variety in the results, but none of the attendants failed completely and nobody managed to find all mistakes in the given time. That means the procedure can cover a broad variety of reading accuracy. Most participants mentioned that there were a few (two or three) items which they would have liked to read twice, but the time given was over before doing so. Of course, there will be an in-depth analysis of these items, and if necessary the presentation time can be adapted, but in principle it is intended that the items are of diverse difficulty and that there are some items that will only be edited correctly by very few participants.

With respect to the first aim stated in the introduction, it can be concluded that the new reading task is of practical usability for studies on noise effects on reading in adults.

Referring to the second aim it can be concluded that there is a significant difference in reading accuracy between both acoustical conditions. Participants were less accurate in editing the items in the speech sound condition than in the silent condition. Even in this small group there is a clear result that participants are disturbed by the speech sound while carrying out a proof reading task. Therefore, it can be stated that a proof reading task in the way as designed here is sensitive to effects by speech noise.

Several explanations are possible why the noise effect on a proof reading task becomes apparent in this study. One reason might be derived from time restriction for each single item (time pressure) in combination with the uninterrupted test run. Both aspects together leave no chance to compensate for inaccuracy or a lack of attention in a noisy condition. Another reason why the proof reading was affected, e.g. in contrast to the results from Schwabe [5], might be the fact that intentionally the sentences contained a great variety of errors, also contextual errors, that are more likely to be affected by noise than errors that do not depend on the context [3]. In addition, due to the variety of errors in the current task, participants could not prepare for specific errors when they started to read an item. A third reason why the effect becomes obvious already in a small group might be the fact that with the speech sound a sound condition was chosen that has a high probability to disturb a cognitive task.

All in all the first steps with the new procedure were successful in the sense that the task with diverse implemented errors was sensitive to an unwanted background speech sound.

4.2 Specific aspects in experimental studies with repeated measurements

Currently a second part of the study with 12 participants is running where six persons carry out the task twice in a silent condition and six twice in the speech sound condition. The aim of this part is to test whether there are training effects between two runs in the same condition. The
results from both parts of the pilot study should be used also to carry out analyses on further aspects in order to identify necessary modifications and optimisations.

First, there should be analyses with respect to the question whether there are specific types of errors that are especially affected by the speech sound condition, and if so, whether the observed effect is mainly caused by a few specific items. We then aim for a deeper insight into the difficulty of single items as well as the parallelism of the two test versions.

A high level of parallelism and definite information about training effects are important requirements for any procedure that is to be carried out twice by one participant in a short time period. Therefore we will not only keep an eye on these issues in the pilot study but also in prospected larger studies.

At the first glance it might seem like going too much into detail when dealing with these aspects. However, such aspects are of great relevance for the interpretation of results from studies on noise effects on humans. If there is a considerable training effect from one run to the next the training effect might mask the noise effect. Such a result or the use of a procedure of low ecological validity might lead to an interpretation that a particular sound did not harm the performance in a particular test and, if such results recur time and again, in the worst case this interpretation can hamper investments and dedications in preventive and protective measures.

Due to the commitment of our institute in the field of safety and health at work and our mandate to develop regulatory proposals to protect employees from adverse health effects it is necessary to keep an eye on the question how statements about effects of noise or the lack of observable effects at the workplace are gained.

4.3 Final remarks

As mentioned above, there are many factors that constitute the workplace environment and the entire work situation, and of course, reading is just one task that has to be carried out in a work-related setting. However, we will stick on the reading task in future, because reading is an essential task in the daily work of many employees at different workplaces. Therefore, the reading task is meant to be a representative task that employees have to carry out with many different purposes.

The task itself is still somewhat artificial, but it is nevertheless closer to reality than a serial recall task for numbers. The task contains many elements that are very realistic. For many employees it is a typical situation that they have only a short moment to get a message from an e-mail or a report from the screen or that proof reading has to be done under time pressure.

Since the first steps were successful it is planned to refine the procedure to a straightforwardly applicable research instrument that might serve to gather detailed information about possible disturbance by different noise characteristics that are typical for work-related background sounds.
Acknowledgments

The authors would like to thank the colleagues from the German Federal Institute for Occupational Safety and Health (BAuA) for their participation in this pilot study and to Ulrich Hold for programming the computer-based procedure.

References


