Communicating information in noise: speaker's and listener's perspective

Annelies BOCKSTAEL[1]; Laurie SAMYN[2]; Paul CORTHALS[2]; Dick BOTTELDOOREN[1]

[1] Ghent University, INTEC-WAVES, Belgium

[2] Ghent University, Speech, Language and Hearing Sciences, Belgium

ABSTRACT

Most environmental studies show that background noise has a negative effect on cognitive performance. The effect of background noise can be two-folded when both listener and speaker are exposed to the same noise. Then not only the perception, but also the quality of the given information might be degraded. This study evaluates the cognitive load for speaker and listener while presenting/retaining information in three types of realistic background noise, city street sounds, multitalker babble and highway noise. For the speaker, cognitive load is quantified by the number of dysfluencies. For the listener, performance on a secondary visual task is evaluated.

For the listener, fluctuating city street sounds have a clearly adverse effect on the visual task. Within the set of presented background sounds, this fragment has the most salient events, potentially attracting the listener's attention. For the speaker, the effect of background noise is less pronounced. This might be due to the specific parameter chosen to evaluate cognitive load, but it is also plausible that the lack of auditory tasks imposed on the speaker makes it more easy to ignore the background noise.

Keywords: Speech intelligibility and interference, task performance, cognitive load I-INCE Classification of Subjects Number(s): 63.3, 63.5

1. INTRODUCTION

Background noise is known to influence task performance. Often, extra mental effort is needed to focus on the task and not on the interfering noise (1).

The effect of background noise depends on the specific interaction of background noise characteristics, task requirements and individual traits. For the background noise, the saliency is especially important, this is how likely it is that the background noise will attract the attention. Saliency not only depends on the overall level of the noise, but also on the temporal structure; sounds with more fluctuation and variation are thought to be more salient (2). Another important aspect is the possible information present in the background noise; interfering speech could be more disturbing, than meaningless background noise (3).

When tasks involve orally presented information, e.g. a typical teaching situation, both sender and receiver have to deal with the background noise. Hence, the noise not only affects the processing of the information by the listener (4), but also the quality of the information given by the talker. In noisy class rooms for example, the quality of the given information might be compromised because the teacher has to spend cognitive resources on strategies to deal with the background noise, leaving less mental capacity for presenting information (5).

The fact that background noise influences listener *and* speaker might hence imply a stronger effect on information processing and task performance. In this experiment, the effect of background noise on cognitive resources is assessed for both speaker and listener.

For the listener, listening effort is often measured following the dual task paradigm (6). Here, the listener has to complete a secondary task together with the main task. The more cognitive resources

1 Annelies.Bockstael@intec.ugent.be

that are (still) available, the better the performance on this secondary task and vise versa. For orally presented information, the main task would be the processing of information and the secondary task can be a visual task.

For the speaker, cognitive load can be quantified by the number of dysfluencies. Research has shown that the percentage of dysfluent speech increases as cognitive load increases; the number of breaks and interjections has been found to increase (7), as well as their duration (7). In addition, the number of incomplete sentences, relaunches and repetitions increases (7).

This study evaluates the cognitive load for speaker and listener while presenting/retaining information in three types of realistic background noise, i.e. fluctuating city street sounds, multitalker babble and monotonous highway noise. For the listener, cognitive load is measured by the performance on a secondary visual task and for the speaker the percentage of dysfluent speech is used.

2. MATERIAL AND METHODS

2.1. Background noise

For the background noise, the following recordings were played back; (1) fluctuating city street noise, (2) monotonous highway traffic noise, and (3) unintelligible multitalker babble. A fourth (4) condition without additional background noise has been added as reference condition.

The background noise fragments were played at an average level of 70 dB SPL, with peaks of 80 dB SPL. At the position of the participants, levels varied between 65 dB SPL and 75 dB SPL.

2.2. Orally presented information

To mimic a teaching situation, the speaker had to read Dutch text fragments on either scientific topics or Dutch literature. Reading of one fragment took between 3 and 5 minutes. The information presented in the fragments was quite specialized but presented in layman's terms. This to ensure that the listener had little a-priori knowledge about the topics, and were capable to understand the presented information.

For the speaker, the text was presented in Calibri font, font size 11 with 1.15 line spacing on A4 paper format. To increase readability, italic and bold fonts were included.

Both speaker and listener were instructed to focus maximally on the content of the lectures. To stimulate this, they had to complete a small exam afterwards.

2.3. Secondary task listener

The secondary task was a visual task: a screen with 14 white squares was shown to the listener. Out of those 14 squares, 6 random squares were shortly colored green one after the other. After each sequence of 6 green squares, the participant was asked to indicate on a screen with only white squares which squares had become green. Both accuracy and reaction time were stored for each trial. During each lecture, at most five instances of the visual task were presented to the listener at random moments.

2.4. Dysfluencies speaker

To score the dysfluencies of the speaker afterwards, speech was recored with a Neumann KM 88 I microphone. With the software Praat, the percentage of dysfluent speech in each speech sample was calculated. The following dysfluencies were included: (1) repetitions, (2) corrections of wrongly read phonemes, syllables, words, part of sentences, and intonation patterns, (3) abnormal long or unusual pauses, (4) interjections, (5) words or word groups read in a hesitating way, and (6) subtle prolonging of fricatives and nasals.

2.5. Test setup

For each test, the two participant, one speaker, one listener, were seated facing each other at 1.20 m distance. The microphone to record the speech was placed on the table using a tripod at 30 cm from the speaker's mouth. Both speaker and listener had their own computer screen respectively showing the text of the lecture and the visual task. The screens did not block the participants' view. Background noise was played by two loudspeakers at 2 m from the participants.

For the first four fragments, one participant was assigned speaker and the other listener. The speaker was instructed to present the information clearly and at the same time retain as much of the information as possible. The listener was told to focus on the information and execute at the same time the visual task without loosing focus on the primary listening task. No information was given on the background noise. After four lectures, the participants switched roles for the next four fragments.

At the end, an exam on the content of all lectures was completed individually by each participant. For each lecture, participants were also asked to rate their interest in the subject on a 5-point scale.

2.6. Participants

To be included in this experiment, participant had to be fluent in Dutch. No hearing loss or severe issues with vision were allowed, as well as problems with speech and language or attention-related disorders. Participants were asked to have a good night sleep prior to the experiment.

The volunteers were recruited via convenience sampling. Duos who already knew each other before were preferred to facilitate the speaker-listener interaction.

In total 60 people participated, 42 female and 18 male. They were on average 21.4 years old, the youngest participants being 18 years old and the oldest 32. All signed the informed consent approved by the ethical committee before testing.

2.7. Statistical analysis

Mixed model linear regression was used to analyze the results. In all the models, the variable participant was included as random factor. The cognitive load of the listener was assessed by including either the score on the visual task, either the reaction time as dependent variable. For the speaker, the percentage of dysfluencies was used as dependent variable. When appropriate, pairwise Tukey post-hoc comparisons were carried out.

3. RESULTS

3.1. Cognitive load listener

Recall that 5 instances of the visual task were presented during each lecture. Each instance consists of 6 green squares the participant had to locate correctly.

For every lecture, the 5^{th} test instance is removed because depending on the speed of the speaker there was not always time enough to present the 5^{th} instance. Therefore, to ensure an equal amount of instances of the visual task for all lectures, only the first four instances are included. In addition, the visual tasks done during the first lecture are also removed for all participants. This is done because a strong learning effect on the performance of the visual task is seen during this first lecture.

The test scores are quite high and little variation is seen, for 84.8 % of all observations, participants place either 5 or 6 of the 6 squares correctly. The lack of variation makes the variable score less suitable to assess difference in performance over different test conditions.

However, the reaction time needed to complete the visual task does vary. Moreover, the reaction time appears to be not correlated with the score. Hence, it is hypothesized that a longer reaction time implies greater cognitive effort to obtain the (mostly almost maximal) score.

Statistical analyses first of all reveal that the following variables had no statistically significant influence on the reaction time: participant gender, participant interest in the lecture, and the topic of the lecture.

Statistical analyses do reveal that the reaction time on the visual task changed significantly over the sequence of the lectures ($p<0.01$). Tukey post-hoc analyses reveal that especially during the lecture presented last the reaction time is lower than during the 2^{nd} ($p=0.04$) and 3th lecture ($p=0.01$). Recall that the first lecture had already been omitted because of the observed learning effect.

The strongest effect on the reaction time is seen for the variable background noise ($p<0.0001$). Pairwise Tukey post-hoc testing shows that reaction time is especially longer for conditions with fluctuating city street noise, especially compared to the condition without additional background noise ($p<0.001$), but also compared to multitalker babble ($p=0.02$) and highway noise ($p=0.03$).

3.2. Cognitive load speaker

For the cognitive load of the speaker, analyses are similar to those for the listener, only this time the percentage of dysfluent speech is used as dependent variable.

In general, results are statistically less pronounced for the speaker compared to the previous analyses of the listeners' cognitive load. Again, no statistical significant effect is seen for participants' gender and for the topic of the lecture.

Descriptive analyses and visual inspection of the data suggest that dysfluencies slightly increase over subsequent lectures, and also that a higher percentage of dysfluency is noted when participants state to be not interested in the topic of the lecture. However, no statistical significant effect is found for either variables ($p>0.05$).

Still, a very moderate effect of background noise is observed ($p<0.05$), with somewhat less dysfluencies in the condition with multitalker babble compared to the condition without background noise (Tukey post-hoc comparison $p=0.05$).

4. DISCUSSION

The most pronounced result from this experiment is the longer reaction time of the listener for visual tasks performed in city street noise, suggesting a higher cognitive load. This type of noise is indeed very salient with a relatively low background level and clearly audible events such as car passages, birds starting to sing, car honks, and so on. It is quite likely that these events trigger the listener's attention away from the tasks and hence increase reaction time.

It could have been hypothesized that multitalker babble would also be disturbing, as it is more similar to the target signal, i.e. the speaker's reading, and hence might require more effort to ignore compared to traffic noise. However, the multitalker babble varies little over time as it consists of a mixture of different voices and unintelligible speech.

For the speaker, the effect of background noise is much weaker. This might be due to the parameter chosen to measure cognitive load, namely the percentage of dysfluent speech in the whole speech fragment. It might be worthwhile to investigate also the variation in duration of dysfluencies and/or the number of dysfluencies. In addition, other parameters assessing vocal effort could also be analyzed.

Apart from the parameters to address speaker's cognitive load, it is intuitively quite plausible that background noise has a greater effect on the listener than on the speaker. The listener has to separate two auditory streams, i.e. the relevant signal and the irrelevant speech. By contrast, no auditory tasks were required from the speaker, so he/she could 'simply' suppress all auditory input.

Finally, thus far the results for speakers and listeners have been analyzed separately. In future, it would be very interesting to see whether there is a statistically significant relationship between the quality of the presented information on the one hand and the performance of the listener on the other.

5. CONCLUSIONS

In this work a clear effect of background noise on listener's performance is found. Especially salient background noise with clearly audible events (city street noise) increases the reaction time needed to complete a visual task successfully. For the speaker, the effect of background noise appears to be less pronounced. This might be related to the sensitivity of the specific parameter used to quantify cognitive load, but could also be explained by the fact that no explicit auditory task was given to the listener, potentially making it more easy to ignore the background noise.

ACKNOWLEDGEMENTS

Annelies Bockstael is post-doctoral fellow of the Research Foundation Flanders (FWO). The support of this organization is gratefully acknowledged.

REFERENCES

1. Szalma JL, Hancock PA. Noise effects on human performance: a meta-analytic synthesis. Psychological bulletin. 2011 Jul;137(4):682.

2. De Coensel B, Botteldooren D, Berglund B, Nilsson ME. A computational model for auditory saliency of environmental sound. The Journal of the Acoustical Society of America. 2009 Apr 1;125(4):2528-.
3. Barker BA, Mosely C. The influence of multiple narrators on adults' listening comprehension. InProceedings of Meetings on Acoustics 2013 Jun 2 (Vol. 19, No. 1, p. 060193). Acoustical Society of America.
4. Schneider BA, Daneman M, Murphy DR, See SK. Listening to discourse in distracting settings: the effects of aging. Psychology and aging. 2000 Mar;15(1):110.
5. Enmarker I. The effects of meaningful irrelevant speech and road traffic noise on teachers' attention, episodic and semantic memory. Scandinavian Journal of Psychology. 2004 Nov 1;45(5):393-405.
6. Fraser S, Gagné JP, Alepins M, Dubois P. Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. Journal of Speech, Language, and Hearing Research. 2010 Feb 1;53(1):18-33.
7. Berthold A, Jameson A. Interpreting symptoms of cognitive load in speech input. Springer Vienna; 1999.