Technical University of Denmark

Representational Similarity Analysis Reveals Heterogeneous Networks Supporting Speech Motor Control

Zheng, Zane; Cusack, Rhodri; Johnsrude, Ingrid ; MacDonald, Ewen; Munhall, Kevin; Vicente-Grabovetsky, Alejandro

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Representational Similarity Analysis Reveals Heterogeneous Networks Supporting Speech Motor Control

Presented During:

359 MT: Representational Similarity Analysis Reveals Heterogeneous Networks Supporting Speech Motor Control

Presented During: O-M3: Language
Monday, June 11, 2012: 6:30 PM - 6:45 PM

Poster No:

359

On Display:

Monday, June 11 & Tuesday, June 12

Authors:

Zane Zheng[1], Rhodri Cusack[2], Ingrid Johnsrude[1], Ewen MacDonald[3], Kevin Munhall[1], Alejandro Vicente-Grabovetsky[2]

Institutions:

[1]Queen's University, Kingston, Canada, [2]Medical Research Council Cognition and Brain Sciences Unit, Cambridge, United Kingdom, [3]Technical University of Denmark, Lyngby, Denmark

Poster Presenter(s):

Zane Zheng - Contact Me
Queen's University
Kingston, Canada

Introduction:

The everyday act of speaking involves the complex processes of speech motor control. One important feature of such control is regulation of articulation when auditory concomitants of speech do not correspond to the intended motor gesture. While theoretical accounts of speech monitoring posit multiple functional components required for detection of errors in speech planning (e.g., Levelt, 1983), neuroimaging studies generally indicate either single brain regions sensitive to speech production errors, or small, discrete networks. Here we demonstrate that the complex system controlling speech is supported by a complex neural network that is involved in linguistic, motoric and sensory processing. With the aid of novel real-time acoustic analyses and representational similarity analyses of fMRI signals, our data show functionally differentiated networks underlying auditory feedback control of speech.

Methods:

We developed a real-time speech-tracking system to deliver three types of auditory signal (clear speech, formant-shifted speech, and signal-correlated noise), which served as temporally-gated feedback during production and as stimuli during listening in a 2x3 factorial design. Twenty subjects either produce a monosyllabic word or listen during a 1600-msec silent period between successive 1600-msec acquisitions in a rapid sparse-imaging paradigm. Each subject was scanned on 72 trials from each of 6 conditions in pseudorandom order. Functional images were analyzed using SPM5 and a custom-made, modular toolbox (MVPaa). Three successive trials of the same condition were modeled as one regressor, leading to 24 ß estimates for each condition. A whole-brain searchlight analysis (Kriegeskorte et al. 2006) was performed on the ß images, comparing regressor-specific multivoxel patterns within each 4mm-radius searchlight using Spearman correlation. The resulting data were condensed into a 6x6 correlation matrix representing the similarity measures within and between conditions. A general linear model was then applied to the similarity measures to test hypothesis-driven predictions of pattern correlations. The resulting ß images were smoothed and entered into random-effects analyses. Clusters that survived the threshold of familywise error correction $p < .05$ were deemed significant.

Results:

We searched for neural patterns that were consistent when processing load was specifically placed on putative systems subserving articulatory control. Clusters in which multivoxel patterns exhibited high similarity between two erroneous feedback types (i.e., formant-shifted feedback and noise feedback) but low similarity between either erroneous feedback and clear speech feedback during production (Fig. 1a) included bilateral cerebellum, right supplementary motor area, and right angular gyrus (Fig. 2, cyan). These similarity patterns were not evident when assessed during listening (Fig. 1b), suggesting that this network encodes an 'error signal' during articulation irrespective of acoustic features of the error. The interaction (Fig. 1c) was confirmed using a paired t-test, which revealed additional areas including bilateral superior temporal gyrus (STG), bilateral inferior frontal gyrus (IFG), and left precentral gyrus. We found that the significant response in the bilateral STG and IFG network (Fig. 2, magenta) was driven by the pattern similarity between the two speech-like stimuli, whereas the response in the left precentral gyrus (Fig. 2, blue) was driven not by either, but by a combination of error-sensitivity and speech-sensitivity (i.e., aspects of cyan and magenta networks). These results were confirmed by ANOVAs demonstrating that the regions exhibiting the interaction comprised 3 functionally differentiated networks ($p = .006$).

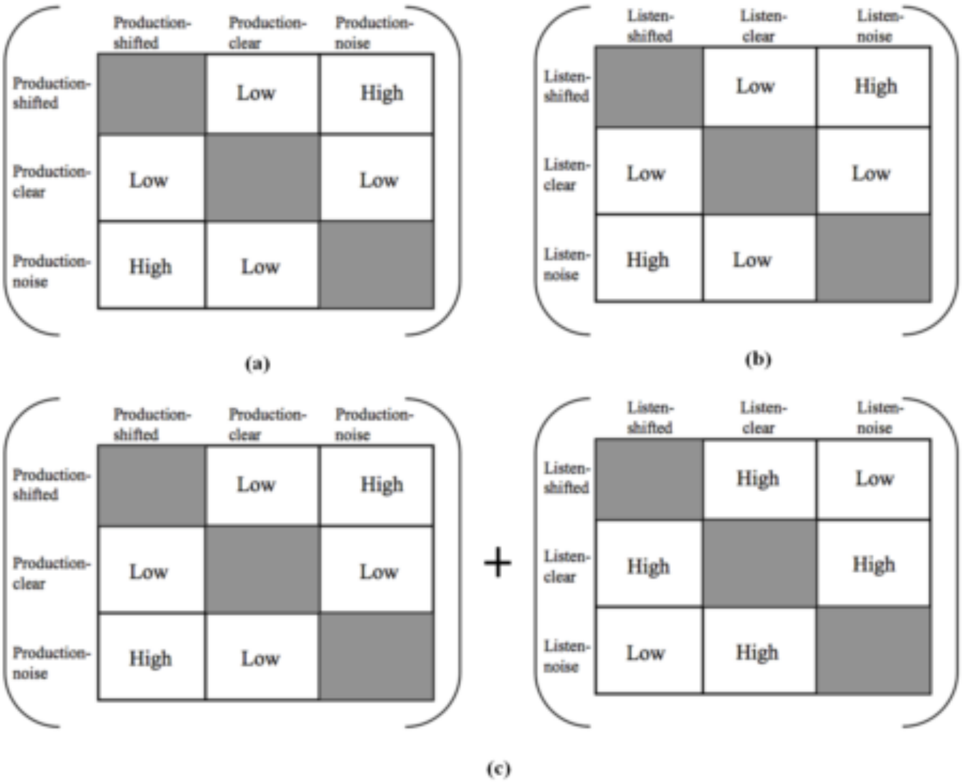

Figure 1. Our analysis was based on an assessment of the model fit between predicted and measured multivoxel pattern similarities within and between conditions. The predicted similarity structure matrices are schematically presented here. The between-condition similarity predictions ensure that the resulting multivoxel patterns contain information that is generalizable across different stimulus types that share a crucial 'feature', not just within a stimulus type. a) The similarity structure matrix that predicts an error-coding network during production, i.e., consistently high similarity between two erroneous feedback types but low similarity between either erroneous feedback and clear speech feedback. b) The similarity structure matrix from a), but with all production conditions replaced by listening conditions. c) The similarity structure matrix for an interaction equivalent to a) - b), highlighting brain regions involved in auditory feedback control of speech. Note that 'high' and 'low' labels on the right matrix were flipped from b). Here the right matrix highlights a speech -sensitivity network during listening, due to the similarity prediction between the two speech-like stimuli: clear speech and shifted speech.

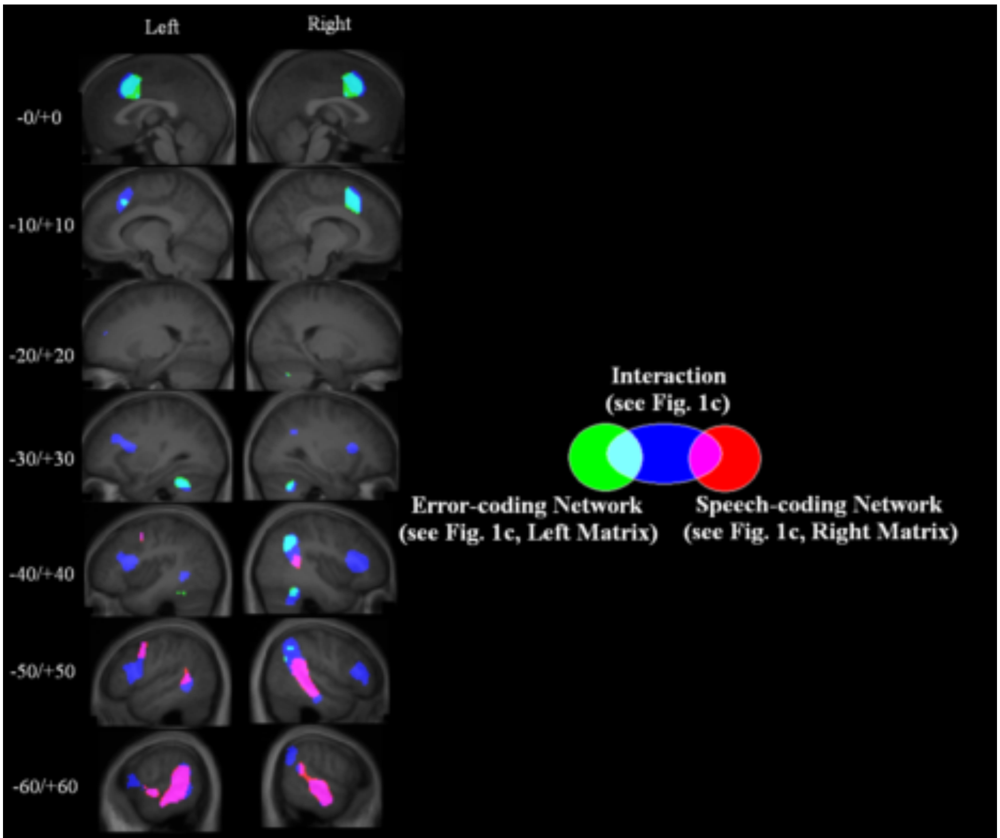

Figure 2. Three networks in which multivoxel patterns exhibited an Auditory Input by Production/Listening interaction are shown. This interaction was inclusively masked by production simple effect (i.e., error-coding network) and by inverse listening simple effect (i.e., speech-coding network) in order to demonstrate their relative contributions: predominantly driven by production simple effect (cyan), predominantly driven by inverse listening simple effect (magenta), and a combination of the two (blue).

Conclusions:

Our findings show for the first time that speech motor control relies on multiple, interactive, and functionally differentiated systems, offering insight into the complexity of the neural processes supporting speech movements.

Language:

Speech Production

Abstract Information

References

Levelt, W.J. (1983), 'Monitoring and Self-repair in Speech', Cognition, vol. 14, no. 1, pp. 41-104.
Kriegeskorte, N., Goebel, R., Bandettini, P. (2006), 'Information-based functional brain mapping', Proceedings of the National Academy of Sciences of the USA, vol. 103, no. 10, pp. 3863-3868.