Title: Neural correlates of informational masking in cocktail party situations

Authors: Marjorie Dole, Michel Hoen, Fanny Meunier, Véronique Boulenger, Jean-Luc Anton, Bruno Nazarian, Muriel Roth, Lionel Collet.

Introduction

The *cocktail party* effect refers to situations in which people must segregate auditory sources. One situation of interest is speech-in-speech comprehension, in which two types of masking are described: *energetic masking* produced when target and noise partially overlap in time and frequency; *informational masking* occurs when information from concurrent flows is of comparable nature. Typically, informational masking is observed when target and background are speech sounds and it has been shown to emerge at different levels of competitions, namely acoustic-phonetic and lexical (Boulenger et al., 2010; Hoen et al., 2007). Using fMRI, Scott et al. (2004) showed increased activations in temporal superior gyrus associated with informational masking, whereas energetic masking involved a network including right posterior parietal cortex, left dorsolateral prefrontal cortex and left frontal pole. The aim of our study is to further characterize cortical networks involved in informational masking.

Methods

We designed an fMRI study using *interleaved silent steady state (ISSS)* imaging (Schwarzbauer et al., 2006) to examine cortical activations during a speech-in-noise intelligibility task. This sparse method allows eliminating the confounding factor produced by the scanner noise on the cortical activations elicited by our auditory stimuli, by acquiring five brain volumes following a silent stimulus presentation period.

Fifteen French native speakers with no hearing or language disorders participated in our experiment. They were asked to carefully listen to 120 stimuli composed of words embedded in different types of noise; two words were then presented on a screen and participants had to select by a button press the word they heard.

To dissociate the different levels at which informational masking occurs, three types of noise were used: a speech-derived Broadband Noise (Noise) producing pure energetic masking of the same magnitude as all following sounds, a 4-talkers babble noise (Cocktail), and the same babble sound but reversed along its temporal axis or reversed babble, keeping some phonetic cues but no lexical or semantic information (Reverse). Baseline was obtained by presenting stimuli in silence (Silence). This led to 4 conditions, with 30 words per condition and divided into three sessions.

Data analysis was performed using SPM5. Images were realigned, normalized to the T1 MNI template and smoothed with a $10mm^3$ Gaussian kernel. For each subject, four regressors of interest (for each condition) were modelled with a Finite Impulse Response with 5 time bins. Individual contrasts were computed for each time bin. At the second (group) level, a random-effect analysis was performed, using a 1-way ANOVA with Time Bin as factor, for each contrast.

Results

Behavioural results first showed that the type of background noise significantly affected participants' performances to identify target words. Identification rate was better in Noise than in Reverse and Cocktail. These latter two conditions also differed from each other, performances being worse when the babble was natural speech (Cocktail) compared to reversed speech (Reverse) in which informational masking is reduced.

Imaging results revealed that specific cortical networks were involved during speech-in-speech comprehension depending on the nature of the background noise. Generally,

informational masking (Cocktail – Noise) activated the left primary auditory cortex (BA 41), the bilateral superior temporal gyrus (BA 22) and the left supramarginal gyrus (BA 40).
Direct contrasts between the Reverse and Cocktail conditions further allowed breaking down informational masking into its different parts. Informational masking mainly at the phonological level (Reverse – Noise), that is when the babble contained only partial phonetic information, activated the left primary auditory cortex (BA 41) and the bilateral middle part of the superior temporal gyrus (BA 22). Informational masking mainly at the lexico-semantic level (Cocktail – Reverse), that is when the babble was composed of linguistic information, engaged the left middle temporal cortex (BA 21) and the posterior part of the superior temporal gyrus (BA 22).

Conclusions
This study was designed to examine cortical activations produced by a speech-in-noise task. Behavioural results showed that performance is modulated depending on the level at which informational masking occurs (phonological or lexico-semantic).
fMRI data revealed that a large part of the bilateral superior temporal gyri responds to informational masking. This was previously observed (Scott et al., 2004). When two components of informational masking were distinguished, we found that masking at the lexical level engaged only the left posterior superior temporal cortex. Given that multi-talker babble contains multiple levels of linguistic information, this is consistent with data from the literature that processing intelligible speech activates area 22 and that lexico-semantic activations are found in area 21. Informational masking at a lower phonological level activated bilateral regions in the superior temporal gyrus, possibly reflecting the processing of acoustic-phonetic information.

References
Boulenger, V., Hoen, M., Ferragne, E., Pellegrino, F. & Meunier, F. (2010). 'Real-time lexical competitions during speech-in-speech comprehension', *Speech Communication*, vol. 52, no 3, pp. 246-253.
Hoen, M., Meunier, F., Grataloup, C.L., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X. & Collet, L. (2007), 'Phonetic and lexical interferences in informational masking during speech-in-speech comprehension', *Speech Communication*, vol. 49, pp. 905-916.
Nairain, C., Scott, S.K., Wise, R.J.S., Rosen, S., Leff, A., Iversen, S.D. & Matthews, P.M. (2003). 'Defining a left-lateralized response specific to intelligible speech using fMRI', *Cerebral Cortex*, vol.13, no 12, pp.1362-1368.
Scott, S.K., Rosen, S., Wickham, L. & Wise, R.J.S. (2004), 'A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception', *Journal of Acoustical Society of America*, vol. 115, no.2, pp. 813-821.
Scott, S.K., Rosen, S., Beaman, C.P., Davis, J.P. & Wise, R.J.S. (2009), 'The neural processing of masked speech: Evidence for different mechanisms in the left and the right temporal lobes', *Journal of Acoustical Society of America*, vol. 125, no.3, pp. 1737-1743.
Schwarzbauer,C., Davis, M.H., Rodd, J.M. & Johnsrude, I. (2006), 'Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI', *NeuroImage*, vol. 29, pp. 774-782.