Speech understanding in everyday life: the role of masking release, binaural unmasking, and linguistic proficiency

S. Theo Goverts, Joost M. Festen, and Tammo Houtgast

ENT-Audiology and EMGO+ institute, VU University Medical Center, P.0. Box 7057, 1007 MB, Amsterdam, The Netherlands

Hearing-impaired subjects experience major problems understanding speech in everyday adverse conditions. In these conditions normal-hearing listeners have substantial benefit from masking release due to fluctuations in the noise-masker, binaural unmasking, and linguistic proficiency. In the present study these effects are investigated in normal-hearing listeners in a *combined* condition, which is relevant for everyday life. Masking release and binaural unmasking are found to be sub-additive, masking release and linguistic proficiency are super-additive. Binaural unmasking and linguistic proficiency are independent.

INTRODUCTION

Hearing-impaired subjects experience major problems understanding speech in everyday adverse conditions. In these circumstances normal-hearing listeners have substantial benefit from fluctuations in the noise-masker (*masking release*) and from spatial separation between the speech and the masker sources (*binaural/spatial unmasking*). Also the role of *linguistic proficiency* in speech intelligibility can be substantial (van Wijngaarden *et al*., 2002). In these effects, both bottom up and top-down processes are involved (Stenfelt and Ronnberg, 2009), that can be affected by hearing loss.

Hearing-impaired listeners are known to have reduced masking release (Festen and Plomp, 1990; George *et al*., 2007). Reduced binaural unmasking was found for some hearing-impaired listeners (Bronkhorst and Plomp, 1990; Goverts and Houtgast, 2009). Hearing loss, especially of congenital origin, may cause reduced linguistic skills, with possible effects on speech intelligibility comparable to those for the non-natives listeners.

Masking release and binaural unmasking are often investigated separately, whereas in everyday life they typically occur together. In an earlier study we demonstrated that masking release and binaural intelligibility level difference (BILD) are not fully additive for normal-hearing listeners (Goverts *et al*., 2007). In a group of 12 normal-hearing listeners we found for different types of masker, that the summed effect of masking release and BILD is significantly more than the enhancement in the combined condition. In that study it was hypothesized that the reduced BILD for fluctuating maskers is caused by a reduced proportion of time for which the instantaneous signal-to-noise ratio is in the range in which binaural unmasking is active. Using this hypothesis of a reduced effective diotic masker it was possible to

Proceedings of ISAAR 2009: Binaural Processing and Spatial Hearing. 2nd International Symposium on Auditory and Audiological Research. August 2009, Elsinore, Denmark. Edited by J. M. Buchholz, T. Dau, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN: 87-990013-2-2. The Danavox Jubilee Foundation, 2010.

predict in a very qualitative way binaural unmasking for a block-modulated masker, a masker with speech-like modulations and an interfering talker.

In the present study, speech understanding is studied in different conditions of masker, interaural presentation and linguistic content for normal-hearing listeners. Especially, the interactions between masking release, binaural unmasking, and linguistic proficiency are investigated. This is just one step towards the final goal: understanding the contribution of auditory and non-auditory factors to everyday speech understanding in combination with hearing impairment. The data that are used are part of larger dataset of a study on the role of linguistic skills in speech intelligibility in normal-hearing listeners.

METHOD

Speech understanding is measured as the speech reception threshold (SRT), i.e. the signal-to-noise ratio (SNR) at which 50% of short sentences is reproduced correctly (Plomp and Mimpen, 1979). In the measurement procedure the SNR is adjusted by adaptively varying the speech level in steps of 2 dB. The SRT is calculated as the average SNR over the last 10 out of 13 sentences.

Dutch and German sentences, recorded by Van Wijngaarden *et al*. (2002) and based on the original Plomp and Mimpen material, are used in the SRT-measurements. For each set of speech-stimuli a stationary masking noise was constructed with a spectrum equal to the long-term average of the sentences. Based on this stationary masker a fluctuating masker was constructed with 16-Hz block-shaped modulation, 50% duty cycle, and infinite modulation depth.

The SRT-measurements are performed with one retest. The noise level was fixed at 65 dB SPL. SRTs were measured for eight different conditions,

- masker: stationary noise and 16-Hz chopped noise
- interaural presentation: N_0S_0 and $N0S_\pi$
- linguistic content: foreign (German) and native (Dutch) speech material, to calculate masking release, binaural unmasking, linguistic proficiency, and their interactions.

As the study aims at investigating the difference between the conditions, the order of conditions was balanced over the subjects. The entire experiment was controlled by a personal computer and the sentences were presented to the subjects over Sony MDR-V900 headphones. Subjects were tested individually in a soundproof room.

24 native Dutch listeners with normal hearing (thresholds between 250 and 8000 Hz < 20 dB), mean age 22 (range 18-29), and university degree participated. The results are compared to earlier unpublished results on the combined effect of masking release and binaural unmasking for 9 native Dutch mild to moderate hearing-impaired listeners, mean age 68 (range 46-83). In this study the Versfeld *et al*. (2000) speech material was used.

RESULTS

The SRT data for the different conditions are given in Fig. 1. An ANOVA of the SRT results shows significant effects for Masker ($F(1,23)=236$, $p<0.001$), Binaural presentation mode ($F(1,23)=71$, $p<0.001$), and Linguistic content ($F(1,23)=88$, $p<0.001$).

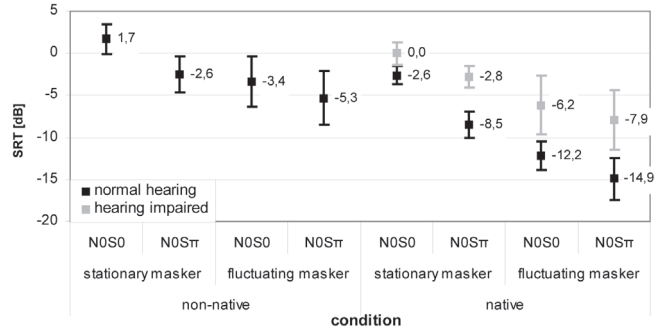

Fig. 1: Average SRT data and standard deviations for conditions of masker, bilateral presentation, and linguistic content. Unpublished hearing-impaired data are shown in grey for comparison.

Significant interactions are found for Masker*Binaural presentation mode ($F(1,23)=30$, $p<0.001$), and Masker*Linguistic content ($F(1,23)=50$, $p<0.001$), but not for Linguistic content*Binaural presentation mode ($F(1,23)=0$, $p=0{,}518$). Scatter-plots for these interactions are given in Fig 2.

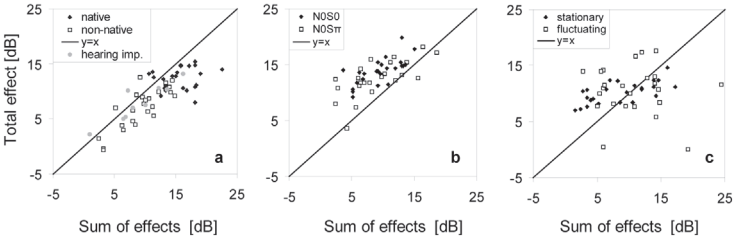

Fig. 2: In each panel the total benefit in the combined condition of two effects is plotted versus the sum of those effects for individual listeners. Panel (a): binaural unmasking and masking release; Panel (b): masking release and linguistics proficiency; Panel (c): binaural unmasking and linguistic proficiency. Data below y=x indicate sub-additivity, above super-additivity.

DISCUSSION

The role of masking release, binaural unmasking, and linguistic proficiency

The present study shows significant effects of masking release, binaural unmasking, and linguistic proficiency on speech understanding for normal-hearing listeners. The size of these effects is in line with typical literature findings for conditions where effects are investigated in isolation. Comparing the SRT in stationary noise and fluctuating noise for diotic presentation mode (N_0S_0) and native language yields a masking release of about 9 dB (e.g. George *et al*., 2007). Comparing SRT in N_0S_0 and N_0S_π for stationary noise and native language yields a value for binaural unmasking of about 6 dB (e.g. Johansson and Arlinger, 2002). Comparing SRT of foreign language and SRT for native language for stationary noise and diotic presentation yields a value for linguistic proficiency of about 4 dB (e.g. van Wijngaarden *et al*., 2002) .

Comparing the most adverse condition (non-native, stationary noise, N0S0) and the condition that turns out to be the most beneficial (native, fluctuating, N0Sπ) yields a large enhancement of speech understanding in the combined condition: nearly 17 dB shift of SNR from 1.7 to -14.9 dB.

Interactions

Figure 2a shows that masking release and binaural unmasking are sub-additive. Comparing the SRT for the conditions (native, N_0S_0, stationary) and (native, N_0S_π, fluctuating) yields a benefit of about 12 dB, which is less than simply adding the effects of masking release and binaural unmasking (9 + 6 = 15 dB). This is in line with the hypothesis of a reduced effective diotic masker causing reduced binaural unmasking for this type of stimuli (Goverts *et al*., 2007). These are both bottom-up processes.

Masking release and linguistic proficiency are super-additive. Comparing the SRT for the conditions (non-native, N_0S_0, stationary) and (native, N_0S_0, fluctuating) noise yields a benefit of about 14 dB, which is more than simply adding the effects of masking release in the non-native condition and linguistic proficiency (4 + 5 = 9 dB). Linguistic skills are beneficial in filling the gaps, which is one of the top-down processes involved in masking release.

No significant interaction between binaural unmasking and linguistic proficiency is found. A reason for this mutual independence might be that binaural unmasking is mainly (if not totally) a bottom-up phenomenon, whereas linguistic proficiency is top-down.

Hearing impairment

The final goal of this research is investigating how hearing impairment influences the role of those effects in everyday life. As an illustration, unpublished SRT data of hearing-impaired listeners in the same conditions of masker and bilateral presentation are added. These data (Fig. 1) show that the size of the effect of hearing impairment is

about the same as that of non-nativeness. However, whereas the reduced scores in the non-native condition must be caused by reduced top-down compensation resources, with hearing impairment the reduced (peripheral) temporal acuity will play a role. This might explain their less pronounced sub-additivity (Fig 2a). This means that less reduction of the effective diotic masker can lead to less reduction of binaural unmasking.

CONCLUSION

It is demonstrated that normal-hearing listeners can have substantial benefit of stimulus properties in speech understanding in everyday life: fluctuations in the masker, interaural phase delay for the speech, and native language. In the combined condition a benefit of about 17 dB is found. Masking release, binaural unmasking and linguistic proficiency don't add up in this condition. Masking release and binaural unmasking are sub-additive, while masking release and linguistic proficiency are super-additive. Binaural unmasking is independent of linguistic proficiency.

Hearing-impaired listeners will have less benefit in such a combined condition. A small pilot study with hearing-impaired listeners has shown that besides the reduced scores on masking release and binaural unmasking, the interaction between those effects seems to be different. This needs to be studied for the other interactions too. Hearing loss, especially of congenital origin, may cause reduced linguistic skills, with effects on speech intelligibility comparable to those for the non-natives listeners. Both native hearing-impaired and non-native normal-hearing listeners have less benefit in conditions with fluctuating maskers.

These results have to be taken into account when interpreting studies on these effects separately. The results support a simplified scheme in which binaural unmasking is a predominantly bottom-up effect and linguistic proficiency is a predominantly top-down effect. Both effects interact with masking release because this effect comprises both bottom-up and top-down aspects.

ACKNOWLEDGEMENTS

We thank Esther Bron for conducting the larger experiment, where the current data are part of. Furthermore we thank Sander van Wijngaarden for providing the speech material and Hans van Beek for technical support.

REFERENCES

Bronkhorst, A. W., and Plomp, R. (**1990**). "A clinical test for the assessment of binaural speech perception in noise," Audiology **29**, 275-285.

George, E. L. J., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., and Houtgast, T. (**2007**). "Auditory and non-auditory factors affecting speech reception in noise by older listeners," J. Acoust. Soc. Am. **121**, 2362-2375.

Goverts, S. T., Delreux, M., Festen, J. M., and Houtgast, T. (**2007**). "The influence of masker type on the binaural intelligibility level difference," in *Hearing, from sensory processing to perception*, edited by B. Kollmeier *et al*. (Springer Berlin Heidelberg).

Goverts, S. T., and Hougast, T. (**2009**). "The BILD of hearing-impaired subjects - the role of suprathreshold coding," *submitted* to J. Acoust. Soc. Am.

Johansson, M. S. K., and Arlinger, S. D. (**2002**). "Binaural masking level difference for speech signals in noise," Audiology **41**, 279-284.

Plomp, R., and Mimpen, A. M. (**1979**). "Improving the reliability of testing the speech reception threshold for sentences," Audiology **18**, 43-52.

Stenfelt, S., and Rönnberg, J. (**2009**). "The Signal-Cognition interface: interactions between degraded auditory signals and cognitive processes," Scand J Psychol. **50**, 385-393.

Versfeld, N. J., Daalder, L., Festen, J. M. and Houtgast, T. (**2000**). "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," J. Acoust. Soc. Am. **106**, 1671-1684.

van Wijngaarden, S. J., Steeneken, H. J. M., and Houtgast, T. (**2002**) "Quantifying the intelligibility of speech in noise for non-native listeners," J. Acoust. Soc. Am. **111**, 1906-1916.