DOES A CONTINUOUS MASKER MAKES SPEECH COMPREHENSION IN NOISE EFFORTFUL?

Antje Heinrich and Bruce A. Schneider
Department of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd., Mississauga, ON, Canada, L5L 1C6
aheinrich@utm.utoronto.ca

1. INTRODUCTION

A common complaint among older adults is that it is difficult for them to understand spoken language, especially in a noisy environment (CHABA, 1988). To comprehend speech in noise listeners not only have to attend to the target speech while ignoring distractors, they also have to extract the meaning from the individual words and phrases and store the result in memory for future use. It is possible that task of segregating the speech signal from the background, because it might require the redeployment of cognitive resources to perceptual processes, could adversely affect memory of the heard material, thereby reducing comprehension on a cognitive level. To see whether the effort of extracting the speech signal from noise adversely affected memory, Murphy, Li, Craik and Schneider (2000) investigated recall performance in a paired-associate memory paradigm when the word pairs were presented in quiet or embedded in a background babble presented at a signal-to-noise ratio (SNR) of –7 dB. Murphy et al. demonstrated that memory performance was especially affected for word pairs presented early on the list (early serial positions) when the word were presented in a background babble compared to the quiet baseline condition. In fact, presenting the words in noise decreased young adults' memory performance to a level comparable to old adults' when they listened to the word pairs in quiet.

Two explanations are conceivable for this result. First, the babble may have degraded the perceptual representation of the words, and an instable or incomplete representation compromised their subsequent processing. Second, the young adults may have had to engage top-down processes when listening to the words in background noise to recover the incomplete or distorted words. These top-down processes may have taken up cognitive resources that otherwise would have been available for the encoding of the words.

We tested both hypotheses. To test the first hypothesis we temporally distorted the word pairs. This manipulation decreased perceptual accuracy of the words to the same extent as the babble, yet, no masker was present from which to extract the words. The results are published elsewhere (Heinrich and Schneider, submitted) and, in short, show that temporal distortion did not adversely affect recall as long as the words were presented on a moderate intensity level. Based on these results we concluded that it was not the degraded perceptual representation of the words themselves that led to the decrement in memory performance in the Murphy et al. study. Rather, it was the effort in extracting the words from the background babble that caused the drop in memory performance. Evidence that a noise masker may involuntarily engage top-down processes to improve accuracy of stimulus perception stems from research on stimulus detection in gated noise. A number of studies show that detection performance is improved when a noise masker is gated on prior to stimulus onset compared to simultaneous onset of stimulus and masker (e.g., Zwicker, 1965). This result suggests that listeners track the noise in order to minimize its effect on stimulus perception. We speculated that a similar mechanism operated when words were presented in babble– the auditory system tracked to background babble between words to enable it to be better able to extract the target words. We reasoned that if this tracking process was effortful (requiring considerable cognitive processing) that fewer resources would be available to store these words in memory.Hence we would expect a drop in memory performance under this condition.

In order to test our hypotheses, we presented the word pairs in discontinuous babble that was present only during word presentation but not in the interval between words. This preserved the masking character of the babble at word presentation but again did prevent the auditory system from effectively being to monitor the babble over longer periods of time in order to minimize its influence on word perception. Subsequently, we compared recall performance from this condition with memory in conditions in which word pairs were presented in quiet and in a continuous babble.

2. METHOD

2.1 Participants

Sixteen young adults (mean age: 19.69 years; s.d.: 1.35, 11 females) were tested in this experiment. All individuals were undergraduate students at the University of Toronto, had completed an average of 15.06 years of education (s.d.: 1.88), and scored 13.38 (s.d.: 1.62) on the Mill-Hill vocabulary test. In exchange for their participation they received $10 per hour. All participants spoke English before the age of five. The demographic data are comparable to the young adults tested by Murphy et al. in the other two conditions.

2.2 Material

The same word pairs as in Murphy et al. (2000) were used. The material consisted of a total of four hundred two-syllable common nouns arranged in 40 lists of 5 word pairs each. The words were digitized at a sampling rate of 20 kHz. For stimulus presentation, words as well as babble, stimuli were delivered through a 16-bit digital-to-analog converter (TDT DD1) followed by a 10-kHz low-pass filter (TDT FT6-2, 60 dB attenuation at 11.5 kHz), a programmable attenuator (TDT PA4), and a weighted signal mixer (TDT SM3). All testing took place in a double walled sound attenuating chamber.

2.3 Procedures

For each listener pure-tone air-conduction thresholds were determined for nine frequencies between 250 and 8000 Hz in the left and right ear. All listeners had pure tone audiometric hearing thresholds within the normal range (< 25 dB HL) for frequencies between 250 and 8000 Hz.. Thresholds for the detection of speech babble were also determined for each individual. A recording of a twelve-talker babble, taken from the modified Speech Perception in Noise (SPIN) test (Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984) was used. The speech threshold was obtained for the right ear first and the left ear second. For the presentation of the words in babble, the presentation level of the words was individually adjusted for each participant so that each word was presented 50 dB above the individual's babble threshold. The babble level was also individually adjusted based on each individual's low-context SPIN threshold to reflect a SNR=-7 dB. For more details on the adjustment procedure please refer to Murphy et al. (2000). For the memory experiment, all participants listened to forty lists containing five word pairs each. In the encoding phase, word pairs were played with a presentation rate of 4 seconds per pair. The two words in a pair were separated by a silence period of 100 ms. The words were randomly paired, any obvious association between two words was avoided. In the retrieval phase, participants were cued with the first word from one of the five previously presented word pairs and were asked to recall the second word of the pair. Only one pair from each list was cued. There was no time limit placed on the recall. Participants were encouraged to guess.

3. RESULTS

The recall performance under discontinuous noise conditions displayed a primacy as well as a recency effect. Correct recall was close to 45% for the first two word pairs in a list, dropped to 38% for the third word pair, and went up to 54% and 79% for serial positions 4 and 5. This current result together with the results from the quiet baseline and continuous noise conditions from the Murphy et al. study was submitted to a mixed measures ANOVA with three background conditions (quiet, continuous noise, discontinuous noise) as between and serial position as within-subject variables. The results show a main effect of serial position ($F(4, 172)= 67.9$, $p<0.001$), condition ($F(2, 43)=4.5$, $p<0.05$), and a significant interaction between the two factors ($F(8, 172)=2.6$, $p<0.05$). Post-hoc test reveal that the main effect of background condition is caused by significant differences in memory recall between quiet and discontinuous noise on the one hand and continuous noise on the other hand. There are no significant differences in memory between the former two conditions. Moreover, the interaction effect is due to the fact that memory performance under continuous noise is considerably decreased for early serial positions (word pairs 1 to 2) compared to the other two conditions but not for serial positions 3 to 5.

4. DISCUSSION

The results demonstrate that it is the continuity of the noise that caused the drop in memory. When the noise was only present during word presentation but not in the interval between words, recall was not affected compared to the baseline condition. Under both these conditions the memory performance is considerably higher than when the words are masked with a continuous masker. Why does a continuous masker cause a drop in recall when a discontinuous masker does not? We assume that the auditory system uses cues available in the continuous babble in order to improve word perception. To do this, it monitors the babble in intervals where no words are presented so that it can use abrupt changes in the auditory stream to identify when words are being presented. However, this process is effortful and resource demanding and resources invested in minimizing the influence of the background noise are not available for memory processing. This lack of resources at encoding leads to the decrement in memory performance

REFERENCES

Bilger, R.B., Nuetzel, M.J., Rabinowitz, W.M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. Journal of Speech and Hearing Research, 27, 32-48.

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics) Working Group on Speech Understanding and Aging, National Research Council (1988). Speech understanding and aging. Journal of the Acoustical Society of America, 83, 859-895.

Murphy, D.R., Craik, F.I.M., Li, K.Z.H. & Schneider, B.A. (2000). Comparing the effects of aging and background noise on short-term memory performance. Psychology and Aging, 15 (2), 323-334.

Zwicker, E. (1965). Temporal effects in simultaneous masking by white-noise bursts. Journal of the Acoustical Society of America, 37(4), 653-663.