Song and Speech – Perception and Covert Production: New Findings using Multi-Voxel Pattern Analysis

D. Goldhahn[1], D. E. Callan[2], G. Lohmann[1], and R. Turner[1]
[1]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2]ATR Computational Neuroscience Laboratories, Kyoto, Japan

Introduction

Multi-Voxel Pattern Analysis (MVPA) (Norman et al., 2006) is a powerful technique for the analysis of functional magnetic resonance imaging data. In contrast to standard univariate methods, which take into account only single voxels, MVPA analyzes the information present in multiple voxels. MVPA can not only identify regions that react more strongly to a condition, but it can also find areas of the brain where the fine spatial pattern of activation of several voxels discriminates between experimental conditions. In this study we use MVPA to analyze brain regions differentially involved with listening to and covert production of song relative to speech. We present new findings that univariate analysis (Callan et al., 2006) failed to discover, and investigate what underlies this discrepancy.

Methods

The experiments were conducted on a 3-T MRI system. A T2*-weighted gradient echo-planar imaging sequence (TE = 30 ms; TR = 4000 ms; flip angle 90°) was used to acquire 36 contiguous axial slices with 3.0x3.0x3.9-mm voxel resolution. Each subject underwent one run of 184 volumes. The first 4 scans were discarded. 16 volunteers with normal music experience were tested using a block design of 5 conditions (passive listening to song, passive listening to speech, covert production of song, covert production of speech, and rest). The stimuli consisted of 6 well known Japanese songs of equal length of about 20 s. Each condition was repeated 6 times, once per song. The aurally presented stimuli were adjusted for identical RMS energy. Production was covert to avoid motion artifacts. For MVPA the data were motion corrected and high pass filtered (cutoff period = 128 s). Analysis was conducted in original image space for each subject separately. Each run was divided into six parts, every part containing all four conditions of one song. The time-steps corresponding to two contrasted conditions were used to train a linear Support Vector Machine. The searchlight algorithm (r = 9 mm) and leave-one-out cross validation were used for whole brain analysis of the following contrasts: 1) listening to singing versus listening to speech 2) covert singing versus covert speech. The results were registered into MNI-space (2x2x2-mm resolution) and a group level one-sample t-test was conducted for each contrast. For univariate analysis the data were motion corrected and registered into MNI-space before smoothing (8x8x8-mm FWHM Gaussian kernel) and high pass filtering (cutoff period = 128s) were carried out. A general linear model (GLM) utilizing a boxcar function convolved with a hemodynamic response function was used for fixed effect analysis of the following contrasts for each subject 1) differences between listening to singing and listening to speech 2) differences between covert singing and covert speech. One-sample t-tests across subjects were conducted for each contrast.

Results and Discussion

For both contrasts the results of the univariate analysis are comparable to Callan et al., 2006. MVPA results resemble those of the GLM analysis, closely for the listening contrast, and more remotely for covert production. When contrasting the listening conditions both analysis techniques agree on several brain regions of differential activity, e.g. in superior temporal gyrus (STG) and superior frontal gyrus bilaterally. Areas such as right globus pallidus, left parahippocampal region and left lateral premotor cortex were revealed only in GLM analysis, while MVPA found discriminative activity in more frontal regions of right STG, as depicted in Figure 1, and in visual areas. When production of speech and song are contrasted, the differences of both analysis methods are more striking. As illustrated in Figure 2, MVPA discovers highly differential activity in large regions of STG bilaterally and in visual areas, while univariate analysis fails to detect activation in either of these regions. Reasons for these differences can be found in the fine spatial pattern of activity, e.g. in visual areas, which cannot be discovered using univariate analysis. Furthermore, the standard method only discovers regions where the sign of changes in MR image intensity is uniform within and across subjects. This condition is not fulfilled in either STG in the production condition contrasts. Average activation in individual subjects is either greater during singing, greater during speaking or similar for both conditions. MVPA detects the discriminative information of a brain region and therefore is not limited by the sign of the signal changes.

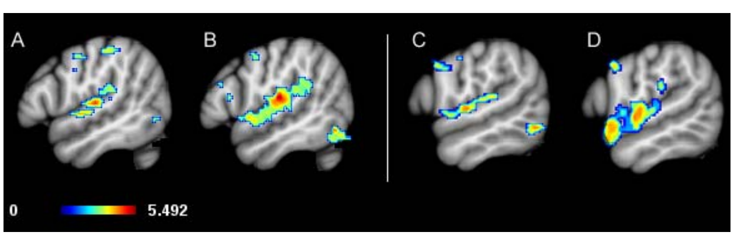

Figure 1: Listening to song versus listening to speech (threshold T=2.576, df=15, minimum spatial extent 10 voxels); A = left STG of GLM-analysis; B = left STG of MVPA; C = right STG of GLM-analysis; D = right STG of MVPA

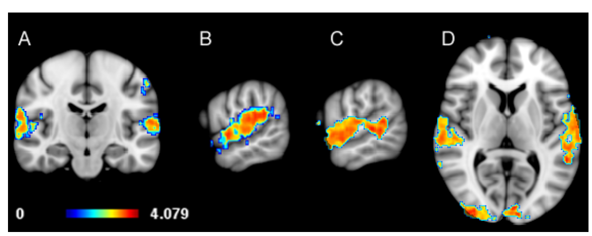

Figure 2: Covert production of song versus covert production of speech (same thresholds as Figure 1); A-D results of MVPA; B = left STG; C = right STG

References

Callan D.E., Tsytsarev V., Hanakawa T., Callan A.M., Katsuhara M., Fukuyama H., Turner R. Song and speech: Brain regions involved with perception and covert production. NeuroImage. 2006;31:1327–1342.

Norman K.A., Polyn S.M., Detre G.J., Haxby J.V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci. 2006;10:424–30.