The effects of noise reduction on cognitive effort in normal-hearing and hearing-impaired listeners

Anastasios Sarampalis[1], Sridhar Kalluri[2], Brent W. Edwards[2], and Ervin R. Hafter[1]

[1] *University of California at Berkeley, Department of Psychology, 3210 Tolman Hall, Berkeley, CA 94702, USA*

[2] *Starkey Hearing Research Center, 2150 Shattuck Ave, Berkeley, CA 94704, USA*

A common complaint of hearing-impaired listeners is difficulty understanding speech in the presence of noise. Digital hearing aids have opened the door to complex signal processing algorithms that attempt to improve the quality, ease of listening, and/or intelligibility of speech in noisy environments. In reality, however, hearing aid users show no intelligibility improvements from single-microphone noise reduction (NR) algorithms, even though they sometimes report that speech sounds easier to understand. A possible explanation for this dichotomy is that NR algorithms replace a function that the human auditory system would otherwise perform. This redundancy means that there is no improvement in intelligibility, but a reduction in listening effort, since fewer cognitive resources would be necessary. We investigated this hypothesis using a dual-task paradigm with normal-hearing and hearing-impa¬ired listeners. They were asked to repeat sentences or words presented in noise while performing either a memory or a reaction-time task. Our results showed that degrading speech by reducing the signal-to-noise ratio increased demand for cognitive resources, demonstrated as a drop in performance in the cognitive task. Use of a NR algorithm mitigated some of the deleterious effects of noise by reducing cognitive effort and improving performance in the competing task.

INTRODUCTION

One of the greatest challenges hearing impaired (HI) listeners face is understanding speech in the presence of noise. Digital hearing aids have allowed the development of signal processing strategies that attempt to address this problem. Single-microphone noise reduction (NR) algorithms are designed to reduce the effects of noise on speech intelligibility by adjusting the gain according to the assumed signal-to-noise ratio (SNR) over time. Though NR algorithms exist in a variety of forms in the hearing-aid industry, there have been few reports of speech intelligibility benefits. In most cases, HI listeners perform as well or better without the NR, than they do with it. They sometimes, however, show preference towards using the NR, because of perceived improvements in sound quality, improved clarity or ease of understanding.

We propose that digital NR algorithms perform an analysis of the auditory scene that is similar in nature to the processing the human auditory system performs when enough cognitive resources are available. As such, in a typical laboratory experiment, NR results in no speech intelligibility improvements because it adds little information to what is already available. Based on theories of attention (Broadbent, 1958) and channel capac-

Anastasios Sarampalis, Sridhar Kalluri, Brent W. Edwards and Ervin R. Hafter

ity (Kahneman, 1973) we use a dual-task paradigm to investigate this hypothesis. In such a paradigm, the primary task is a speech intelligibility task, where the listeners repeat sentences or words presented to them over headphones. A second, simultaneous, task requires participants to either try to remember some items or to respond to visual stimuli while their reaction times are measured. Changes in the amount of effort required to perform one of the tasks (for example, understanding speech) will manifest as changes in performance in the other tasks (for example changes in reaction times).

The present study aimed to use this objective measure of listening effort to test the hypotheses that:

- The presence and amount of background noise affects how much effort is required to understand speech.

- Digital NR algorithms reduce the effort of listening to speech in the presence of noise.

EXPERIMENT I: EFFECTS OF NOISE REDUCTION ON RECALL OF SPOKEN WORDS

Listeners

Twenty-five native American English speakers with thresholds lower than 15 dB HL at all audiometric frequencies participated in Expt I. Their average age was 20 years.

Stimuli

The Revised SPIN sentences were used in Expt 1 (Bilger *et al.*, 1984). In each condition, 48 sentences were played to the listener over headphones (Sennheiser HD580). Half of them contained context information making the last word predictable, while the other half did not. They were presented at a level of 65 dB SPL, either in quiet, or in the presence of 4-speaker babble. When babble was present, the SNR was -2 or 2 dB and the materials were either left unprocessed or were processed using the Ephraim-Malah NR algorithm (Ephraim and Malah, 1984, 1985).

Procedure

The paradigm used in Expt 1 is based on that of (Pichora-Fuller *et al.*, 1995). Listeners were tested in a double-walled sound-attenuating chamber. They were instructed to repeat the last word of each sentence, as they believed they heard it. The next sentence was presented one second after their response. They were also asked to remember their responses as they would be asked to recall them later. After every 8 sentences, a visual cue prompted the listener to recall as many of the previously reported words as they could, verbally and in any order they preferred.

Results

Average speech-intelligibility scores were calculated for each condition and are presented in Fig. 1. As expected, performance was best for sentences with the higher SNR and for sentences with context. Moreover, for the sentences without context information,

the processed materials were harder to understand than the unprocessed materials.

Average memory-performance scores are presented in Fig. 2. Performance was best when the sentences were presented in quiet (dashed reference lines). For no-context sentences, the addition of noise (2dB SNR) dropped memory performance by 7-10%. The 4-dB reduction in SNR decreased performance further, by 5-8%. For high-context sentences, the addition of noise (2dB SNR) dropped performance by 10-12%. For unprocessed sentences there was a further drop in performance when the SNR was reduced by 4 dB, however memory performance for the processed sentences did not drop with the decrease in SNR.
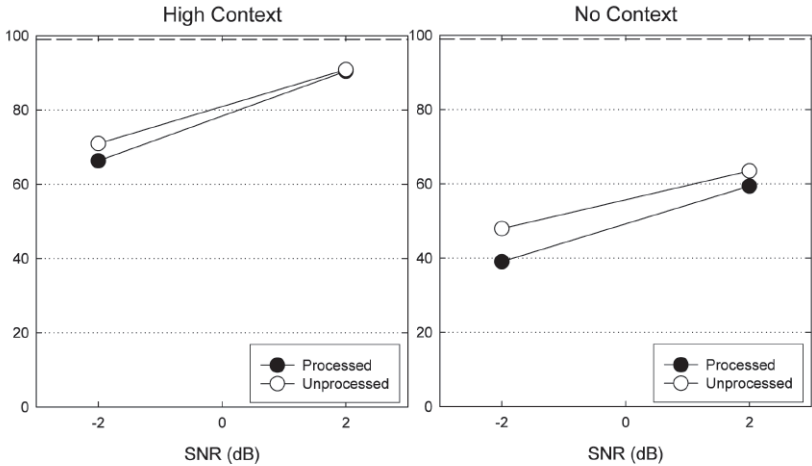

Fig 1: Average speech intelligibility performance for 25 NH listeners as a function of SNR, NR processing, and context. The dashed reference line shows performance in quiet.
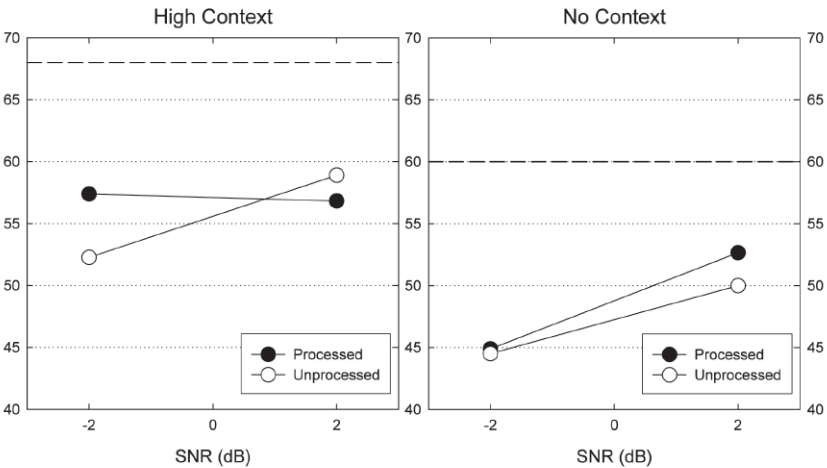

Fig. 2: Average recall performance for 25 NH listeners as a function of SNR, NR processing, and context. The dashed reference line shows performance in quiet.

Figure 3 shows average memory performance data, re-plotted as a function of word position in the sentence list (1 to 8). The memory curves for no-context sentences show clear recency (better performance for recent items) and primacy (better performance for early items) effects for sentences presented in quiet, but only recency effects for all sentences presented in noise. For high-context sentences, the recency effect is seen in all five conditions. The primacy effect is strong for sentences presented in quiet, but becomes smaller with the addition and increase of noise. The primacy effect for the -2dB SNR condition is larger with NR than with the unprocessed sentences, suggesting that the NR helped the rehearsal processes involved in short-term memory tasks.
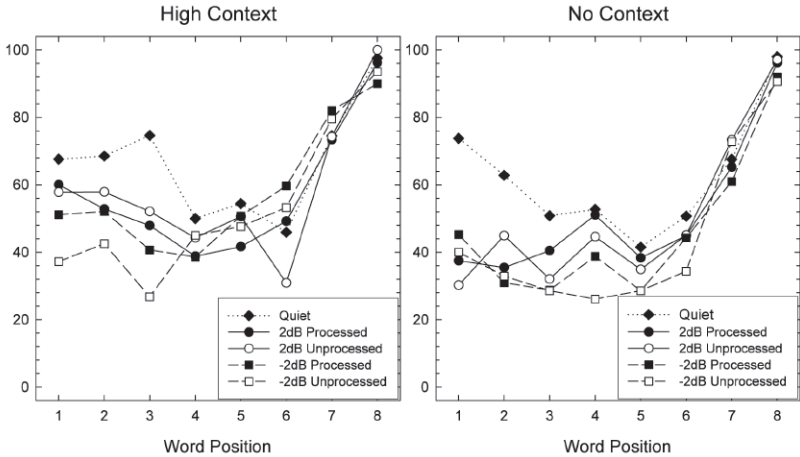

Fig. 3: Average recall performance for 25 NH listeners as a function of word position, SNR, NR processing, and context.

EXPERIMENT II: EFFECTS OF NOISE REDUCTION ON REACTION TIMES TO VISUAL STIMULI

Listeners

Twenty-five native American English speakers with audiometric thresholds lower than 15 dB HL and six listeners with hearing impairment took part in this experiment. The HI listeners were divided in two groups according to their speech intelligibility performance in noise (based on pilot data). It was also the case that the three listeners in Group 1 were all approximately 60 years of age and all listeners in Group 2 were approximately 80 years of age. Their audiograms are shown in Fig. 4. The average age of the NH listeners was 21 years.

Stimuli

The IEEE (IEEE, 1969) sentences recorded in a male voice were used in Exp. 2. With NH listeners, the sentences were presented at 65 dB SPL, either in quiet or in the presence of 4-speaker babble with SNRs of -6, -2, and 2 dB. When in noise, the speech was either left unprocessed or processed with the Ephraim-Malah NR algorithm (Ephraim and Malah, 1984,1985). With HI listeners, the target sentences were presented at 65

dB HL and the noise, when present, was set at -2, 0, and 2 dB for Group 1, whereas they were set at 2, 4, and 6 dB for Group 2. These were chosen to cover similar speech intelligibility performance ranges. Fifty sentences were played over headphones (Sennheiser HD580), in each experimental condition.
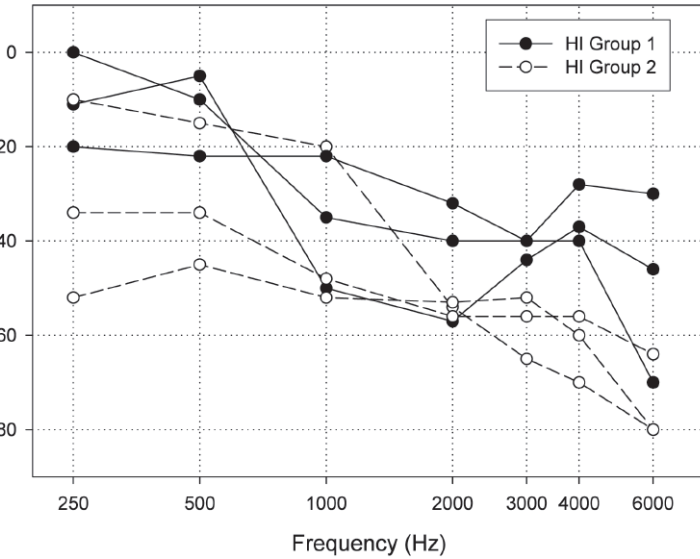

Fig. 4: Audiometric thresholds for the six HI listeners who participated in Experiment 2.

Procedure

Listeners were tested in a double-walled, sound-attenuating booth. They were instructed to repeat each presented sentence, as they thought they heard it. Concurrently to the listening task, they were instructed to perform a visual reaction-time task. Specifically, they were shown a window with two boxes on a monitor. At quasi-random intervals a digit appeared in one of the two boxes. The participant's task was to press the button with an arrow pointing to the direction of the digit, if the digit was an even digit, or the button with an arrow pointing away from the digit, if the digit was an odd digit. They were told to respond as quickly and as accurately as possible. Accuracy and response times were recorded.

Results

Figure 5 shows average speech intelligibility results as a function of SNR, processing condition, and listener group. As in Expt 1, speech intelligibility performance with NH listeners dropped as the SNR was reduced (panel a). The processing had very little effect on performance. Similar results were observed with the second group of HI listeners (panel c). On the other hand, Group 1 performed significantly better with the unprocessed than the processed sentences (panel b).
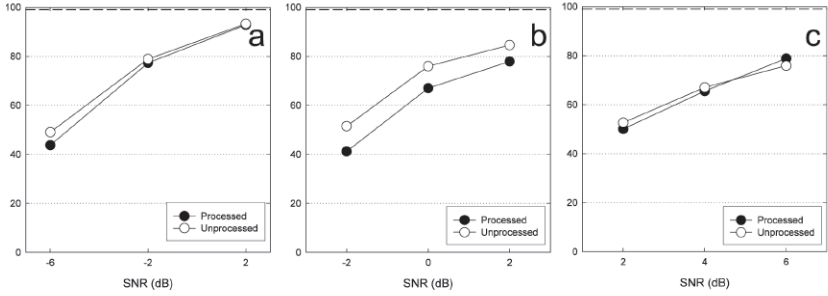
Fig. 5: Average speech intelligibility performance for 25 NH listeners (panel a) and two groups of 3 HI listeners (panels b and c), as a function of SNR and NR processing. The dashed reference lines show performances in quiet.

Figure 6 shows average reaction times (RTs) as a function of SNR, processing, and listener group. NH listeners (panel a) responded in 620 ms approximately, when the sentences were presented in quiet. Adding noise to the sentences at an SNR of 2 dB increased RTs by approximately 40 ms. With unprocessed sentences, each 4dB drop in SNR resulted in an increase in RTs of approximately 40 ms. With sentences that were processed with the NR algorithm, a drop in SNR from 2 to -2 dB resulted in a similar 40-ms increase in RTs, but there was no further increase with a subsequent 4dB drop in SNR. The first group of HI listeners (panel b) responded as quickly to the visual stimuli as the NH group, when the sentences were presented in quiet. Moreover, RTs were slower in the presence of noise than in quiet, for all SNRs, though there was less evident a relationship. Typically, however, RT's were faster when the sentences were processed than when they were unprocessed. The second group of HI listeners (panel c) was approximately 400 ms slower in their RTs than the other two groups, in the quiet condition. The presence of noise slowed down response times further, and there was a trend for higher RTs with lower SNRs. Unlike the other two groups, however, RT performance was typically slower when the sentences were processed, than when they were unprocessed.
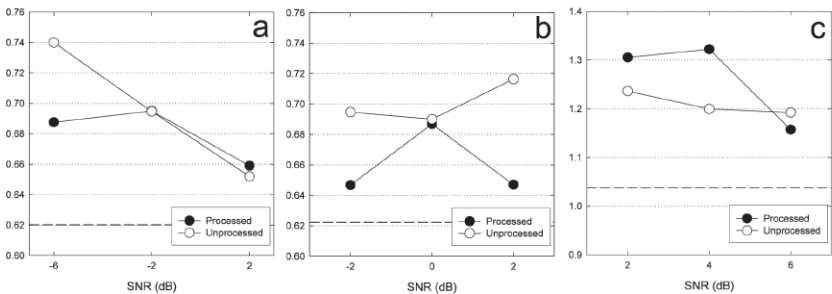

Fig. 6: Average reaction times for 25 NH listeners (panel a) and two groups of 3 HI listeners (panels b and c), as a function of SNR and NR processing. The dashed reference lines show performance in quiet.

574

CONCLUSION

The results of these experiments suggest that the presence of background noise increases the effort required to understand speech. Furthermore, the amount of listening effort increased as the SNR decreased, at least with NH listeners. The Ephraim-Mallah (1984, 1985) NR algorithm did not provide any speech intelligibility benefits for either NH or HI listeners. In fact, for some HI listeners, the NR was detrimental to speech intelligibility performance. With NH listeners the Ephraim-Mallah (1984, 1985) NR algorithm can reduce listening effort, allowing more resources to be made available to other processes, such as recall and speeded motor responses. It is unclear why the two groups of HI listeners showed such different responses to NR. One possibility is that the effectiveness of signal processing depends upon cognitive function. Lunner (2003) reported that listeners with higher scores on a cognitive test made better use of a NR scheme and, in the present study, it was the younger HI listeners whose reaction times improved with NR while reaction times with the older HI group grew worse. Is this necessarily a warning that older listeners cannot use signal processing effectively or is it, perhaps, that they require more practice with a new listening task? This will have to be studied in the future with more subjects. Nevertheless, with NH and younger HI listeners, the reaction-time data support our original hypothesis, namely that the relation between noise reduction and listening effort in extraction of speech from noise can be seen in its effects on other cognitive functions.

REFERENCES

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (**1984**). "Standardization of a Test of Speech-Perception in Noise" J. of Speech and Hear. Res. **27**, 32-48.

Broadbent, D. E. (**1958**). "Perception and Communication," Pergamon Press, London.

Ephraim, Y., and Malah, D. (**1984**). "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator," IEEE Transactions on Acoustics, Speech and Signal Processing **32**, 1109-1121.

Ephraim, Y., and Malah, D. (**1985**). "Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator," IEEE Transactions on Acoustics Speech and Signal Processing **33**, 443-445.

IEEE (**1969**). "IEEE Recommended Practice for Speech Quality Measurements," IEEE Transactions on Audio and Electroacoustics **Au17**, 225-246.

Kahneman, D. (**1973**). "Attention and Effort," Prentice-Hall, New Jersey.

Lunner, T. (**2003**). "Cognitive function in relation to hearing aid use," Intern. J. of Audiol. **42**, S49-S58.

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (**1995**). "How Young and Old Adults Listen to and Remember Speech in Noise," J. Acoust. Soc. of Am. **97**, 593-608.