

SUPPLEMENTARY MATERIAL

Supplementary Table 1

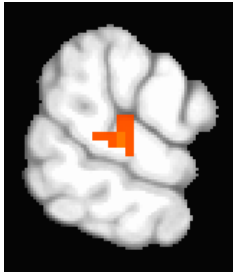
Behavioral Performance and Post-test Questionnaire Ratings for the Control Behavioral Experiment

	Live	Social	Standard	$F_{(2,36)}$	Pairwise Comparisons
BEHAVIORAL PERFORMANCE					
Accuracy (%)	84.72 (11.16)	82.87 (19.3)	86.57 (11.12)	.442	Live=Soc=Std
RT (ms)	1384 (244)	1421 (360)	1438 (312)	.348	Live=Soc=Std
POST-TEST QUESTIONNAIRE RATINGS					
Liveness	3.42 (.95)	3.13 (.85)	2.39 (.88)	9.89***	Live=Soc>Std
Likeability	3.84 (1.07)	4.00 (1.20)	3.39 (1.12)	2.02	Live=Soc=Std
Engagement	4.58 (.61)	4.47 (.77)	4.16 (.83)	2.11	Live=Soc=Std

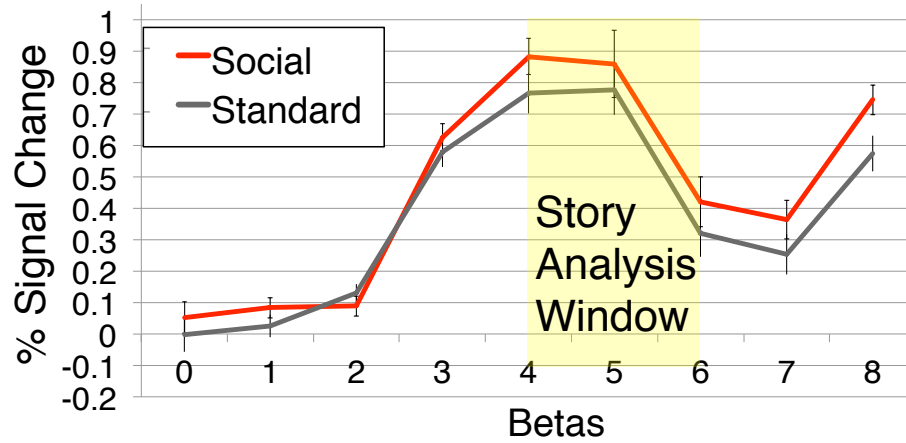
Note: Values are mean (standard deviation). Due to a technical error, task data were not recorded for one participant, resulting in N=18 for accuracy and RT measures and N=19 for post-test ratings. All post-test questionnaire ratings are based on items scored on a 1 to 5 scale. The control study had a 3s response window due to timing constraints at the behavioral session. Post-hoc pairwise comparisons were made using a Tukey's test with an alpha of .05. ***, $p < .001$. Soc=Social; Std=Standard.

Control Comparison

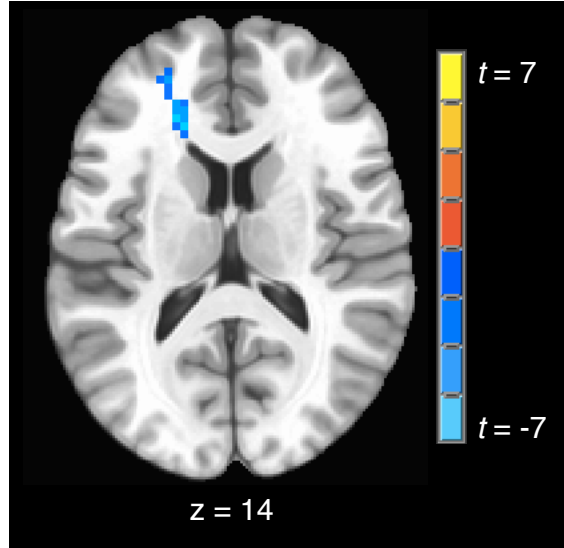
Right STG



x=65



Supplementary Figure 1. Time series plot for a selected cluster defined from group-level comparison of the two recorded speech conditions. Individual spline estimates are depicted for a region active for the Social > Standard comparison in the Story Window (Betas 4-6; voxel-wise correction of $p < .01$ and cluster correction of $p < .05$). Error bars represent standard error of the mean. s=seconds; STG=superior temporal gyrus.



Supplementary Figure 2. Whole-brain analysis examining the relation between age and the magnitude of the Live > Social contrast. Using a voxel-wise correction of $p < .01$ and cluster correction of $p < .05$, a single significant cluster was identified. This cluster demonstrated a negative relation between age and sensitivity to live speech.