Feasibility of decoding covert speech in ECoG with a Transformer trained on overt speech

Shuji Komeiji[1], Takumi Mitsuhashi[2], Yasushi Iimura[2], Hiroharu Suzuki[2], Hidenori Sugano[2], Koichi Shinoda[3], and Toshihisa Tanaka[1,*]

[1]Department of Electronic and Information Engineering, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei-shi, Tokyo, 184-8588, Japan
[2]Department of Neurosurgery, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo, 113-8421, Japan
[3]Department of Computer Science, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8550, Japan
[*]tanakat@cc.tuat.ac.jp

Comparison 1	Comparison 2	p-value (Uncorrected)	p-value (Corrected)	Effect size
Transformer (“P $\rightarrow$ P”)	BLSTM (“P $\rightarrow$ P”)	4.0×10^{-4}	8.1×10^{-4}	0.93
Transformer (“O $\rightarrow$ O”)	BLSTM (“O $\rightarrow$ O”)	3.9×10^{-4}	3.9×10^{-4}	0.51
Transformer (“C $\rightarrow$ C”)	BLSTM (“C $\rightarrow$ C”)	1.7×10^{-2}	8.5×10^{-2}	0.60
Transformer (“P $\rightarrow$ C”)	BLSTM (“P $\rightarrow$ C”)	0.5	5.0	0.03
Transformer (“O $\rightarrow$ C”)	BLSTM (“O $\rightarrow$ C”)	2.2×10^{-2}	1.3×10^{-1}	0.70
Transformer (“C $\rightarrow$ C”)	Shuffle (“C $\rightarrow$ C”)	1.0×10^{-2}	4.0×10^{-2}	1.02
Transformer (“P $\rightarrow$ C”)	Shuffle (“P $\rightarrow$ C”)	1.7×10^{-1}	1.3	0.38
Transformer (“O $\rightarrow$ C”)	Shuffle (“O $\rightarrow$ C”)	6.7×10^{-4}	2.0×10^{-3}	1.34
Transformer (“P $\rightarrow$ C”)	Transformer (“C $\rightarrow$ C”)	6.3×10^{-2}	4.4×10^{-1}	0.78
Transformer (“O $\rightarrow$ C”)	Transformer (“C $\rightarrow$ C”)	4.9×10^{-1}	4.4	0.07

Table 1. Statistics for ten comparisons in the main text. All comparisons were assessed using a one-sided Wilcoxon signed-rank test, with the null hypothesis that “Comparison 2 outperforms Comparison 1.” The p-values underwent Holm-Bonferroni correction for the ten comparisons. The table reports both uncorrected and corrected p-values, along with effect sizes calculated using Cohen’s d. In the “Comparison 1 and 2” column, acronym abbreviations (P, O, and C) represent Perception, Overt, and Covert, respectively. For example, “P $\rightarrow$ P” signifies “Perception $\rightarrow$ Perception.”

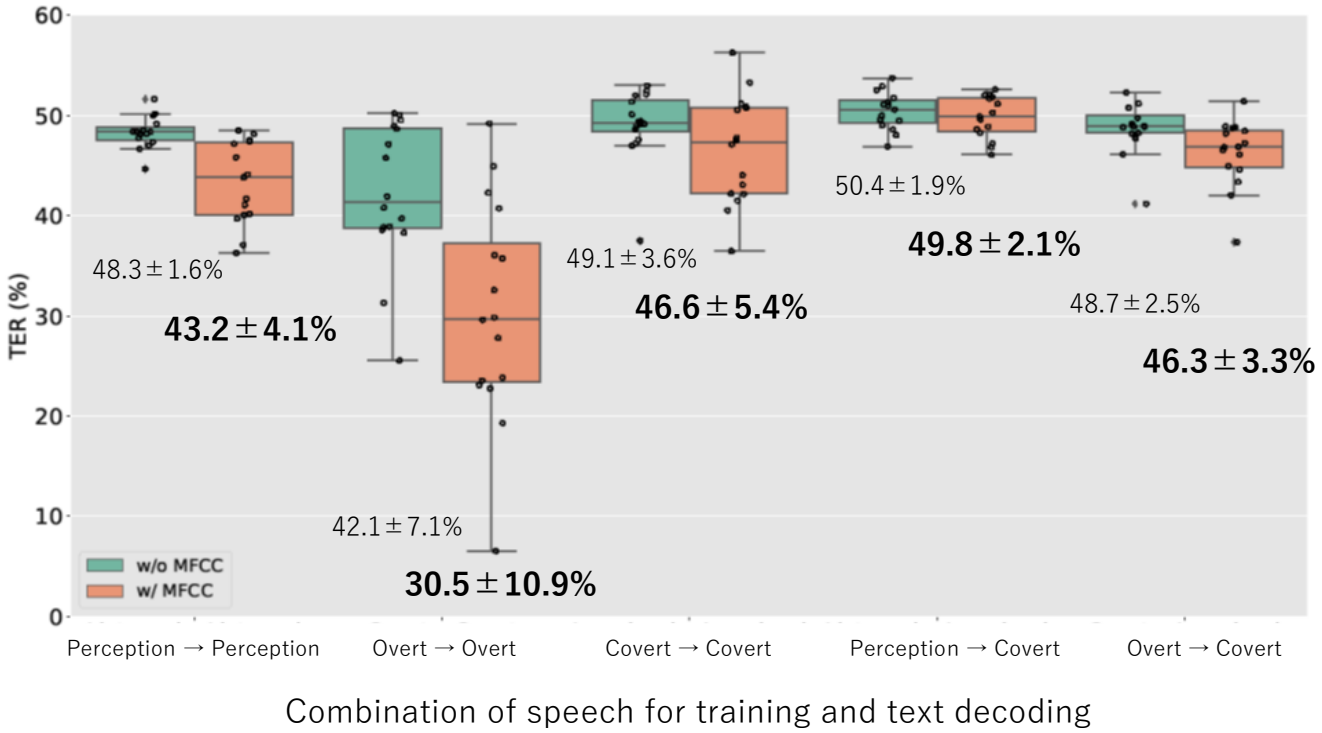

Figure 1. The decoding performance comparing use with no use of MFCCs: The figure presents a boxplot depicting the distribution of TERs for five combinations of training and decoding. Each dot on the plot represents an individual participant. The green boxes represent the TER distribution of the Transformer with no MFCCs, whereas the orange ones represent the Transformer with MFCCs. The numbers in the figure indicate the average ± standard deviation.

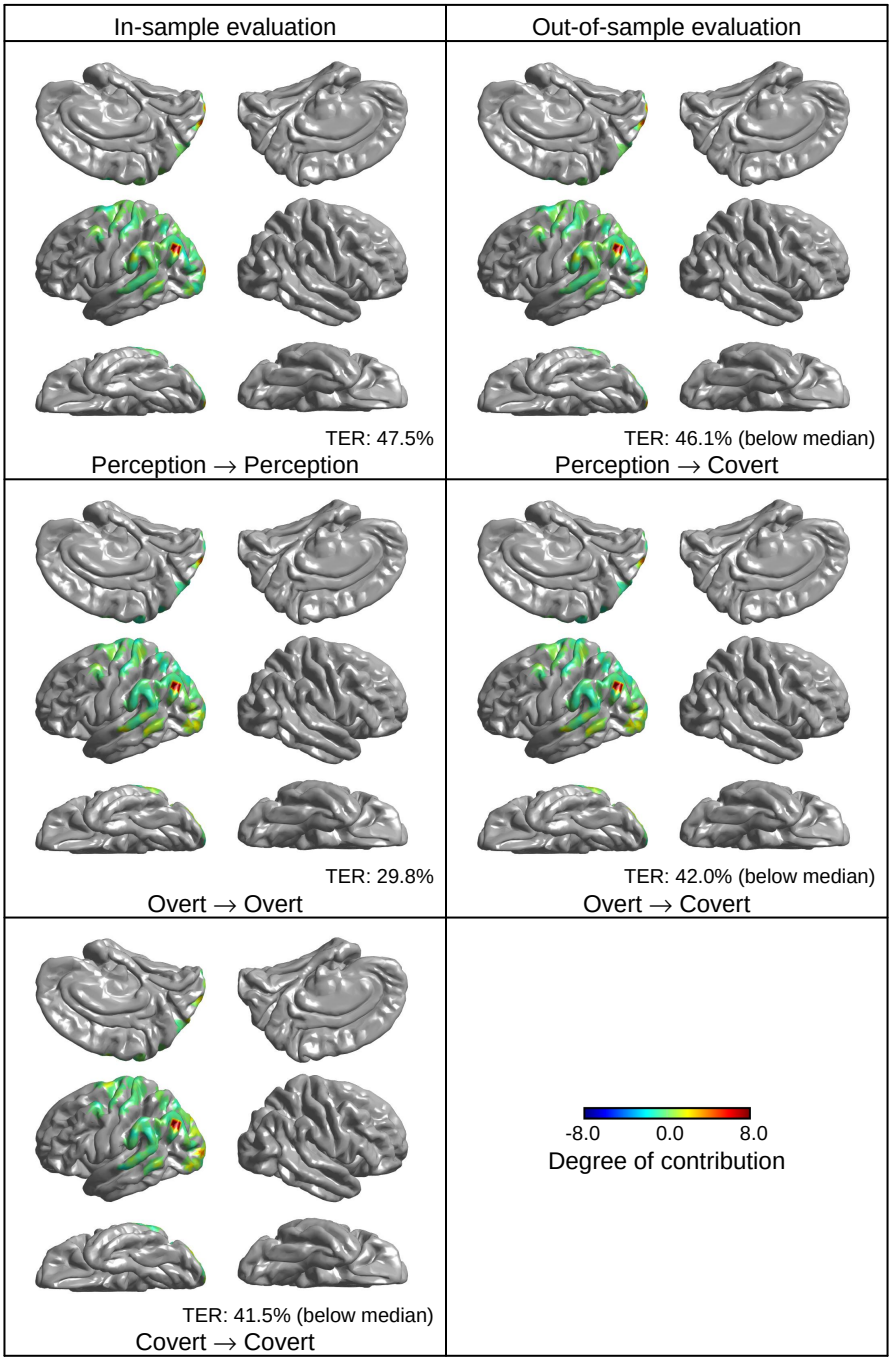

Figure 2. Electrode contributions for js1. TERs for js1 in "Covert → Covert," "Perception → Covert, " and "Overt → Covert" combinations are below the median among all participants. The inferior parietal cortex emerges as the most significant contributing area for all combinations.

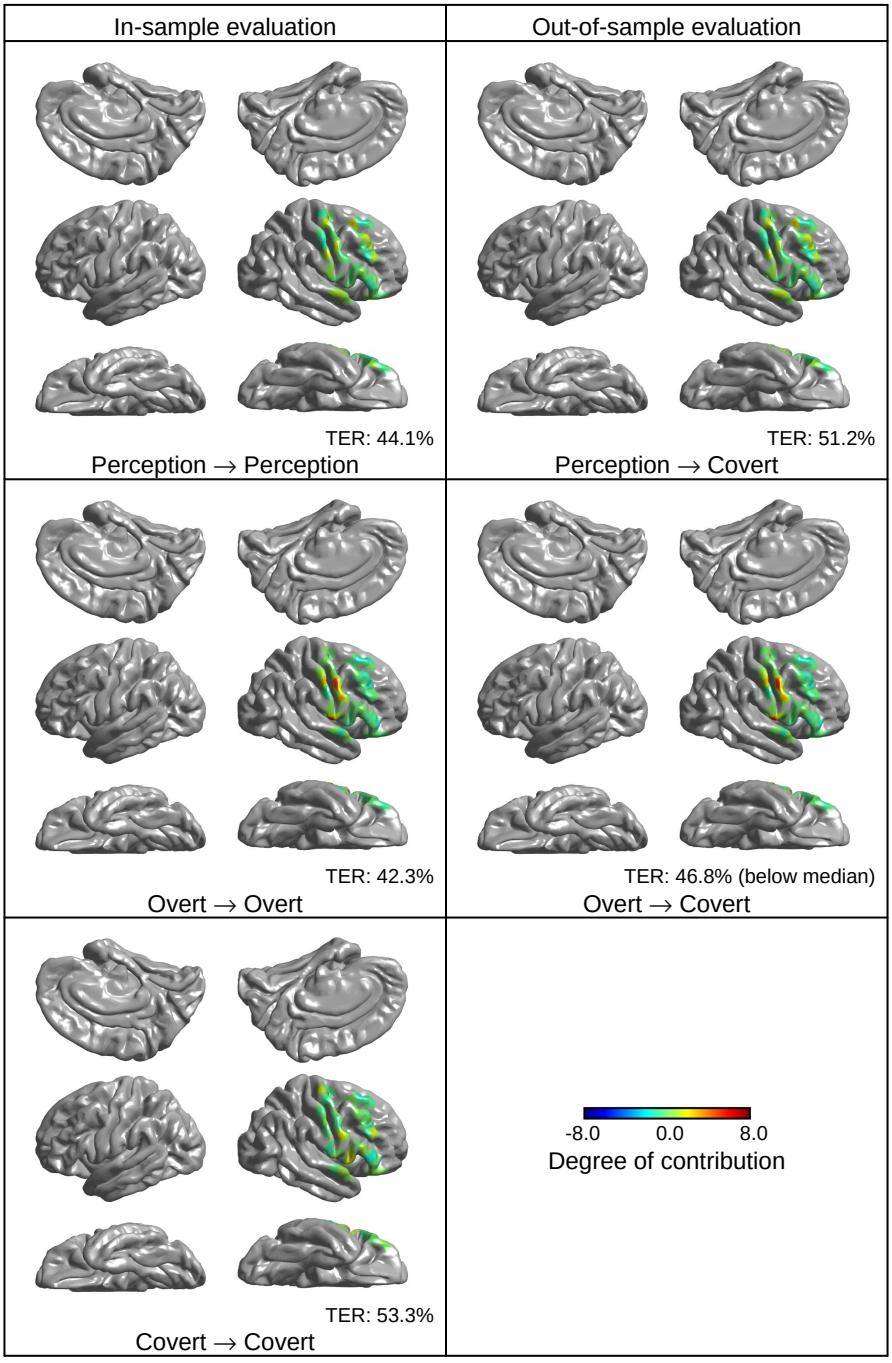

Figure 3. Electrode contributions for js2. Only the TER for js2 in the "Overt → Covert" combination is below the median among all participants. The SMC in the right hemisphere emerges as the most significant contributing area for this combination.

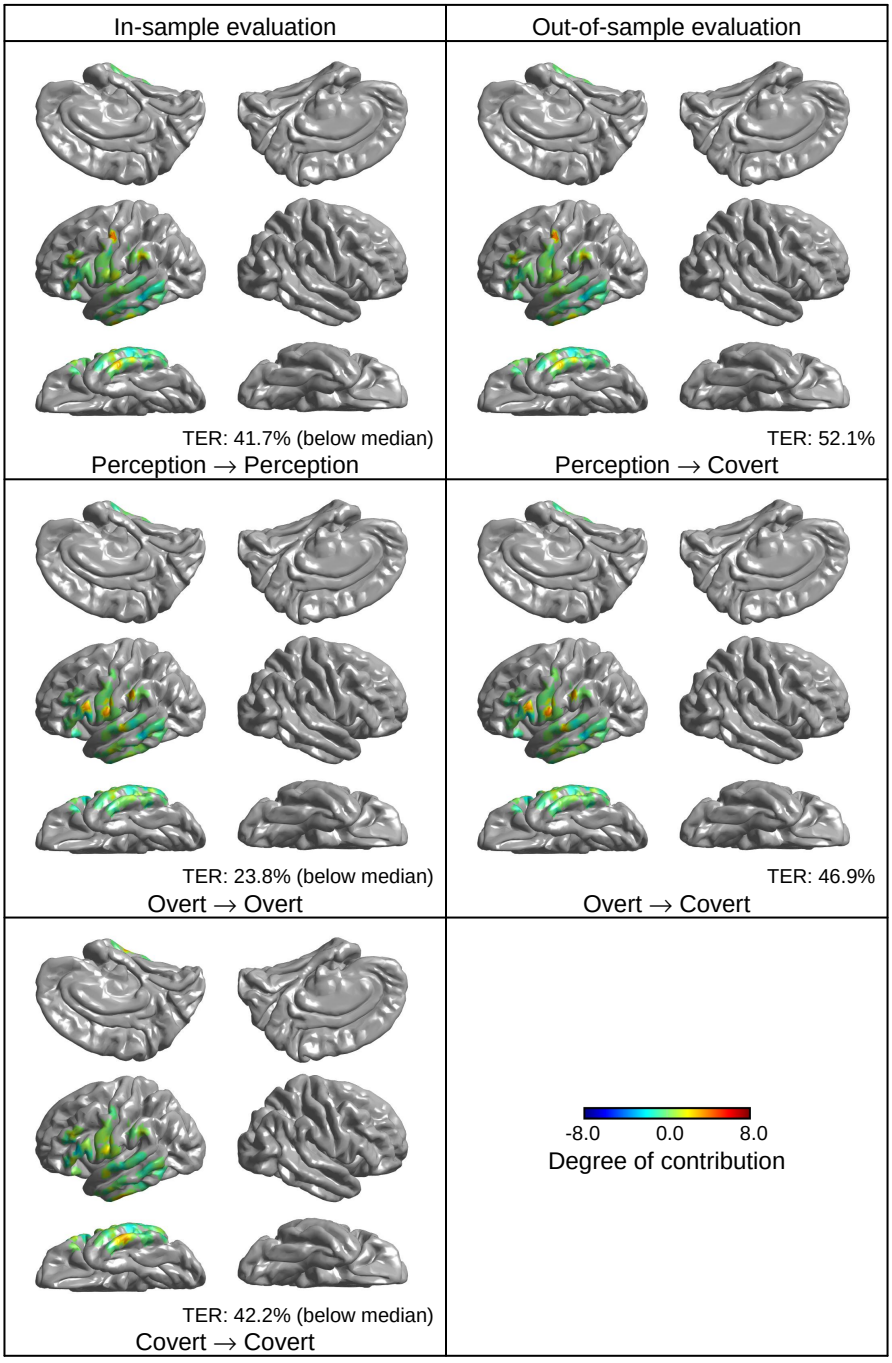

Figure 4. Electrode contributions for js3. TERs for js3 in "Perception → Perception," "Overt → Overt," and "Covert → Covert" combinations are below the median among all participants. The SMC for "Perception → Perception," and Broca's area, the SMC, and Wernicke's area for "Overt → Overt," emerge as significant contributing areas. The inferior temporal gyrus for "Covert → Covert" is a slightly contributing area.

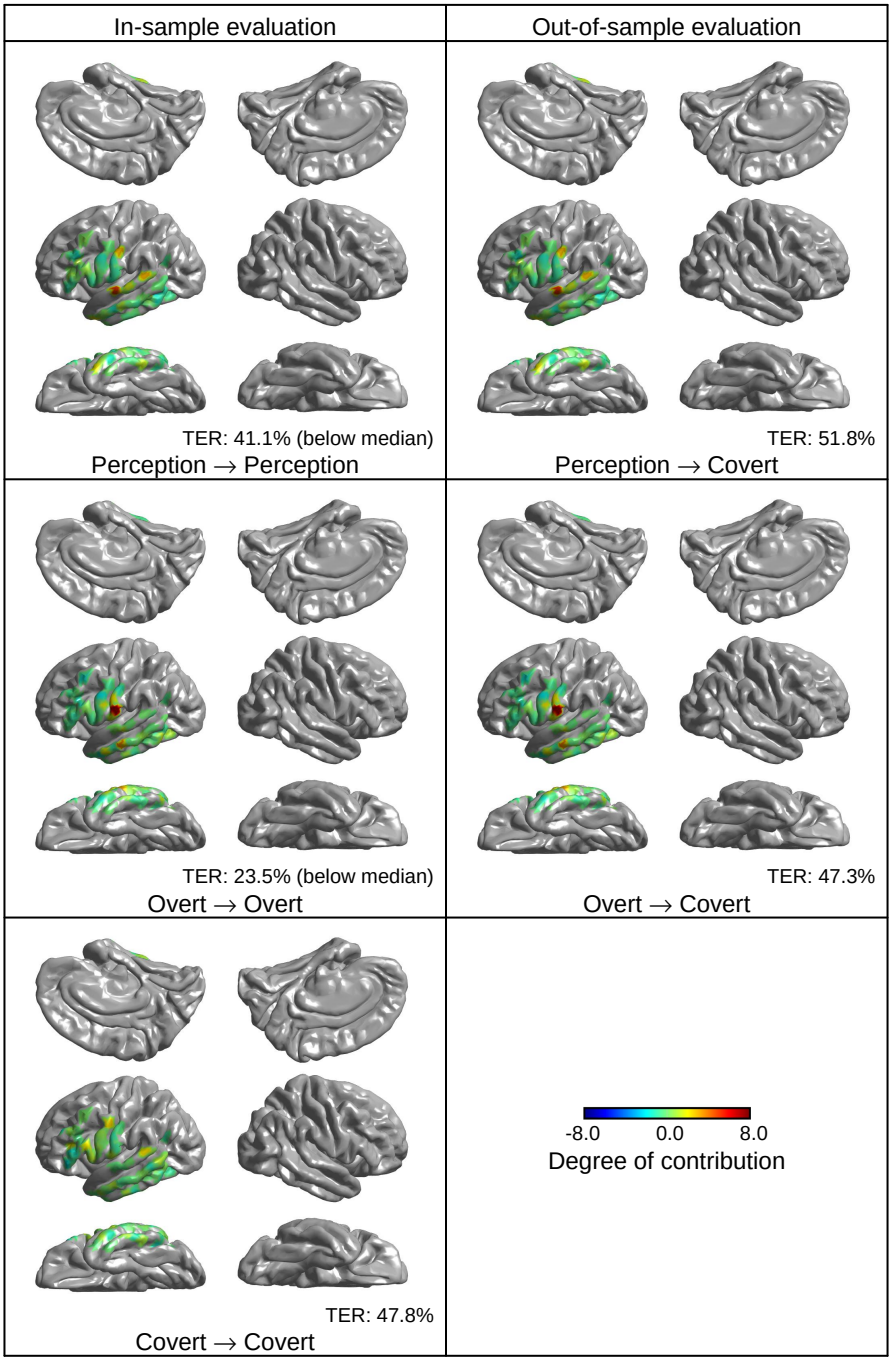

Figure 5. Electrode contributions for js4. TERs for js4 in "Perception → Perception" and "Overt → Overt" combinations are below the median among all participants. The STG for "Perception → Perception" and the SMC for "Overt → Overt" emerge as the most significant contributing areas.

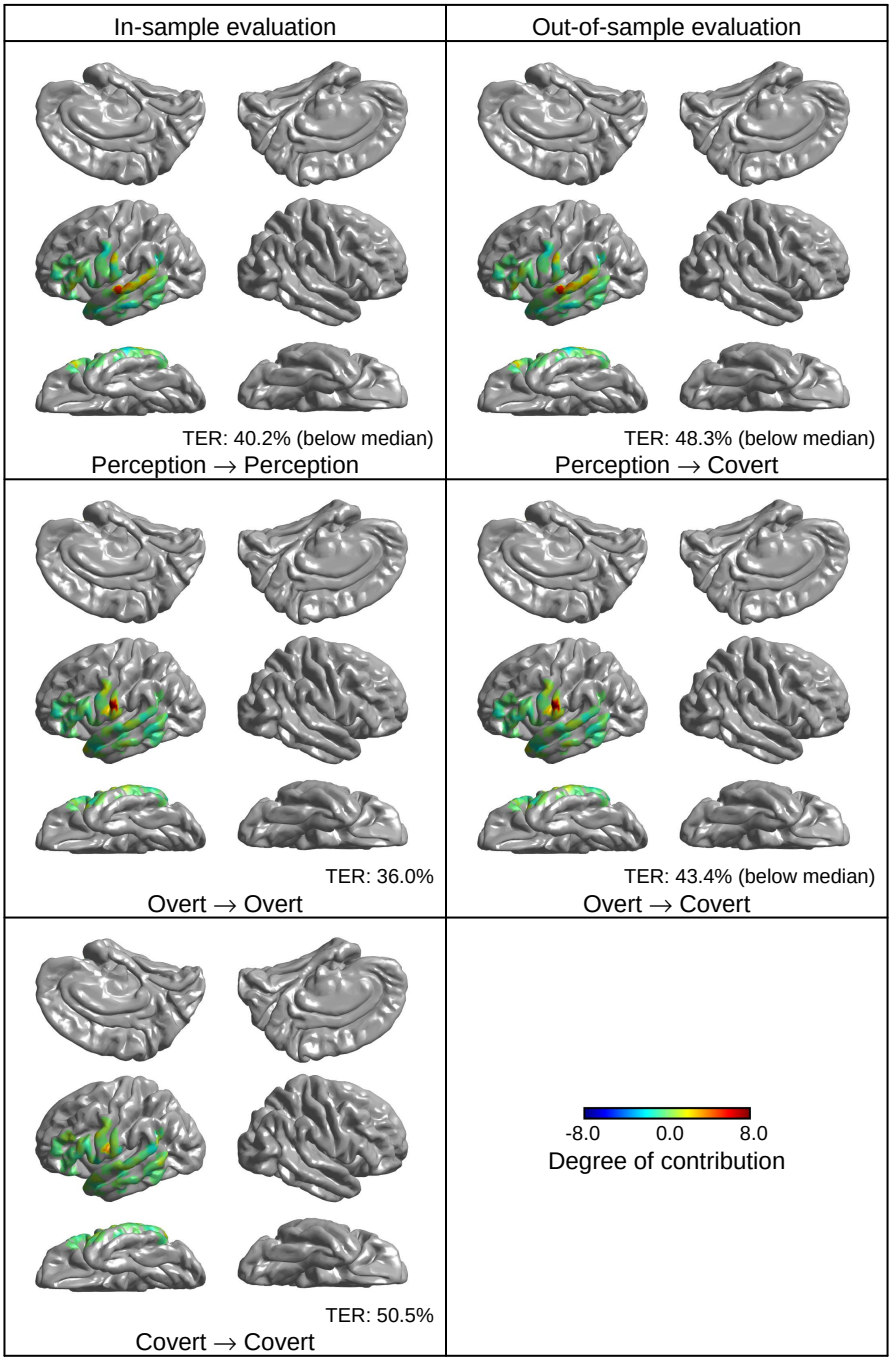

Figure 6. Electrode contributions for js5. TERs for js5 in "Perception → Perception," "Perception → Covert," and "Overt → Covert" combinations are below the median among all participants. The STG for "Perception → Perception" and "Perception → Covert," and the SMC for "Overt → Covert" emerge as the most significant contributing areas.

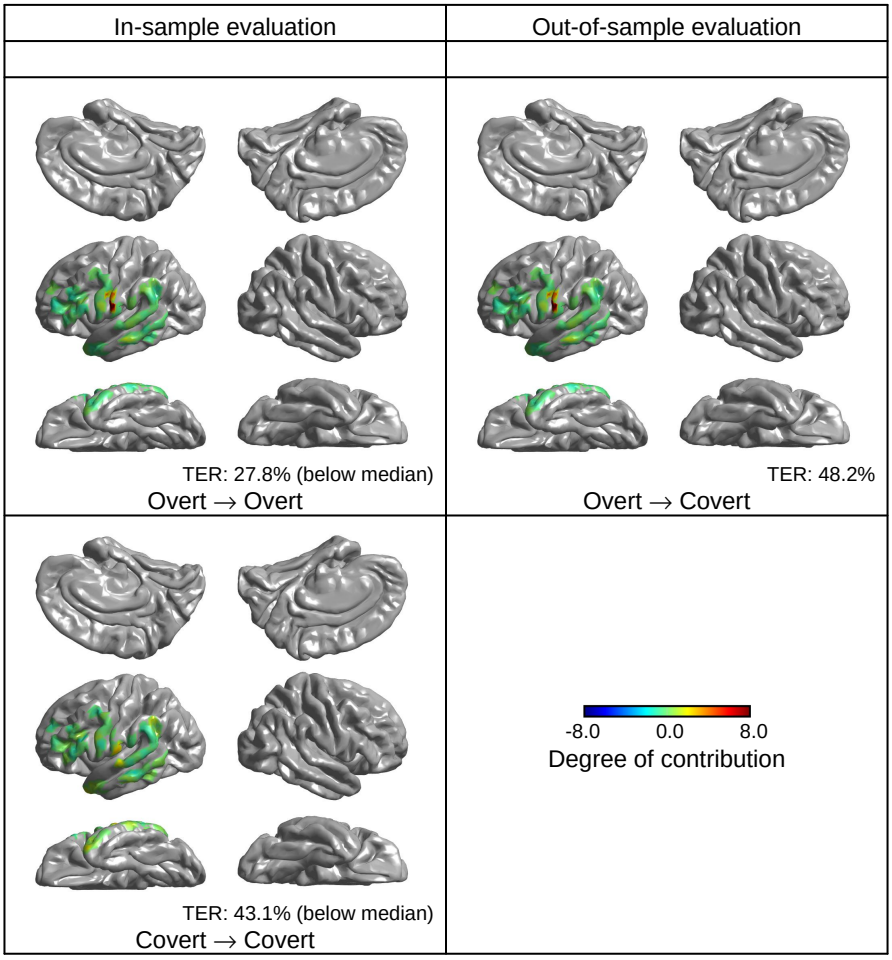

Figure 7. Electrode contributions for js6. TERs for js6 in "Overt → Overt" and "Covert → Covert" combinations are below the median among all participants. The SMC for "Overt → Overt" emerges as the most significant contributing area. No significant contribution area was found for "Covert → Covert."

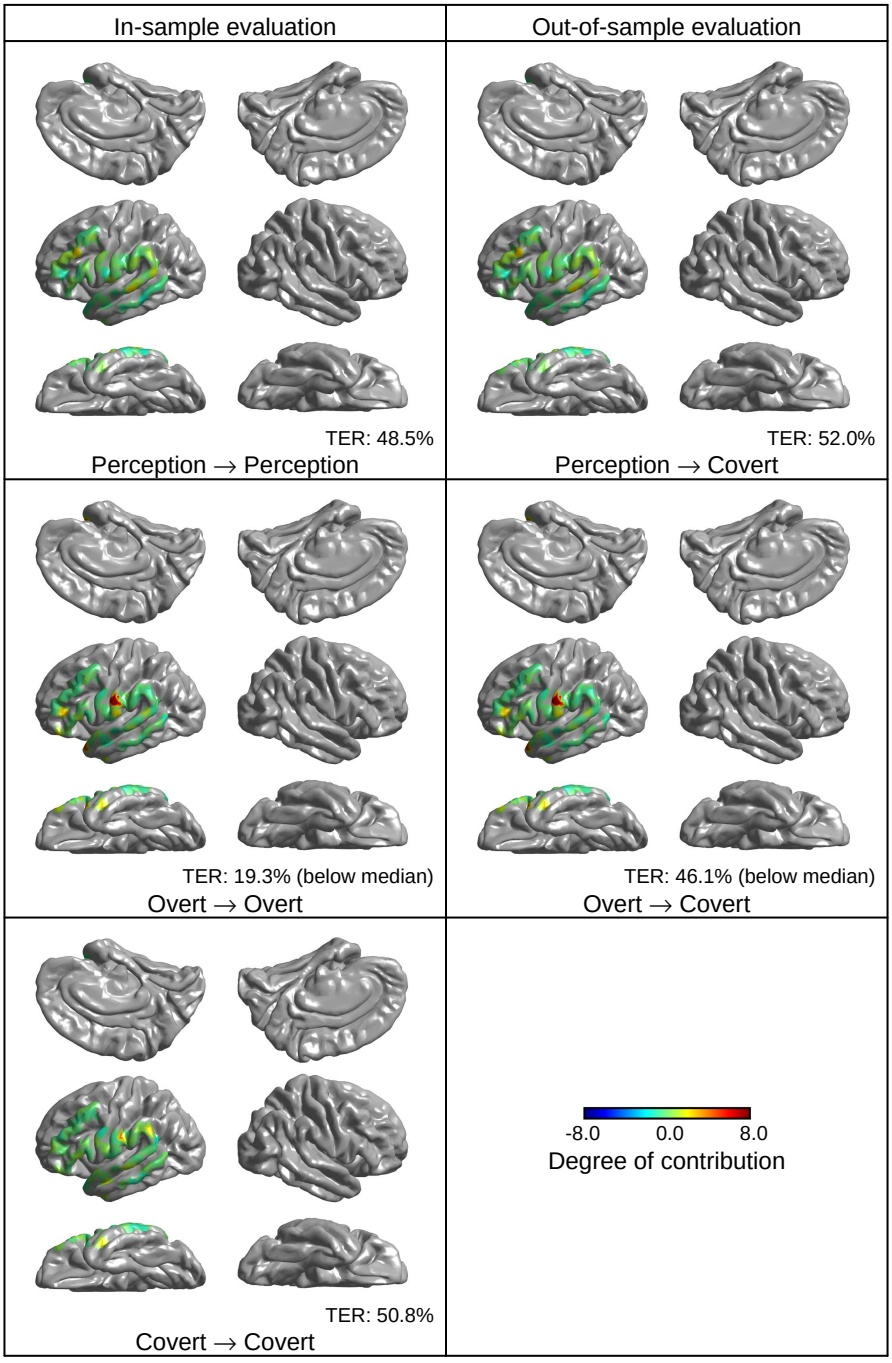

Figure 8. Electrode contributions for js7. TERs for js7 in "Overt → Overt" and "Overt → Covert" combinations are below the median among all participants. The SMC for these combinations emerges as the most significant contributing area.

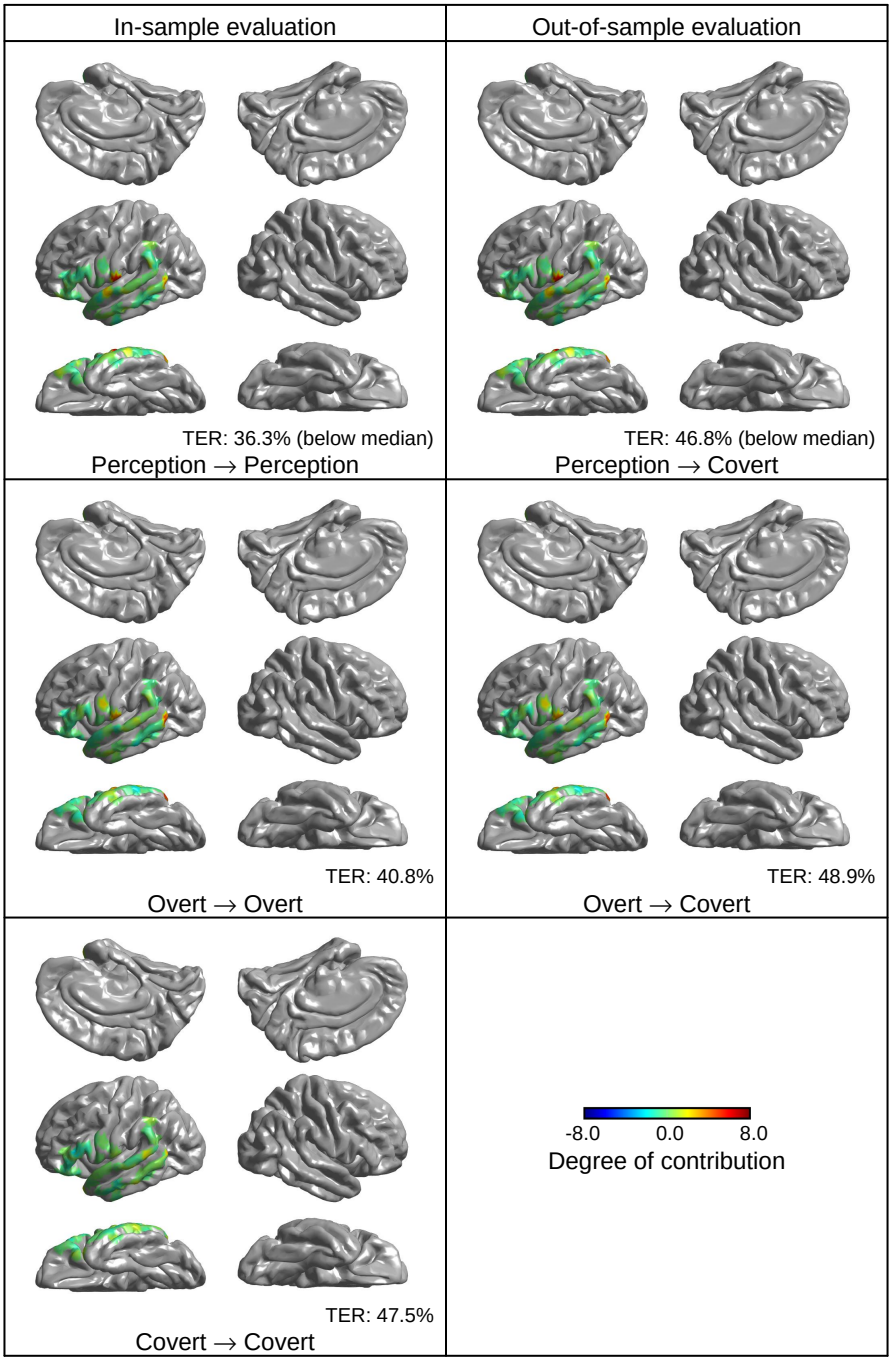

Figure 9. Electrode contributions for js8. TERs for js8 in "Perception → Perception" and "Perception → Covert" combinations are below the median among all participants. The STG for these combinations emerges as the most significant contributing area.

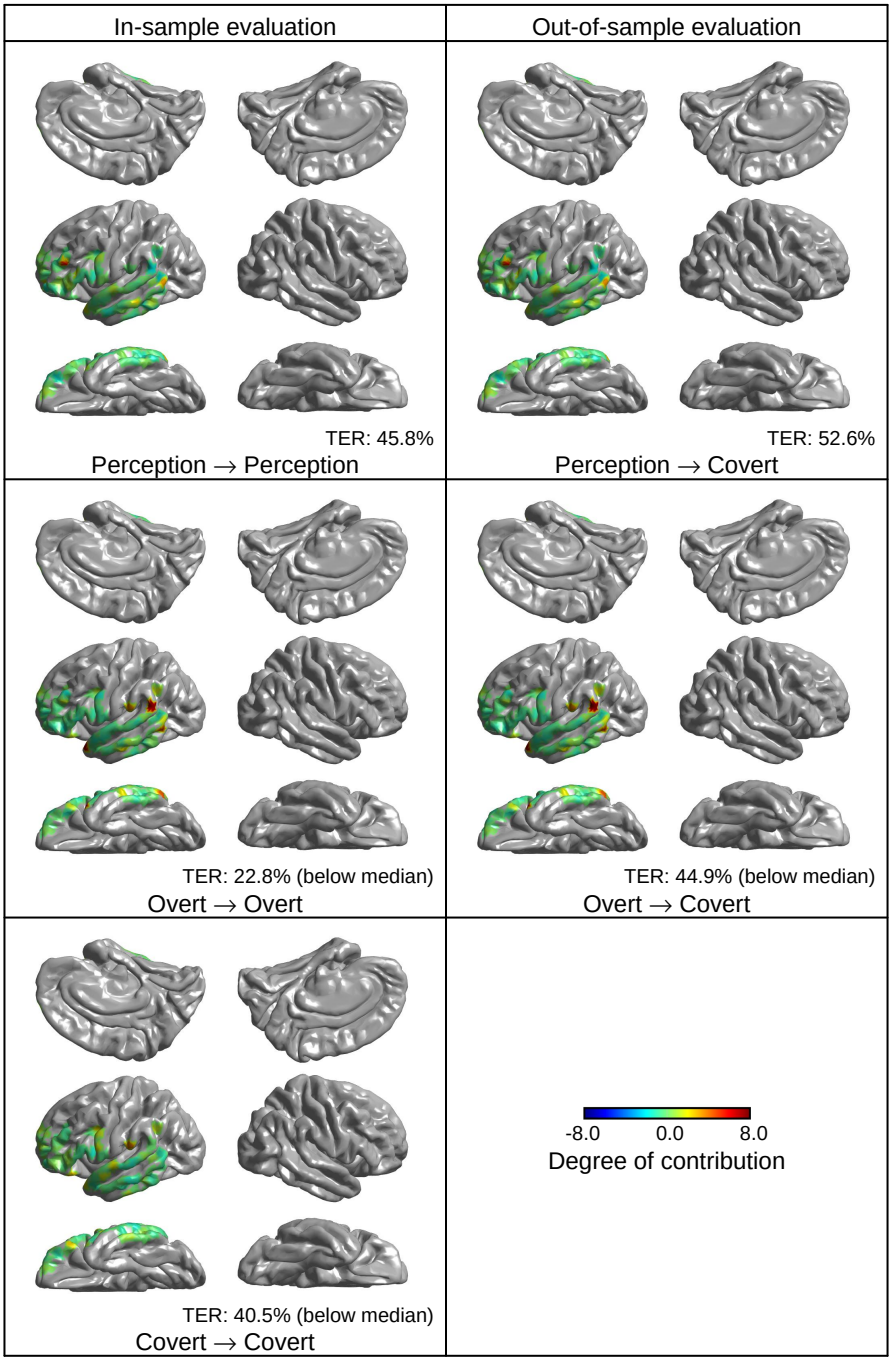

Figure 10. Electrode contributions for js9. TERs for js9 in "Overt → Overt," "Covert → Covert," and "Overt → Covert" combinations are below the median among all participants. Wernicke's area for these combinations emerges as the most significant contributing area.

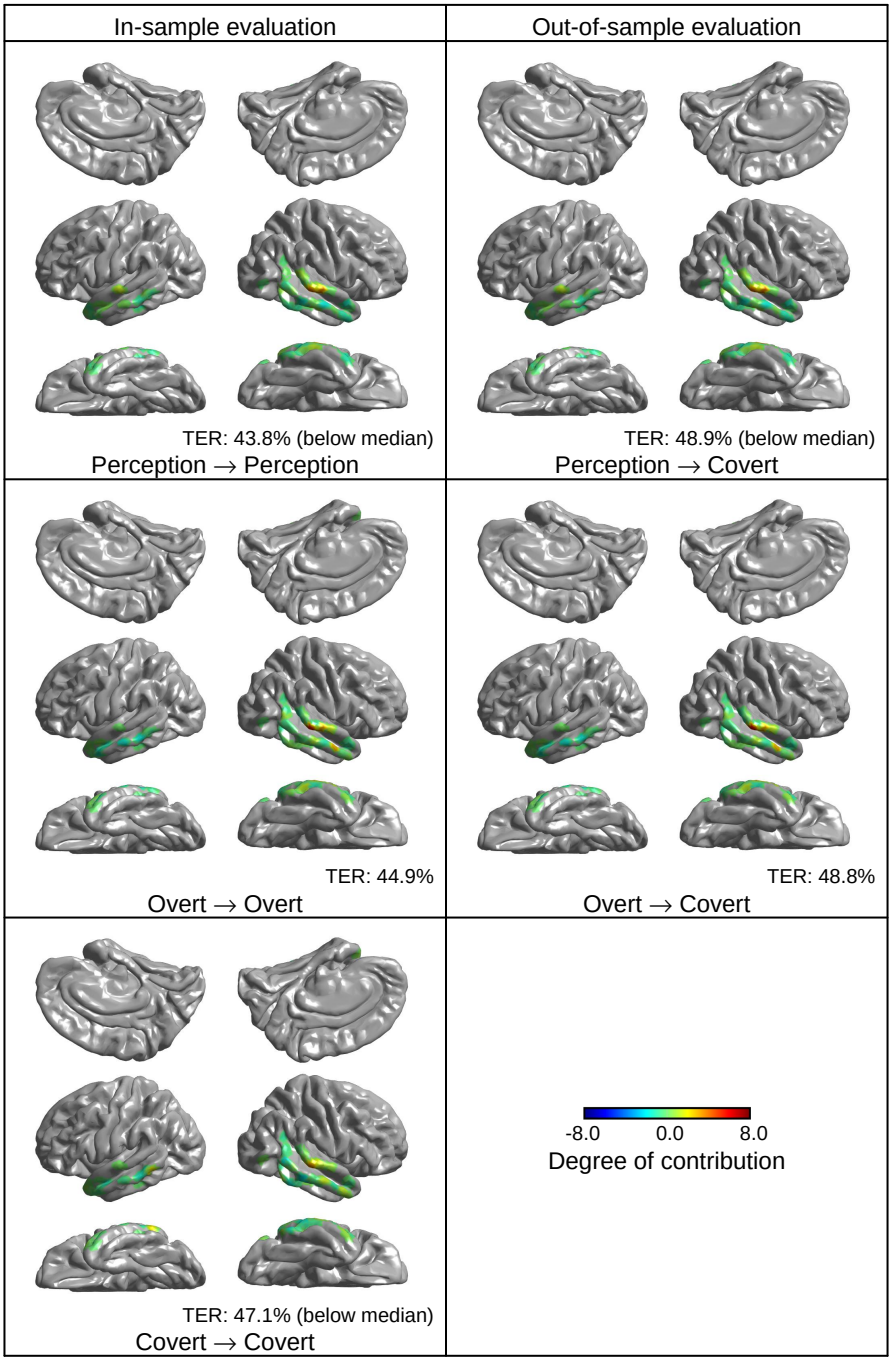

Figure 11. Electrode contributions for js10. TERs for js10 in "Perception → Perception," "Covert → Covert," and "Perception → Covert" combinations are below the median among all participants. The STG for these combinations emerges as the most significant contributing area.

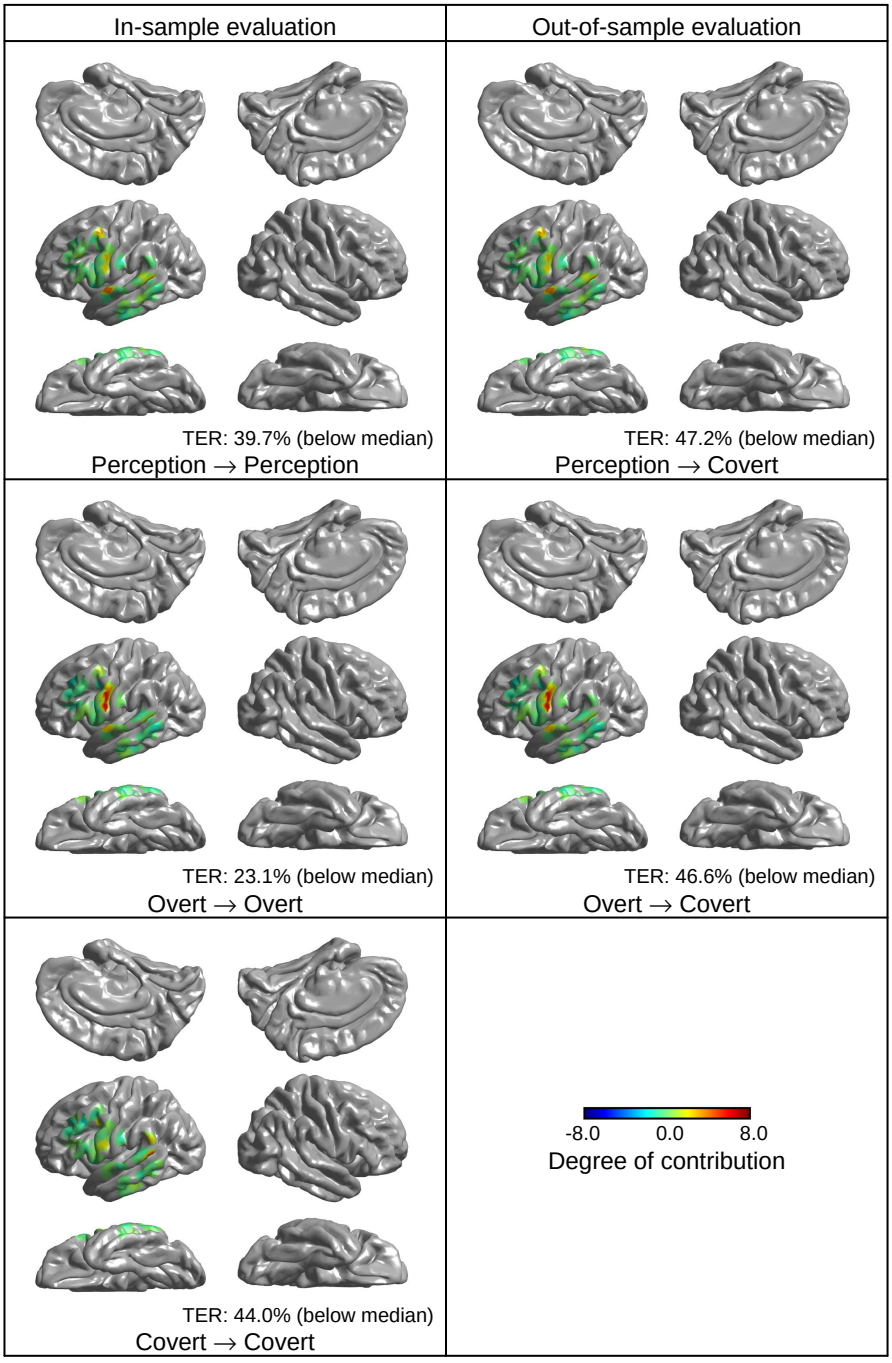

Figure 12. Electrode contributions for js11. TERs for js11 in all five combinations are below the median among all participants. STG for "Perception → Perception," "Perception → Covert," and "Covert → Covert" and the SMC for both "Overt → Overt" and "Overt → Covert" emerges as the most significant contributing area.

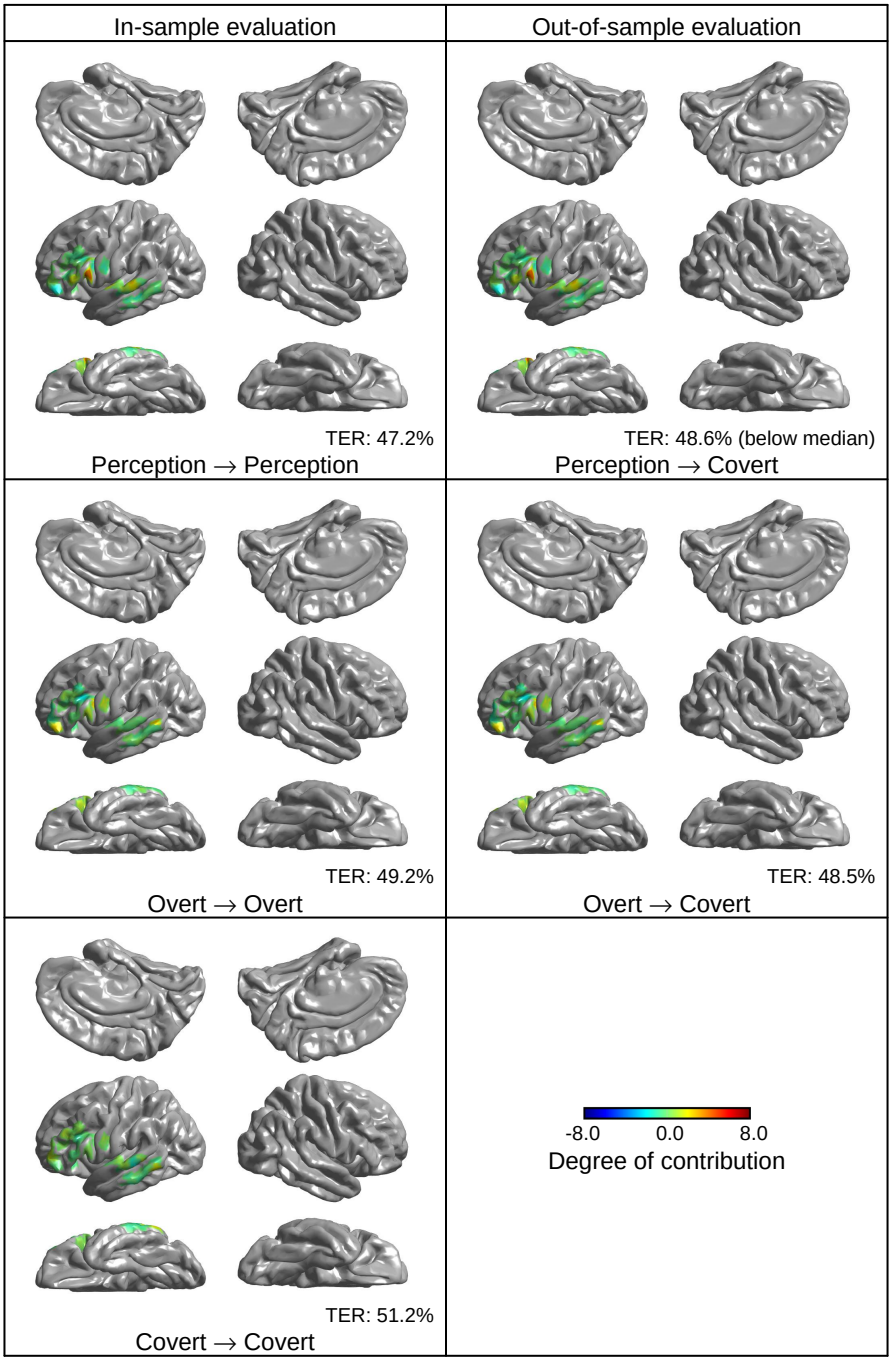

Figure 13. Electrode contributions for js12. Only the TER for js12 in the "Perception → Covert" combination is below the median among all participants. Broca's area for this combination emerges as the most significant contributing area.

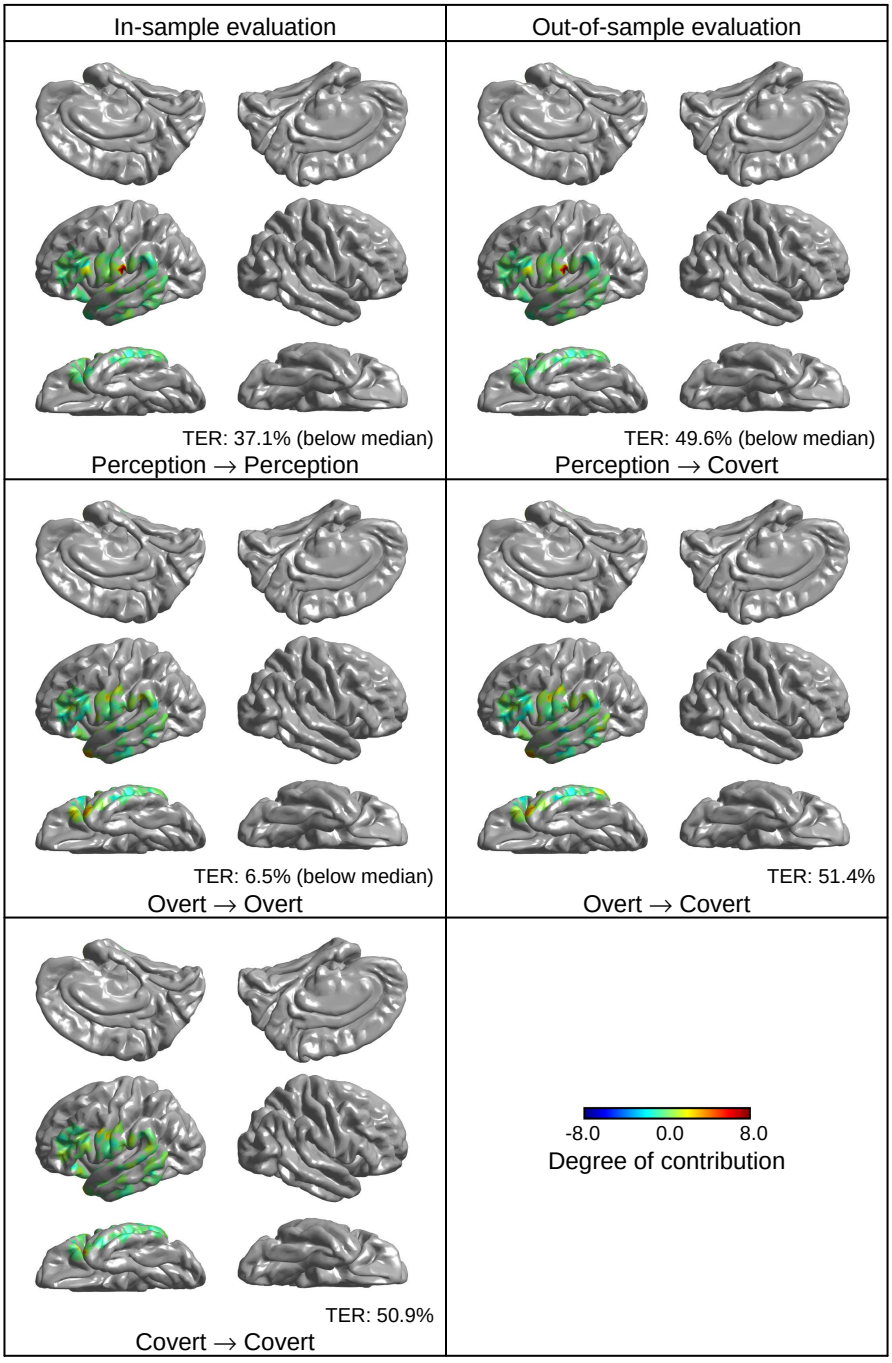

Figure 14. Electrode contributions for js13. TERs for js13 in the "Perception → Perception," "Overt → Overt," and "Perception → Covert" combinations are below the median among all participants. The SMC for both "Perception → Perception" and "Perception → Covert," the inferior temporal gyrus for "Covert → Covert" emerge as the most significant contributing area.

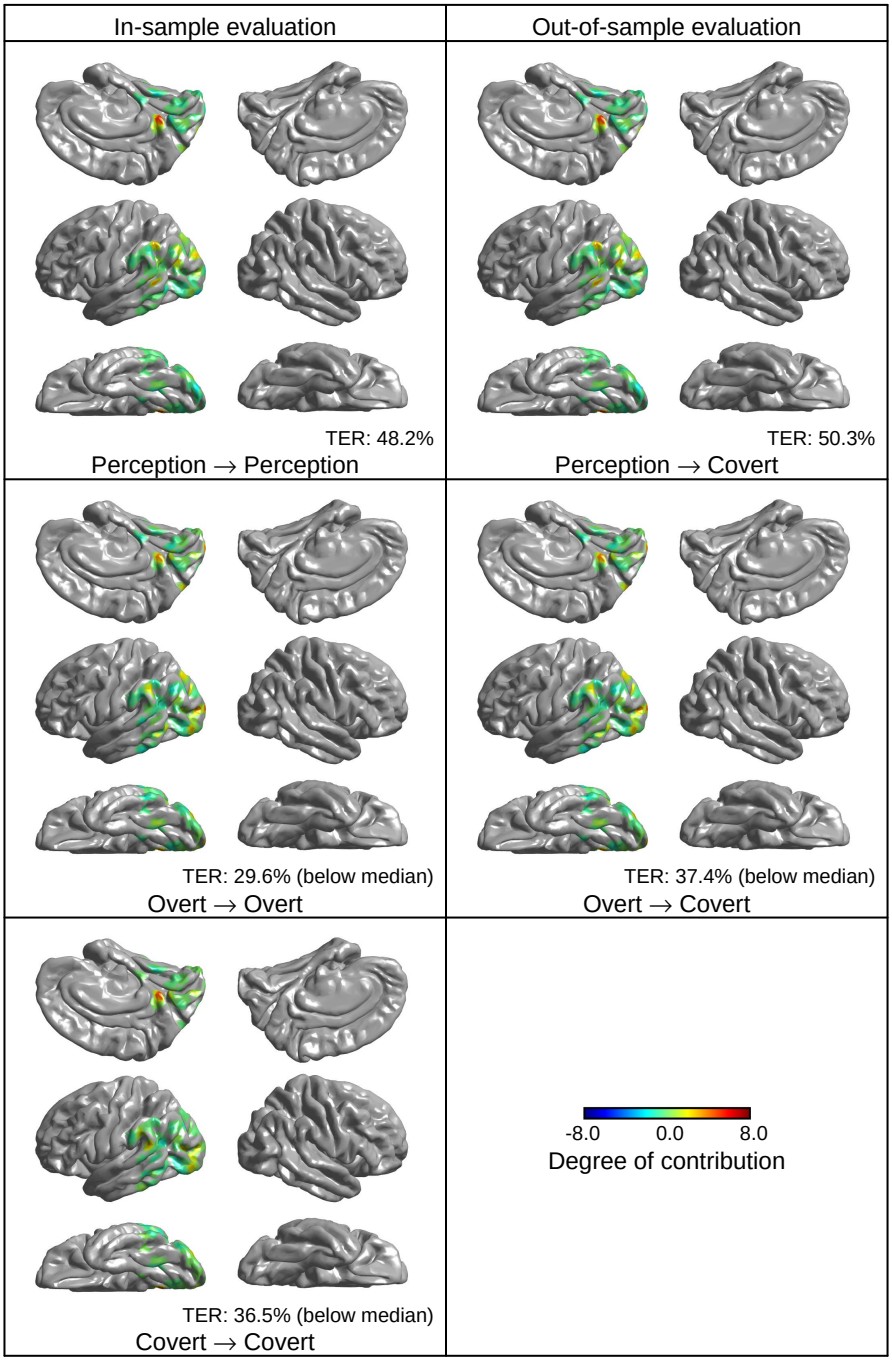

Figure 15. Electrode contributions for js14. TERs for js14 in the "Overt → Overt," "Covert → Covert," and "Overt → Covert" combinations are below the median among all participants. The occipital lobe for these combinations emerges as the most significant contributing area.

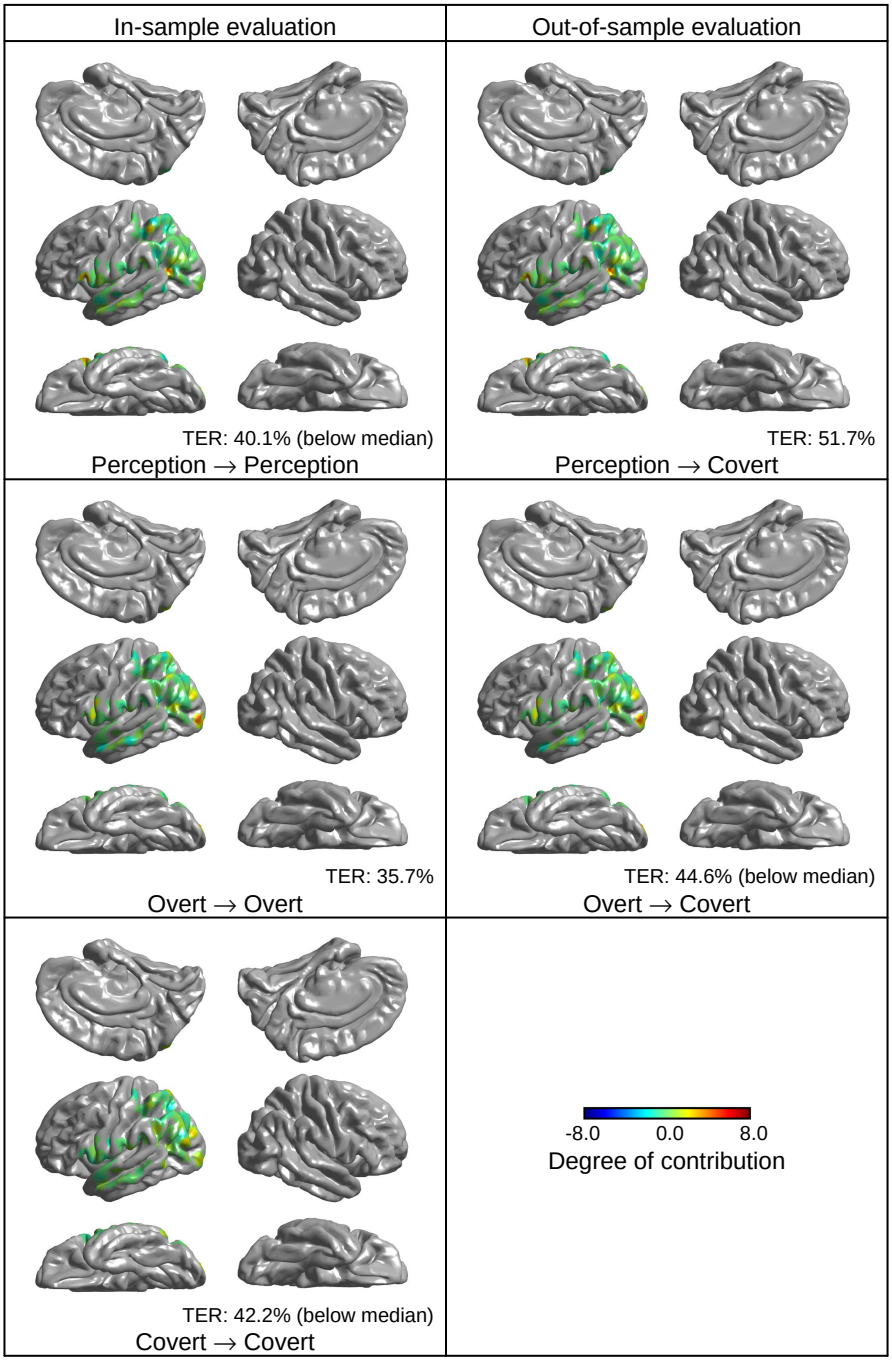

Figure 16. Electrode contributions for js15. TERs for js15 in the "Perception → Perception," "Covert → Covert," and "Overt → Covert" combinations are below the median among all participants. Wernicke's area for "Perception → Perception" and the occipital lobe for "Overt → Covert" emerge as the most significant contributing area. No significant contribution area was found for "Covert → Covert."

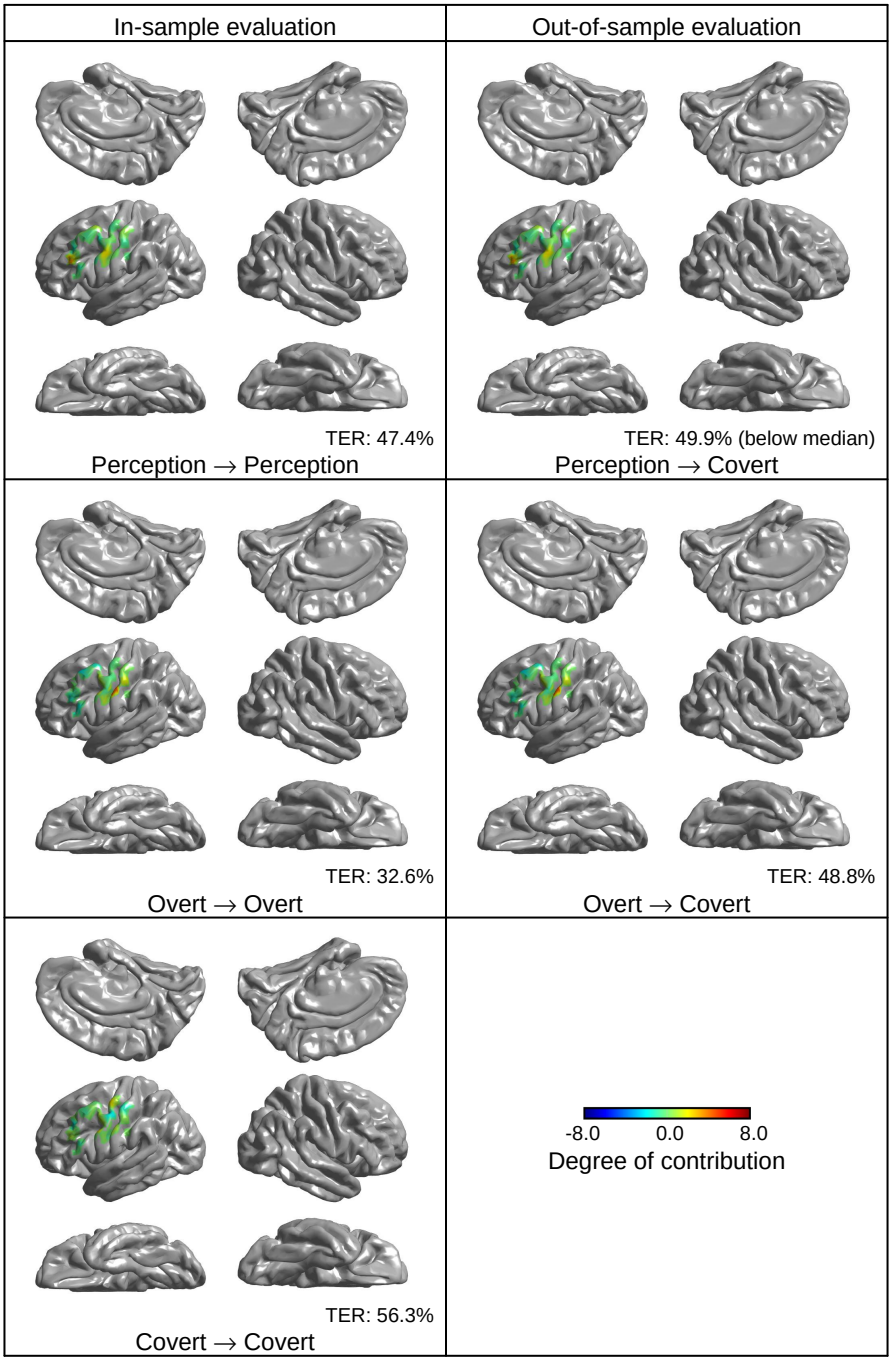

Figure 17. Electrode contributions for js16. Only the TER for js16 in the "Perception → Covert" combinations is below the median among all participants. No significant contribution area was found for this combination.