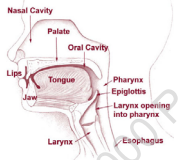
Brain structures differentially responsible for controlling overt and covert speech

Simon A. Overduin[1] & **Frank H. Guenther**[1,2,3,4] ([1]Department of Cognitive and Neural Systems, Boston University; [2]Division of Health Sciences & Technology, Harvard-MIT; [3]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital; [4]Sargent College of Health and Rehabilitation Sciences, Boston University)

82.3

Introduction

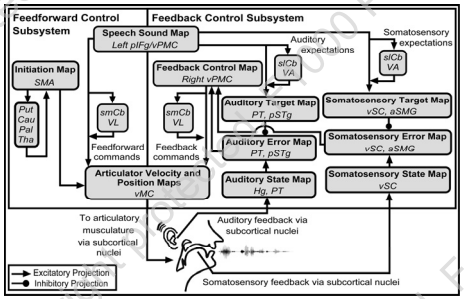


Overt speech requires coordination of three groups of muscles: those for respiration, phonation, and articulation. Neuroimaging has been used to probe the mechanisms underlying this control. The production of **covert** (or "inner") speech has received less attention, even though it often substitutes for overt speech in cognitive tasks as it avoids movement artifacts (e.g. [1]).

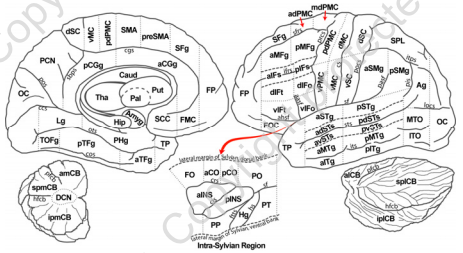
Multiple imaging studies have included both overt and covert speech **articulation** [2-10], with inconsistent results. Together they suggest that many brain areas are more active in overt speech, including somatosensory, primary motor, premotor, supplementary motor, cingulate, auditory, and insular cortex, as well as thalamic, striatal, and cerebellar areas. Few if any are more active in covert speech.

These studies did not control for differences unrelated to articulation. No attempts were made to control subjects' **respiration**, even though breathing-related BOLD variations occur in brain regions with high blood volume, including gray matter [11]. Other studies have found that active breathing control activates all the same regions as the "overt speech" network listed above (e.g. [12, 13]). Overt and covert speech also differed in terms of **voicing**, which may modulate regions including premotor and temporal cortex, basal ganglia, thalamocortical circuitry, and cerebellum [14]. Overt and covert speech differed, too, in sensory **feedback**.

Here, we minimize these confounds by comparing covert speech to unvoiced, feedback-masked overt speech with the same breathing pattern. Speech was whispered to reduce bone conduction as well as involvement of voicing muscles [15], and was acoustically masked with speech-shaped noise [16, 17]. Respiration was monitored using a pneumatic belt [11].



DIVA (Directions into Velocities of Articulators) is a neural network model integrating theories and data on speech acquisition and production [18-20]. Its predictions:



Normal Overt Speech vs. Baseline: The basic feedforward network including the speech sound map (left plFg/vPMC), articulatory velocity and position maps (vMC), auditory and somatosensory state and target maps (Hg/PT/pSTg, vSC/aSMg), subcortical loops (cerebellar smCb and slCb, VL/VA thalamic nuclei), and initiation map (SMA) will be active in overt speech, even if whispered rather than voiced.

Normal vs. Masked Overt Speech: Speech with vs. without feedback will more strongly engage auditory state and error cells in superior temporal cortex (Hg/PT/pSTg). With audition masked, there will be more feedback control activity (right vPMC, smCb, and vMC) *if* subjects make any behavioral correction. Absent such correction, masked production will not differentially engage any area.

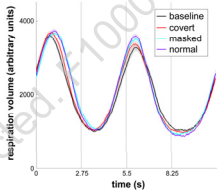
Masked Overt vs. Covert Speech: Overt speech will differentially engage SMA initiation cells, vMC, vSC, and the feedback circuit (right vPMC, smCb and vMC) in the case of a correction. Covert speech will not differentially engage any area.

Methods

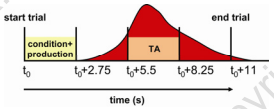
Subjects 16 normal right-handed English speakers (8 male, av. 24 years old).

Task Subjects read tri-syllabic pseudowords (e.g. "ba·de·gu") silently (***Covert***) or in a whispered voice, with feedback audible (***Normal***) or ***Masked***. Conditions plus a silent ***Baseline*** ("**·**·**·**") were pseudorandomly ordered. Stimuli were presented visually with the condition cue (color), throughout the 2.75-s production period. Subjects completed 9 runs each with 32 trials.

Respiration Breathing was measured by a pneumatic belt (Siemens). During 1 or 2 practice runs, subjects were trained to regulate the depth and timing of their breaths in each condition.



Audio Subjects' speech spectra were used to create speech-shaped noise, played through headphones (Stax) during each run. Volume was adjusted so they could not hear their whisper except when overlaid on noise, on a random half of trials.



fMRI Data were collected with a Siemens 3T Trio scanner (TR 2.75 s, 45 slices, 3-mm^3 resolution). Data were sampled sparsely, limiting activation induced by scanner noise [21] and speech movement artifacts [22].

Analysis Volumes were realigned, normalized in stereotactic space, smoothed (8-mm FWHM Gaussian) and analyzed for random effects using SPM2 and MATLAB. Results are given as normalized effect sizes ($p < 0.05$, FDR corrected).

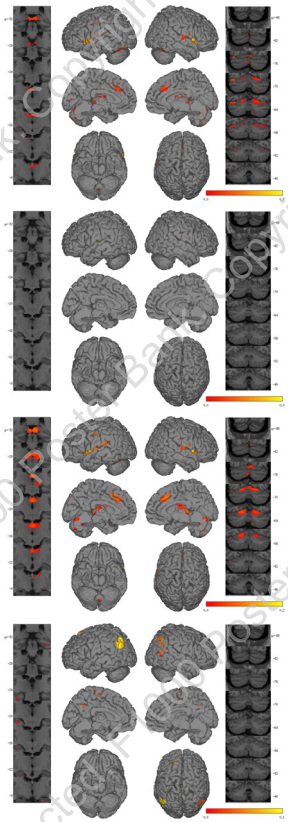
Results

Normal Overt Speech vs. Baseline
In unmasked whispered speech, most of the basic feedforward network was active: L (and R) lFg/vPMC, vMC, pSTg/Hg/PT (R), vSC, aSMg (L), smCb, slCb, and VL/VA. Additional responses were observed in L occipital cortex, anterior cingulate, R anterior insula, R inferior cerebellum, and vermis. PreSMA but *not* SMA was active.

Normal vs. Masked Overt Speech
In normal vs. auditory-masked overt speech there was no difference besides a trend for bilateral pSTG/Hg to be more active when speech was heard ($p < 0.001$ uncorrected). In masked vs. unmasked speech no areas were more active. Behaviorally, no subject showed a difference between mean or peak acoustic RMS values (all $p > 0.08$).

Masked Overt vs. Covert Speech
In auditory-masked overt vs. covert speech, much of the basic feedforward network was active: vPMC, vMC, pSTg/Hg/PT (R), vSC, aSMg, smCb, slCb, VL/VA. So were most of other areas noted above (L occipital, anterior cingulate, R anterior insula, vermis, preSMA). The strongest activations were those in smCb and R anterior insula.

In Covert vs. Masked Overt speech, there were several areas strongly active in covert production, in contrast to the prediction of none. The largest BOLD responses were SMA (with pdPMC/dMC/dSC) and angular gyrus (Ag/OC) bilaterally. Additional activity was observed in L precuneus (PCN), L anterior dorsal PMC and SFg, R posterior MTg, and para- and hippocampal regions.



Discussion

Whispered and voiced speech activate the same feedforward network.
The additional activations seen in occipital cortex may have been due to reading, while those in anterior cingulate, anterior insula and inferior cerebellum have each been related to the detection of sensory errors in speech [16,20,23].

Tonic auditory masking of speech does not require a corrective response.
Subjects did not measurably modify their speech online when it was masked, and no area implicated in feedback correction was active. Another study [16] found STG more active with masked speech, opposite to the pattern found here. However, there the masking noise was turned on only during overt production, potentially causing a greater auditory cortical response than the tonic noise used here. DIVA's inhibitory projections from auditory target to error cells may explain why no error signal is observed in the absence of significant acoustic input.

The entire feedforward network may be de-activated in covert speech.
The above analyses validated unvoiced, masked speech and allowed us to do a controlled comparison with covert articulation. If there is a structure gating overt production, it may act earlier than motor cortex, since other areas downstream of the putative speech sound map were inactive in covert speech. How inner speech can be "heard" without auditory target activation remains to be explored.

Rather than initiate overt speech, SMA may inhibit it during covert speech.
No single area was more active in overt than covert speech. Speech initiation may be reflected in a network of areas including the cerebellum [24] and anterior insula [25]. SMA's critical role in speech initiation [26], evident from neurological studies [27], may actually be one of motor inhibition [28]. Other areas more active in covert speech require further study, including the postcentral gyrus, precuneus, and MTg (previously found to be more active in covert speech [10]), and the angular gyrus (traditionally thought to convert written text into inner speech [29]).

Acknowledgments

- Jon Brumberg, Shanqing Cai, Satrajit Ghosh, Elisa Golfinopoulos, Carrie Niziolek, Jason Tourville
- Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, MIT
- NIH R01 DC-007683, NIH R01 DC-002852 (PI: Frank Guenther)

References

1. Hinke RM et al. (1993) *Neuroreport* 4, 675-8.
2. Hirano S et al. (1996) *Neuroreport* 8: 363-7.
3. Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W (2000) *Neuroreport* 11: 1997-2000.
4. Rosen HJ, Ojemann JG, Ollinger JM, Petersen SE (2000) *Brain Cogn* 42: 201-17.
5. Huang J, Carr TH, Cao Y (2001) *Hum Brain Mapp* 15: 39-53.
6. Palmer ED et al. (2001) *NeuroImage* 14: 182-93.
7. Sakurai Y et al. (2001) *Cogn Brain Res* 12: 161-5.
8. Stippich C, Ochmann H, Sartor K (2002) *Neurosci Lett* 331: 50-4.
9. Ehrsson HH, Geyer S, Naito E (2003) *J Neurophysiol* 90: 3304-16.
10. Shuster LI, Lemieux SK (2005) *Brain Lang* 93: 20-31.
11. Birn RM, Diamond JB, Smith MA, Bandettini PA (2006) *NeuroImage* 31: 1536-48.
12. McKay LC, Evans KC, Frackowiak RSJ, Corfield DR (2003) *J Appl Physiol* 95: 1170-8.
13. Simonyan K, Saad ZS, Loucks TMJ, Poletto CJ, Ludlow CL (2007) *NeuroImage* 37: 401-9.
14. Schulz GM, Varga M, Jeffires K, Ludlow CL, Braun AR (2005) *Cereb Cortex* 15: 1835-47.
15. Paus T, Perry DW, Zatorre RJ, Worsley KJ, Evans AC (1996) *Eur J Neurosci* 8: 2236-46.
16. Christoffels IK, Formisano E, Schiller NO (2007) *Hum Brain Mapp* 28: 868-79.
17. Miller GA (1947) *Psychol Bull* 44: 105-29.
18. Guenther FH (1994) *Biol Cybern* 72: 43-53.
19. Guenther FH, Ghosh SS, Tourville JA (2006) *Brain Lang* 96: 280-301.
20. Golfinopoulos E, Tourville JA, Guenther FH (2009) *NeuroImage* (in press).
21. Bandettini PA et al. (1998) *Magn Reson Med* 39: 410-6.
22. Birn RM, Bandettini PA, Cox RW, Jesmanowicz A, Shaker R (1998) *Magn Reson Med* 40: 55-60.
23. Tourville JA, Reilly KJ, Guenther FH (2008) *NeuroImage* 39:1429-43.
24. Riva D (1998) *Cortex* 34: 279-87.
25. Shuren J (1993) *J Neurol* 240: 216-8.
26. Ziegler W, Kilian B, Deger K (1997) *Neuropsychologia* 35: 1197-208.
27. Bohland JW, Guenther FH (2006) *NeuroImage* 32: 821-41.
28. Dinomais M et al. (2009) *Neuroreport* (epub).
29. Geschwind N (1965) *Brain* 88: 237-94.

Contact

Simon Overduin (overduin@bu.edu, http://speechlab.bu.edu)