Overlap behaviour in task-oriented dialogue

Vincent Aubanel[1,2], Martin Cooke[1,2], Catherine Mayo[3] and Robert Clark[3]

[1]Language and Speech Laboratory, University of the Basque Country, Vitoria, Spain
[2]Ikerbasque (Basque Foundation for Science)
[3]Centre for Speech Technology Research, University of Edinburgh

`v.aubanel@laslab.org`

Abstract

Speakers change the way they speak depending on the surrounding environment. When masking noise obstructs the communication channel between interlocutors, they consistently engage in Lombard speech, whose spectral characteristics are well described (e.g. [1, 2, 3]) and are believed to result in better intelligibility through energetic masking reduction (e.g. [4]). However, less is known about how speakers adapt to the temporal characteristics of a fluctuating masker, and whether any such changes aid communication.

In this study pairs of speakers were recorded while engaged in a sudoku-solving task in quiet and in several masking conditions. Maskers were either a competing talker or speech modulated noise with identical temporal characteristics, chosen to investigate the informational masking potential of the masker. The silence density of each masker was also manipulated by adjusting the durations of pauses in the masker to 33% or 66% of the overall duration of the masker.

In all masking conditions, speakers displayed a reduction of overlap with the masker relative to a baseline computed from the masker and speech produced in quiet (Figure 1). The overlap reduction tended to be larger for less
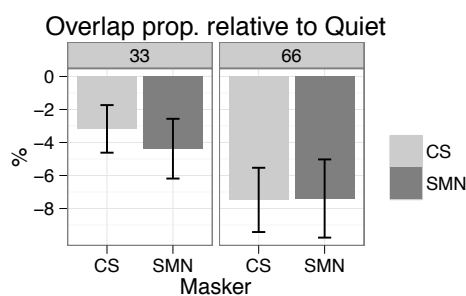
Figure 1: Overlap reduction relative to quiet in two silence densities: (33% and 66%), for a masker of competing speech (CS) and speech modulated noise (SMN).

dense maskers, and was obtained, at least for the competing talker case, in spite of an increase of speech activity in the masking condition. These results confirm [5], and indicate that speakers were able to exploit the temporal silences of the masker to convey information necessary to complete the task.

While denser maskers led to stronger Lombard effects, such as greater increase of $F0$ and $F1$ compatible with the energetic masking potential of the masker, less dense maskers induced a qualitatively different behaviour in speakers, which we characterise as a "wait and talk strategy", demonstrated by a decrease of speaker onsets following masker onsets and an increase of these events following masker offsets.

Taken together, these results suggest that talkers respond to their environment by making speech modifications which have the potential to help communicate with their interlocutors.

References

[1] W. Van Summers, D. B. Pisoni, R. Bernacki, R. Pedlow, and M. Stokes, "Effects of noise on speech production: Acoustic and perceptual analyses," *J. Acoust. Soc. Am.*, vol. 84, no. 3, pp. 917–928, 1988.

[2] J. C. Junqua, "The Lombard reflex and its role on human listeners and automatic speech recognizers," *J. Acoust. Soc. Am.*, vol. 93, no. 1, pp. 510–524, 1993.

[3] M. Garnier, "Communiquer en environnement bruyant : de l'adaptation jusqu'au forçage vocal ," Ph.D. dissertation, Université Paris 6, 2007.

[4] Y. Lu and M. Cooke, "Speech production modifications produced by competing talkers, babble, and stationary noise," *J. Acoust. Soc. Am.*, vol. 124, no. 5, pp. 3261–3275, 2008.

[5] M. Cooke and Y. Lu, "Spectral and temporal changes to speech produced in the presence of energetic and informational maskers," *J. Acoust. Soc. Am.*, vol. 128, no. 4, pp. 2059–2069, 2010.