BRIEF REPORT

Neural evidence for predictive coding in auditory cortex during speech production

Kayoko Okada^{1,2} · William Matchin³ · Gregory Hickok²

Published online: 10 April 2017
© Psychonomic Society, Inc. 2017

Abstract Recent models of speech production suggest that motor commands generate forward predictions of the auditory consequences of those commands, that these forward predictions can be used to monitor and correct speech output, and that this system is hierarchically organized (Hickok, Houde, & Rong, *Neuron*, 69(3), 407–422, 2011; Pickering & Garrod, *Behavior and Brain Sciences*, 36(4), 329–347, 2013). Recent psycholinguistic research has shown that internally generated speech (i.e., imagined speech) produces different types of errors than does overt speech (Oppenheim & Dell, *Cognition*, 106(1), 528–537, 2008; Oppenheim & Dell, *Memory & Cognition*, 38(8), 1147–1160, 2010). These studies suggest that articulated speech might involve predictive coding at additional levels than imagined speech. The current fMRI experiment investigates neural evidence of predictive coding in speech production. Twenty-four participants from UC Irvine were recruited for the study. Participants were scanned while they were visually presented with a sequence of words that they reproduced in sync with a visual metronome. On each trial, they were cued to either silently articulate the sequence or to imagine the sequence without overt articulation. As expected, silent articulation and imagined speech both engaged a

left hemisphere network previously implicated in speech production. A contrast of silent articulation with imagined speech revealed greater activation for articulated speech in inferior frontal cortex, premotor cortex and the insula in the left hemisphere, consistent with greater articulatory load. Although both conditions were silent, this contrast also produced significantly greater activation in auditory cortex in dorsal superior temporal gyrus in both hemispheres. We suggest that these activations reflect forward predictions arising from additional levels of the perceptual/motor hierarchy that are involved in monitoring the intended speech output.

Keywords Cognitive neuroscience · Speech production

It is well-established that speech production critically involves sensory systems. Evidence for this assertion comes primarily from various demonstrations of the modulatory or disruptive effect of altered auditory feedback on speech production, including late onset deafness, delayed feedback or formant-shifted feedback (Guenther, Hampson, & Johnson, 1998; Hickok, 2012a; Hickok, Houde, & Rong, 2011; J. Houde & Nagarajan, 2011; J. F. Houde & Jordan, 1998; Perckell, 2012; Tremblay, Shiller, & Ostry, 2003) and from similar demonstrations in the somatosensory domain (Waldstein, 1989).

It has been hypothesized that forward predictive coding is a mechanism for this involvement (Stuart, Kalinowski, Rastatter, & Lynch, 2002; Yates, 1963). Forward predictive coding (see also similar terms like forward model, internal model, efference copy, or *corollary discharge*) refers to the idea that motor plans or commands lead to predictions of the sensory consequences of those plans/commands via motor-to-sensory neural projections, which in turn serve to facilitate error detection and correction in motor control (Burnett, Senner, & Larson, 1997; J. F. Houde & Jordan, 1998).

✉ Gregory Hickok
greg.hickok@uci.edu

¹ Department of Psychology, Loyola Marymount University, Los Angeles, CA, USA

² Department of Cognitive Sciences, University of California at Irvine, Irvine, CA, USA

³ Department of Linguistics, University of California at San Diego, La Jolla, CA, USA

Neural evidence for the existence of a forward predictive coding mechanism in speech motor control has come from experiments showing either a suppression of the auditory response, measured using magnetoencephalography (MEG), to self-produced speech compared to externally delivered speech (Tremblay et al., 2003) or from an enhanced response, measured using functional magnetic resonance imaging (fMRI), to self-produced altered versus unaltered speech (Hickok, 2012a; J. Houde & Nagarajan, 2011; Tourville, Reilly, & Guenther, 2008). A complication in interpreting the suppression effect data is that it is hard to precisely control the acoustic stimulus in self- versus externally produced speech. This is less of a concern in experiments that compare unaltered (predicted) versus altered (unpredicted) speech, but the observed localization of this effect did not fall squarely in auditory cortex but rather in what appears to be an auditory-motor integration area, area Spt (Desmurget & Grafton, 2000; Kawato, 1999; Wolpert, 1997; Wolpert, Ghahramani, & Jordan, 1995). This does not challenge the claim for the existence of a forward predictive code, but does question the role of auditory cortex proper in this system.

Here we use a very simple fMRI paradigm to provide straightforward evidence for the existence of a forward predictive coding mechanism in speech production and to home in on the level(s) of motor planning that give(s) rise to it. We asked participants to read a list of words and either imagine speaking them without overtly articulating or to overtly articulate the words without phonating (see Fig. 1). Thus both conditions are matched for acoustic input (i.e., no speech input). Previous behavioral research has shown that these two tasks engage different levels of linguistic/motor planning. Imagined speech engages lexical-level processes but not lower level phonological-level processes, whereas silently

articulated speech engages both levels of processing (Aliu, Houde, & Nagarajan, 2009; Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000; Heinks-Maldonado, Nagarajan, & Houde, 2006; J. F. Houde, Nagarajan, Sekihara, & Merzenich, 2002; Numminen, Salmelin, & Hari, 1999). We reasoned that engaging motor-phonological processes should generate a forward prediction of the acoustic consequences of the executed (silent) speech whereas engaging lexical-level processes should not. If true, we should then see differential activation in auditory cortex for silently articulated speech compared to imagined speech despite the fact that neither condition involves any speech input.

Method

Subjects

Twenty four participants (15 females) between 18 and 40 years of age were recruited from the University of California, Irvine (UCI) community and received monetary compensation for their time. The volunteers were right-handed, native English speakers with normal or corrected-to-normal vision, no known history of neurological disease, and no other contraindications for MRI. Informed consent was obtained from each participant prior to participation in the study in accordance with guidelines from UCI Institutional Review Board, which approved this study. Four subjects were omitted from data analysis: two subjects had excessive motion ($>3^\circ$ of movement), one subject reported excessive errors on the task ($>50\%$ error rate), and one subject was eliminated because of response box failure to log responses.

Fig. 1 Example of a single trial. Subjects were presented with a tongue twister sequence that remained on screen for 3 seconds, followed by a cue to either articulate the sequence or imagine the sequence. They recited each word in sync with the visual metronome (Color figure online)

Stimuli and task

fMRI was used to monitor blood oxygenation (BOLD) changes elicited by reciting tongue twisters. The use of tongue twisters increases the probability of speech errors and therefore the load on error detection systems. The experiment was closely modeled on previous behavioral studies using these stimuli (Oppenheim & Dell, 2008, 2010), and participants were scanned while they recited a set of four words (e.g., lean reed reef leach) in sync with a visual metronome. Two speech production conditions were included, one in which speech was articulated without phonating (silent articulation) and one in which speech production was imagined without articulation (imagined). On each trial, a tongue twister phrase was visually presented on screen for 3 seconds, and then subjects were cued to silently articulate the sequence or imagine saying the sequence without mouth movements (see Fig. 1). The cue was a cartoon face that remained on screen for 500 ms, and a red arrow pointed either to the head or to the lips. An arrow pointing to the head cued the participants to imagine saying the word, and an arrow pointing to the lips cued the subjects to silently articulate the words. A red fixation appeared on screen 500 ms after cue offset, and this served as the visual metronome and flashed at a rate of 2/s. Participants recited one word per fixation in sync with the metronome. Following Oppenheim and Dell (2010), after recitation, participants indicated with a button press whether or not they correctly produced the sequence.

A total of 32 sets of tongue twisters were used in the study. Although not of primary interest in this study, the word lists in the experiment varied in lexicality. Lexical bias refers to the tendency for word errors to create a real word instead of a nonword (e.g., scenarios in which the target word is *reef* but slips to *leaf* is more likely than scenarios in which the target word is *wreath* and slips to *leath* because *leath* is a nonword). These stimuli were designed so that if an error occurred on the third word of each sequence, the outcome would yield either a real-word error (*lean* would induce the error *leaf* instead of *reef*) or a nonword error (*lean* would induce the error *leath* instead of *wreath*). The specific metrics of the stimuli are described elsewhere (Oppenheim & Dell, 2010).

A single trial was 8 seconds in length, and there were 36 trials in each session. Each session consisted of an equal number of tongue twister phrases, which were silently articulated or imagined. The experiment consisted of eight experimental sessions, and each session consisted of 16 trials of each type, which were randomly presented along with four rest trials (fixation). The study started with a short practice session with 10 trials to familiarize subjects with the task. Subjects were scanned during the practice session to acclimatize them to the fMRI environment. The study ended with a high-resolution structural scan, and the entire experiment was 1 hour in length. Stimuli presentation and timing was controlled using Cogent

software (http://www.vislab.ucl.ac.uk/cogent_2000.php) implemented in MATLAB 7.1 (MathWorks, Inc, USA) running on a dual-core IBM Thinkpad laptop.

Imaging

MR images were obtained in a Philips Achieva 3T (Philips Medical Systems, Andover, MA) fitted with an eight-channel RF receiver head coil, at the Research Imaging Center scanning facility at the University of California, Irvine. Images during the experimental sessions were collected using Fast Echo EPI (sense reduction factor = 2.0, matrix = 112×112 mm, TR = 2.0 s, TE = 25 ms, size = $2.5 \times 2.5 \times 2.5$ mm). A total of 1,152 echo planar images (EPI) were collected over eight sessions, and 41 slices provided whole brain coverage. After the functional scans, a high resolution T1-weighted anatomical image was acquired with an MPRAGE pulse sequence in axial plane (matrix = 256×256 mm, TR = 8 ms, TE = 3.6 ms, flip angle = 8° , size = $1 \times 1 \times 1$ mm).

Data analysis

Data preprocessing and analyses were performed using AFNI software (Cox, 1996). First, motion correction was performed by creating a mean image from all of the volumes in the experiment and then realigning all volumes to that mean image using a six-parameter rigid-body model (Cox & Jesmanowicz, 1999). The images were then high-pass filtered at 0.008 Hz and spatially smoothed with an isotropic 8-mm full width half maximum (FWHM) Gaussian kernel. The anatomical image for each subject was coregistered to his or her mean EPI image. Data analysis proceeded in two steps. First, multiple regression analysis was performed at the single subject level, parameter estimates for events of interest were obtained, and then these data were transformed into standardized space for group-level analysis.

First-level analysis was performed on the time course of each voxel's BOLD response for each subject using AFNI software (Cox, 1996). Regression analysis was performed using AFNI's 3dDeconvolve function, and the regressors were created by convolving the predictor variables representing the time course of stimulus presentation with a gamma variate function. A total of 15 regressors were entered into the analysis. The first eight regressors were used to model the following experimental trial types—articulation: similar onset, nonword errors; articulation: similar onset, word errors; articulation: dissimilar onset, nonword errors; articulation: dissimilar onset, word errors; imagining: similar onset, nonword errors; and imagining: similar onset, word errors; imagining: dissimilar onset, nonword errors; and imagining: dissimilar onset, word errors. The ninth regressor corresponded to all of the trials in which the subjects reported making an error. Thus, the regressors modeling the speech production conditions

only include trials in which subjects report accurately reproducing the sequence of words. An additional six regressors representing the motion parameters determined during the realignment stage of processing were entered into the model. Parameter estimates for the events of interest were obtained and statistical maps were created.

For group-level analysis, these statistical maps for each participant were transformed into standardized space (Talairach & Tournoux, 1988) using a Talairach template and resampled into 2-mm³ voxels. We performed *t* tests to examine group differences between silent articulation and imagining the word lists. We also examined the lexicality effect and phonemic similarity effects. Group-level activation maps were created, and a corrected significance level was set at $p < .05$. This threshold was determined using 3dFWHMx and 3dClustStim (Cox, 1996) to estimate the smoothness of the noise, then combines minimum cluster size and *p* threshold ($p < .001$) to correct for multiple testing.

Results

Silent articulation and imagining speech

First, we examined the neural regions activated by silent articulation and by imagining separately. This was done to ensure that regions previously implicated in speech production were engaged during the two tasks and to examine differences in activation patterns between the two tasks. As expected, silent articulation engaged a wide network of regions previously implicated in speech production. We found significant activation in inferior and middle frontal gyri in the left hemisphere, bilateral precentral gyrus, bilateral inferior parietal cortex including angular gyrus and supramarginal gyrus, insula, basal ganglia, and cerebellum. We also found

significant activity in left superior temporal gyrus. Imagining word lists activated a similar network of areas, such as left inferior and middle frontal cortex, bilateral precentral gyrus, bilateral inferior parietal cortex, and cerebellum. Figure 2 illustrates the activation patterns associated with the two tasks, and areas significantly activated for each task are reported in Table 1.

Articulation versus imagining

A contrast of silent articulation with imagining word lists yielded greater activity in bilateral superior and middle temporal gyrus, precentral gyrus, anterior cingulate, bilateral parahippocampal gyrus, cerebellum, and basal ganglia. Each of these areas was more active during silent articulation than imagined speech, and the reverse contrast of Imagining > Articulation did not yield any significant regions. What is particularly interesting is the large clusters of activation in bilateral auditory cortex, given that both tasks involved silent recitation. Subjects did not have overt speech feedback in the experiment. This shows that engaging motor articulators changes the activation pattern in auditory cortex compared to imagining/thinking about speech. Figure 3 shows the regions more activated by articulation, and Table 1 lists the Talairach coordinates of the significant clusters.

Lexicality effect

Although not the main focus of our study, the lexicality effect was examined by comparing activation associated with reciting word lists biased to produce real-word errors versus non-word errors. A lexicality effect was not observed at a threshold corrected for multiple comparisons. However, at a lowered threshold ($p < .001$, uncorrected), this contrast revealed a lexicality effect in a region previously implicated in lexical-level

Silent Articulation

Imagining

Fig. 2 Silent articulation and imagining word lists. Group activation map ($N = 20$) overlaid on a template brain illustrating regions significantly activated in the Articulation > Baseline, and Imagining > Baseline conditions ($p < .05$, corrected) (Color figure online)

Table 1 Summary of brain regions significantly activated in the Articulation > Imagining contrast ($p < .05$, corrected)

REGION	Peak voxel			Approximate Brodmann	Cluster size
	<i>x</i>	<i>y</i>	<i>z</i>		
Articulation > Imagining					
Right precentral gyrus	59	−7	38	BA 6	11,606
(contiguous cluster includes the insula, superior and middle temporal gyri)					
Right cerebellum	17	−59	−22		2722
(contiguous cluster include left cerebellum)					
Left cingulate	−1	11	32	BA24	1619
Right superior frontal gyrus	7	−7	72	BA 6	474
Right superior frontal gyrus	7	71	10	BA 10	78
Lift middle temporal gyrus	−45	−53	0	BA 19/37	52
Lift inferior frontal gyrus	−37	29	−2	BA 47	29
Lift inferior frontal gyrus	−47	27	2	BA 45	27
Imagining > Articulation					
Left medial frontal gyrus	−11	41	−16	BA 11	367
Right cerebellum	15	−9	−28		33
Articulation > baseline					
Cerebellum (bilateral)	−45	−71	−22		28,731
(cluster includes bilateral visual cortex)					
Superior frontal gyrus (bilateral)	−1	−1	54	BA 6	2251
Left inferior parietal cortex	−43	−49	54	BA 40	1247
Right middle frontal gyrus	35	45	22	BA 10	203
Left middle temporal gyrus	−47	−49	−2	BA 37	146
Right inferior parital cortex	43	−45	42	BA 40	124
Right middle temporal gyrus	47	−35	2	BA 22	40
Imagining > baseline					
Cerebellum (bilateral)	−45	−69	−24		13,762
Left precentral gyrus	−47	−3	50	BA 6/46	3342
(cluster includes middle/inferior frontal)					
Medial frontal gyrus	1	1	54	BA 6	1356
Left inferior parietal cortex	−45	−49	52	BA 40	895
Right putamen	19	3	8		251
Right precentral gyrus	55	−3	42	BA 6	157
Right middle frontal gyrus	33	45	24	BA 10	111
Right inferior parietal cortex	43	−43	42	BA 40	94
Left thalamus	−11	−17	4		63
Left putamen	−27	−17	−6		37

Also reported are regions significantly activated during each task separately: Silent Articulation > Baseline and Imagining > Baseline. Reported are Talairach coordinates for the peak voxel in significant clusters in the group analysis ($N = 20$) and approximate Brodmann areas (BA)

Fig. 3 Evidence of predictive coding. Group activation map ($N = 20$) overlaid on a template brain illustrating regions significantly activated in the contrast Articulation > Imagining ($p < .05$, corrected). There was

greater activation in bilateral auditory cortex when subjects were silently articulating word lists compared to imagining word lists. (Color figure online)

processes, the posterior middle temporal gyrus in the left hemisphere (peak coordinate $[-53 -59 6]$). That is, when the word list was biased to produce nonword errors (e.g., producing *leath* instead of *leaf*), greater activation was observed in pMTG compared to a word list biased to produce word errors, and this effect was observed on error-free trials (i.e., errors were modeled separately). What makes this finding potentially interesting is that the difference in activation pattern cannot be attributed to lexical differences in the words presented to the subject or to lexical differences in what was spoken; in both conditions, subjects viewed the same set of words (although in different combinations) and spoke real words. Instead, what is driving the activation difference is the *potential* for a word versus a nonword error that was not overtly committed. If this effect replicates in future studies it suggests that the system detected the distinction internally and this could only be the case if in fact an internal error was committed and then corrected prior to accurate output. Figure 4 illustrates the lexicality effect. This should be further investigated in future work.

Discussion

We observed a substantial effect of speech production condition, with the silent articulation condition generating more activation than the imagined speech condition in a wide network of brain regions. Many of these are unsurprising, such as the greater activity in primary sensorimotor cortex (predominantly in the right hemisphere, however), the cerebellum, and subcortical nuclei all of which play a role in overt movement control. Most interestingly, we also found robust activation differences (Silent Articulation > Imagined Speech) in the superior temporal gyri bilaterally, including portions of auditory cortex, despite the fact that there was no difference in auditory input in the two conditions. One

Fig. 4 Lexicality effect. Group activation map ($N = 20$) overlaid on a template brain illustrating regions significantly activated in the Nonword > Word contrast ($p < .001$, uncorrected). This activation includes only those trials in which subjects report correctly reciting the tongue twister sequence. Error trials are omitted. (Color figure online)

possible explanation of this effect is that it reflects a mismatch error signal: articulatory plans result in an internal forward prediction of the acoustic consequences of speech articulation, which fail to arrive. Mismatch error signals have been reported previously under conditions of altered auditory or somatosensory feedback (Golfinopoulos et al., 2011; Tourville et al., 2008) and similar inferences regarding forward prediction of sensory consequences in speech have been used to explain differences in the acoustic response to self-versus other-generated speech (Heinks-Maldonado et al., 2006; J. F. Houde et al., 2002; Ventura, Nagarajan, & Houde, 2009). If overt articulation (i.e., the actual execution of motor speech plans) results in stronger forward predictions than imagined speech, this could explain the observed activation in auditory cortex in the articulation versus imagined speech condition. This then, would be a reflection of a lower level of feedback control than the higher level circuit revealed by the lexicality manipulation, consistent with recent hierarchical models of feedback control in speech and motor control generally (Diedrichsen, Shadmehr, & Ivry, 2010; Grafton & Hamilton, 2007; Hickok, 2012a).

Of course, a forward predictive mechanism is not the only possible explanation for our findings. One might argue more broadly, for example, that auditory imagery is evoked more strongly during actual articulation than imagined speech, leading to the observed activation difference. However, one then might ask why this should be the case; why is the auditory system so compelled to image the acoustic correlates of articulated speech? And the best available answer to this question is that there are strong computational arguments from motor control that such “imagery” serves an important function in speech production.

Second, we found a possible brain region that is sensitive to the well-established lexical bias effect in speech production (although slightly under threshold). When participants were biased to make nonword errors, there was enhanced activation in posterior middle temporal gyrus (pMTG), a region previously implicated in lexical-semantic processing. Neuroimaging and neuropsychological studies have demonstrated this region’s involvement in semantic tasks (Diedrichsen et al., 2010; Grafton & Hamilton, 2007), word retrieval, and naming (Binder, Desai, Graves, & Conant, 2009; Binder et al., 1997; Rodd, Davis, & Johnsrude, 2005). Conditional on future replication, we interpret this pattern as evidence for the existence of neural network for detecting and correcting word-level errors prior to overt production, broadly consistent with recent proposals regarding hierarchically organized internal feedback control circuits for speech production (DeLeon et al., 2007; Hillis et al., 2001). The pMTG was not the only region to show a differential response to word versus non word biased lists, suggesting that it is part of a larger network. One of these activations implicates a structure, the cerebellum, that is

widely believed to play a role in internal models for motor control (Hickok, 2012a, 2012b)—a fact that is broadly consistent with the interpretation of our data. Future work should investigate this further.

The main finding in this study is that auditory cortex activity is modulated as a function of silent motor speech articulation. We hypothesized that if forward predictions are generated at different levels of the speech motor control hierarchy, we should find more activity in some portions of auditory cortex during silently articulated speech compared to imagined speech because predictions are being generated from multiple levels, even though neither condition involves any auditory input. In line with our predictions, we found greater activation in several regions of auditory cortex during silent articulation compared with imagined speech. We suggest that these activations reflect forward predictions arising from additional levels of the perceptual/motor hierarchy that are involved in monitoring the intended speech output.

References

- Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, 21(4), 791–802. doi:10.1162/jocn.2009.21055
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796. doi:10.1093/cercor/bhp055
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17(1), 353–362.
- Burnett, T. A., Sennar, J. E., & Larson, C. R. (1997). Voice F0 responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice*, 11(2), 202–211.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29(3), 162–173. doi:10.1006/cbmr.1996.0014
- Cox, R. W., & Jesmanowicz, A. (1999). Real-time 3D image registration for functional MRI. *Magnetic Resonance in Medicine*, 42(6), 1014–1018. doi:10.1002/(SICI)1522-2594(199912)42:6<1014::AID-MRM4>3.0.CO;2-F
- Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Hum Brain Mapp*, 9(4), 183–191.
- DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., ... & Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*, 130(5), 1408–1422. doi:10.1093/brain/awm011
- Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Science*, 4(11), 423–431.
- Diedrichsen, J., Shadmehr, R., & Ivry, R. B. (2010). The coordination of movement: Optimal feedback control and beyond. *Trends in Cognitive Science*, 14(1), 31–39. doi:10.1016/j.tics.2009.11.004
- Golfinopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *NeuroImage*, 55(3), 1324–1338. doi:10.1016/j.neuroimage.2010.12.065
- Grafton, S. T., & Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4), 590–616. doi:10.1016/j.humov.2007.05.009
- Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, 105, 611–633.
- Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, 17(13), 1375–1379. doi:10.1097/01.wnr.0000233102.43526.e9
- Hickok, G. (2012a). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13(2), 135–145. doi:10.1038/nrn3158
- Hickok, G. (2012b). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders*, 45(6), 393–402. doi:10.1016/j.jcomdis.2012.06.004
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, 69(3), 407–422. doi:10.1016/j.neuron.2011.01.019
- Hillis, A. E., Kane, A., Tuffiash, E., Ulatowski, J. A., Barker, P. B., Beauchamp, N. J., & Wityk, R. J. (2001). Reperfusion of specific brain regions by raising blood pressure restores selective language functions in subacute stroke. *Brain and Language*, 79(3), 495–510. doi:10.1006/brln.2001.2563
- Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, 279, 1213–1216.
- Houde, J. F., & Nagarajan, S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5(82). doi:10.3389/fnhum.2011.00082
- Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *J Cogn Neurosci*, 14(8), 1125–1138. doi:10.1162/089892902760807140
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6), 718–727.
- Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, 265(2), 119–122.
- Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, 106(1), 528–537. doi:10.1016/j.cognition.2007.02.006
- Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & Cognition*, 38(8), 1147–1160. doi:10.3758/MC.38.8.1147
- Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, 25(5), 382–407. doi:10.1016/j.jneuroling.2010.02.011
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavior and Brain Sciences*, 36(4), 329–347. doi:10.1017/S0140525X12001495
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, 15(8), 1261–1269. doi:10.1093/cercor/bhi009
- Stuart, A., Kalinowski, J., Rastatter, M. P., Lynch, K., & S. Pt. 1. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, 111, 2237–2241.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging*. Stuttgart: Thieme.
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3), 1429–1443. doi:10.1016/j.neuroimage.2007.09.054

- Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, 423(6942), 866–869. doi:[10.1038/nature01710](https://doi.org/10.1038/nature01710)
- Ventura, M. I., Nagarajan, S. S., & Houde, J. F. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neuroscience*, 10, 58. doi:[10.1186/1471-2202-10-58](https://doi.org/10.1186/1471-2202-10-58)
- Waldstein, R. S. (1989). Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. *Journal of the Acoustical Society of America*, 88, 2099–2144.
- Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Science*, 1(6), 209–216. doi:[10.1016/S1364-6613\(97\)01070-X](https://doi.org/10.1016/S1364-6613(97)01070-X)
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880–1882.
- Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, 60, 213–251.