Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 4pSCa: Auditory Feedback in Speech Production II

4pSCa3. Neural evidence for state feedback control of speaking

John F. Houde*, Naomi S. Kort, Caroline A. Niziolek, Edward F. Chang and Srikantan S. Nagarajan

***Corresponding author's address: Otolaryngology -- Head and Neck Surgery, University of California San Francisco, 513 Parnassus Ave. S362, San Francisco, CA 94143, houde@phy.ucsf.edu**

An important recent development in neuroscience has been the use of models based on state feedback control (SFC) to explain the role of the central nervous system in motor control. In SFC, control is based on internal feedback of an estimate of the dynamic state of the thing (e.g., arm) being controlled. Within the internal loop, the state is predicted from outgoing motor commands and corrected by comparing the feedback expected to result from this state with actual incoming sensory feedback. SFC has received scant attention in the speech community, but the indirect role it suggests for feedback can account for much of what is known about the role of feedback in speech motor control. Our lab has been investigating how well SFC also accounts for the neural correlates of auditory feedback processing during speaking. Our principal approach has used magnetoencephalography to record the cortical activity of speakers as they hear themselves speaking, but recently, we have also completed an auditory feedback study based on electrocorticography. Many of the results of these studies have supported the SFC model, but some have posed challenges for it, which will be discussed. (Support provided by NSF grant BCS-0926196 and NIH grant R01-DC010145.)

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4799495]
Received 22 Jan 2013; published 2 Jun 2013
Proceedings of Meetings on Acoustics, Vol. 19, 060178 (2013)

In pursuing an understanding of the neural mechanisms involved in speech motor control, a central issue to be resolved is how auditory feedback is processed during speaking. The behavioral phenomena associated with auditory feedback suggest its role in speaking is not entirely straightforward. Clearly, auditory feedback is important for learning to speak, since, in general, children born deaf don't learn to speak unless fitted with cochlear implants [1]. But later in life, after the speaking skill has developed, the role of auditory feedback becomes more complex. On one hand, intelligible speech can be produced in the absence of auditory feedback. This is demonstrated by the speech of those deafened in adult life, which can remain intelligible for decades [2], and by the clarity of speech produced when masking noise completely blocks hearing [3]. But this does not mean that auditory feedback is ignored after speaking is learned. The effect of hearing masking noise on speech, while not necessarily affecting its clarity does cause speakers to raise the volume of their speech (the so-called Lombard effect) [3]. An opposite effect can be seen when speakers hear amplified feedback of their speech. In this case, the amplified feedback causes them to lower the volume of their speaking [4]. And in the lab, several other artificial feedback manipulations can be used to show the sensitivity of speech motor control to auditory feedback. These include the long-studied dysfluency-producing effect of delaying auditory feedback [5], as well as the more recently observed phenomena of compensation and long-term adaptation to realtime alterations of the pitch and spectral features of auditory feedback [6-11].

In sum, the paradox about auditory feedback is that it appears that it is not necessary for speech production, but if auditory feedback is present, it needs to be correct or speech will be affected. For these reasons, current models of speech motor control relegate auditory feedback to a more indirect role, with an inner feedback loop within the CNS that directly controls speech output, and actual sensory feedback forming a slower, possibly delayed and intermittent, external loop that updates the internal feedback loop [12-14]. Such models can be described as variations on the general theory of state feedback control (SFC), developed in the domain of modern control engineering theory [15-17]. SFC models have become more prevalent in many domains of motor control research [18-20], and here we consider the applicability of SFC to modeling speech motor control. We describe the SFC model in some detail below and then discuss neuroimaging and electrophysiological evidence in support of this model.

SPEECH MOTOR CONTROL AS STATE FEEDBACK CONTROL

As shown in Figure 1, our SFC model is based on the classic process of observer-based state feedback control, and we have postulated roles in this process for different areas of cortex [21, 22]. In our model, a speech goal arising from frontal and premotor cortex (**blue** lines) enables a series of state estimation functions($\widehat{\mathbf{vtdyn}}(\mathbf{u}_{t-1}, \hat{\mathbf{x}}_{t-1})$, $\widehat{\mathbf{vtout}}(\hat{\mathbf{x}}_t)$, $\mathbf{K}_t(\tilde{\mathbf{y}})$) for a control law $\mathbf{U}_t(\hat{\mathbf{x}}_t)$ in motor cortex. This control law generates controls $\mathbf{u}_t$ of the vocal tract musculature based on an estimate $\hat{\mathbf{x}}_t$ of its current state, and the state estimate is continually updated in a prediction/correction process as controls are applied. First, in the prediction (**green**) direction, efference copy of the last controls $\mathbf{u}_{t-1}$ sent to the vocal tract are input to a mapping $\widehat{\mathbf{vtdyn}}(\mathbf{u}_{t-1}, \hat{\mathbf{x}}_{t-1})$ in ventral premotor cortex that predicts the next state $\hat{\mathbf{x}}_{t|t-1}$ of the vocal tract. This mapping reflects what premotor cortex has learned about the dynamics of the vocal tract (i.e., its responses to controls) when it is directed to produce the current speech goal, and its output – the next state prediction – can, by itself, function as a state estimate in the absence of other sensory feedback. When sensory feedback is available, it's compared with a feedback expectation derived from the next state prediction, with the difference used to correct the next state prediction. To simplify the figure, we have shown this process as occurring in generic "sensory cortex", but this is meant to represent similar processes occurring in both the somatosensory (parietal) and auditory (temporal) cortices. We postulate that high order sensory cortex learns a mapping $\widehat{\mathbf{vtout}}(\hat{\mathbf{x}}_t)$ from next state predictions $\hat{\mathbf{x}}_{t|t-1}$ to expected feedback $\hat{\mathbf{y}}_{t-\hat{\mathbf{N}}}$, which includes learning of the sensory input processing delay $\hat{\mathbf{N}}$, possibly as a distinct separate step $\boxed{\mathbf{z}^{-\hat{\mathbf{N}}}}$ occurring via the interaction of premotor cortex and the cerebellum. Then, in the correction (**red**) direction, expected feedback $\hat{\mathbf{y}}_{t-\hat{\mathbf{N}}}$ is compared with actual feedback $\mathbf{y}_{t-\mathbf{N}}$ in primary sensory cortex, with the resulting feedback prediction error $\tilde{\mathbf{y}}_{t-\hat{\mathbf{N}}}$ passed back to high order sensory cortex. There, another mapping $\mathbf{K}_t(\tilde{\mathbf{y}})$ is learned between feedback prediction errors $\tilde{\mathbf{y}}_{t-\hat{\mathbf{N}}}$ and

corrections $\hat{\mathbf{e}}_t$ to the next state prediction $\hat{\mathbf{x}}_{t|t-1}$. State correction $\hat{\mathbf{e}}_t$ is then sent back to premotor cortex, where it is combined with $\hat{\mathbf{x}}_{t|t-1}$ to update the state estimate $\hat{\mathbf{x}}_t$ used by the state feedback control law $\mathbf{U}_t(\hat{\mathbf{x}}_t)$ in motor cortex.

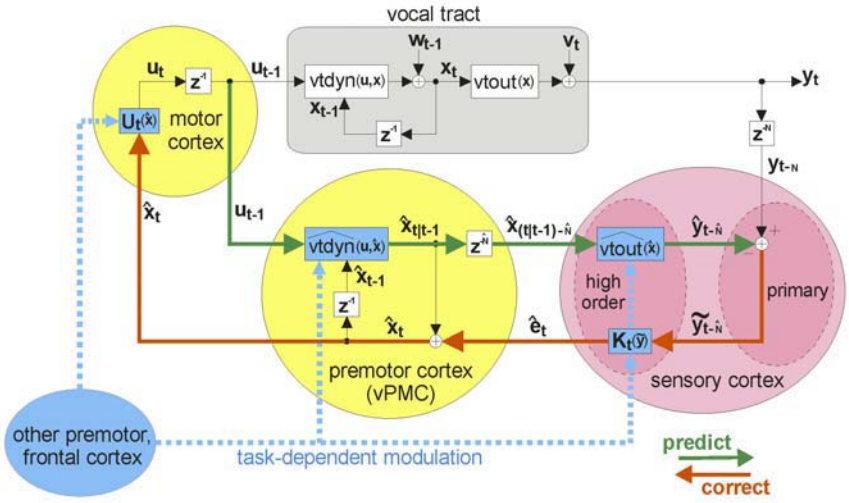

FIGURE 1. Model of speech motor control as a state feedback control (SFC) process.

This SFC model of speech motor control explains many the behavioral phenomena associated with how auditory feedback interacts with speech production, but the model also makes a number of predictions about the neural processing of auditory feedback during speaking. First, the model predicts the now well-known speaking-induced suppression (SIS) effect, where the response of a subject's auditory cortices to his/her own self-produced speech is significantly smaller than their response to similar, but externally produced speech (e.g., tape playback of the subject's previous self-productions. This effect has been seen using positron emission tomography (PET) [23-25], electroencephalography (EEG) [26, 27], magnetoencephalography (MEG) [28-33], and electrocorticography (ECoG) [34].

In a series of recent experiments, we have been testing further predictions made by the SFC model about neural processing of auditory feedback. These experiments are based on examining speakers' responses (both neural and behavioral) to brief alterations of the pitch of their audio speech feedback. Such alterations are known to generate short-latency compensatory responses in speakers (Figure 2a) [6]. We have also simulated our SFC model reacting to a pitch perturbation (Figure 2b), and have verified that, in order to be stable in the presence of noise, an observer-based state feedback controller will only partially compensate for feedback alterations, which is what is seen in the speakers' responses.

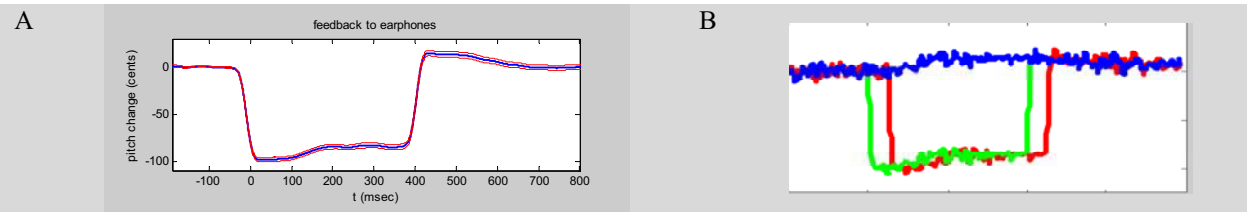

FIGURE 2. Pitch feedback perturbation response: A) mean pitch change (in cents) of subject's feedback in response to 400ms, -100 cent pitch perturbation, inducing a partial compensation response that cancels out roughly 15% of the perturbation. B) simulated single trial SFC model response to -100 cent pitch feedback perturbation. **blue**: pitch output of the model, **green**: feedback received by the model, showing applied feedback perturbation, **red**: internally delayed reception of sensory feedback in the model.

Prior studies have reported a response to feedback alterations (like that shown in Figure 2) that in some sense is the opposite of the SIS effect. Some auditory cortical areas have an enhanced response to feedback alterations during speaking, compared to their response to hearing the same altered feedback when passively listening [35, 36]. This effect has been called speech perturbation response enhancement (SPRE), and has been seen

in an ERP study based on EEG [35], as well as an fMRI study which localized the effect to the dorsal speech processing stream SPT region (left pSTG, left vSMG, right mSTG) [36]. Our SFC model explains SPRE as arising from high-level auditory areas where the speaking-dependent gain $\mathbf{K}_t(\tilde{\mathbf{y}})$ maps feedback prediction errors into state estimate corrections.

CORTICAL RESPONSES TO ALTERED PITCH FEEDBACK: AN MEG STUDY

One recent test of our SFC model was an experiment where magnetoencephalography (MEG) was used to record subjects' cortical responses while the pitch of their ongoing phonation was perturbed. Subjects sat with their head in the MEG scanner (275-channel, whole-head, VSM Medtech, Canada) while wearing insert earphones. At a visual prompt, they phonated /a/ with their speech being picked up by a microphone and fed through a DSP back to their earphones. The DSP perturbed the pitch of in real time using a custom-built program (details of this program have been described elsewhere [37, 38]).

The experiment consisted of interleaved speak and listen conditions. In each speak condition trial, subjects phonated /a/ for roughly 3 seconds. At a randomly jittered latency within their phonation, the DSP was directed to perturb pitch for 400ms by +/- 100 cents (see Figure 2A). After 74 speak condition trials, the feedback recorded was played back, and the subject passively listened (listen condition). This cycle of speak and listen conditions was then repeated. The MEG data acquired was analyzed using virtual channel analysis [39, 40]. As shown in figures 3A and 3B, four virtual channel locations were chosen in each hemisphere to match postulated locations of our SFC model components.

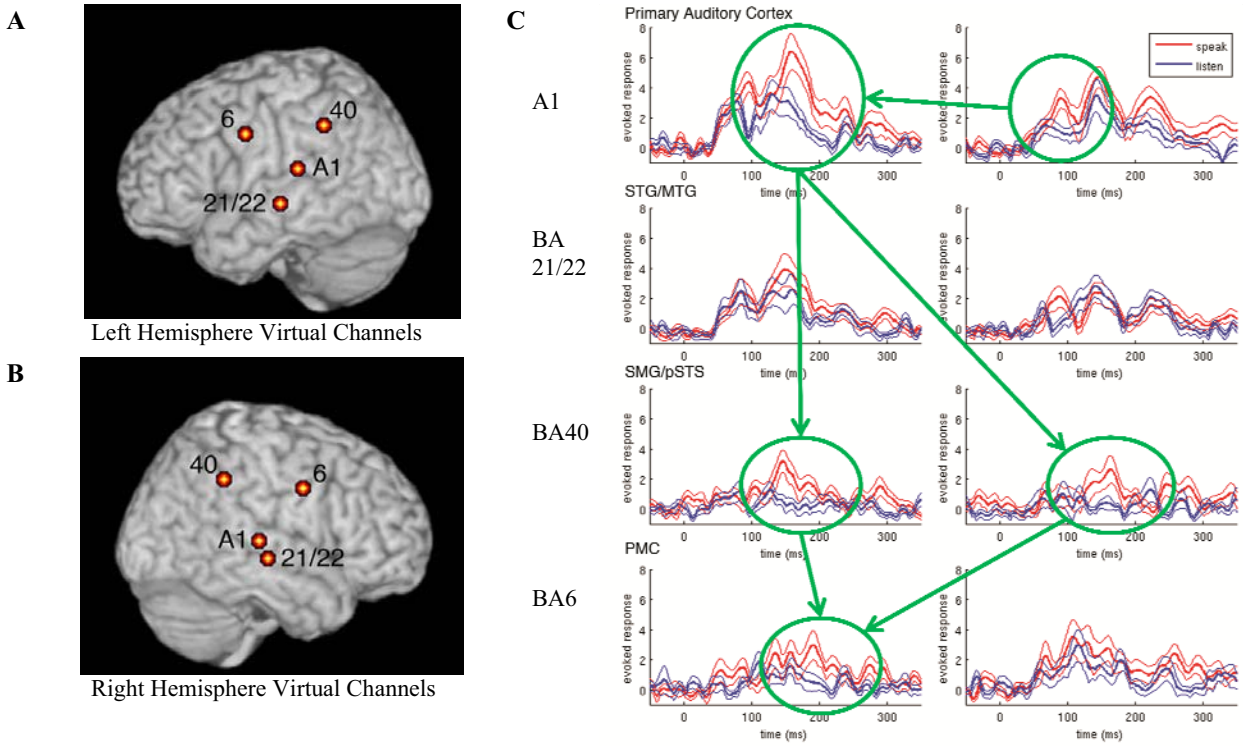

FIGURE 3. MEG study results. **A, B**: Locations for vitual channel analysis of MEG data: premotor cortex (BA6), the inferior parietal lobe (BA40), the superior-to-middle temporal gyrus (BA21/22), and primary auditory cortex (A1). Except for A1, which was localized from responses to 1 KHz tones, the virtual channels were set at the centroids of their respective Brodmann areas. **C**: Neural responses to pitch perturbation. Plots show timecourse of evoked RMS power for each condition (**red**: speak, **blue**: listen) in each virtual channel location (rows) in each hemisphere (left column: left hemisphere, right column: right hemisphere), with perturbation onset at 0ms. **Green ovals**: intervals exhibiting SPRE in the plots; **Green arrows**: sequence of SPRE activations suggested by the results.

Results across a group of subjects (n=10) are summarized in Figure 3C, which shows the timecourses of evoked power, averaged over subjects, at the virtual channel locations in response to the onset of the pitch perturbation. The plots show that many of the virtual channels exhibit intervals where the response in the speaking condition (**red**) is enhanced compared to the response in the listen condition (**blue**), replicating the study results discussed earlier that found the SPRE effect. But the plots also show that the pattern of SPRE responses is not uniform across cortical areas. In auditory regions, primary auditory cortex exhibits SPRE in both hemispheres, although there is evidence that SPRE in the right hemisphere may be seen earlier (prior to 100ms) than in the left hemisphere, where a large speak condition enhancement is seen between 100 and 200ms. As indicated by the green arrow in the A1 row, this may suggest that right A1 is initially most sensitive to pitch deviations from expected feedback, and passes this feedback mismatch on to left A1. After A1, SPRE is next seen bilaterally in IPL/SMG/pSTS (BA40) cortex (again between 100 and 200ms, post-perturbation), followed finally by a unilateral SPRE response at around 200ms in left PMC (BA6).

In general, this sequence of SPRE responses along what has been called the "dorsal auditory stream" (A1, IPL/SMG/pSTS, PMC) [22, 41, 42] is consistent with the connectivity of cortical areas postulated in the SFC model. The lack of SPRE activity in STG/MTG (BA21/22) is also consistent with the hypothesis that areas of the so-called "ventral auditory stream" are not involved in processing auditory feedback. It is also not surprising that a motor-related area (PMC) would show a greater response to feedback perturbations in the speaking condition, since it's likely to be involved in generating a compensatory response. However, several aspects of the results were surprising and suggest modifications to our SFC model. First, SPRE was hypothesized to arise in higher-order auditory cortex, not A1 (although the results do show greater SPRE, as a percent of listen activity, in higher-order areas). Also, the model has been agnostic about the cortical laterality of its various operations, but a distinct pattern of laterality was seen: SPRE in higher auditory cortex (IPL/SMG/pSTS) was bilateral, yet the motor response SPRE in PMC was left-lateralized.

CORTICAL RESPONSES TO ALTERED PITCH FEEDBACK: AN ECOG STUDY

We were able to investigate some of the issues raised by the MEG study in more detail in a recent study of pitch perturbation responses that used electrocorticography (ECoG) [38]. Patients with intractable forms of epilepsy often having electrocortical grids (ECoGs) temporarily implanted in their brains for the clinical purpose of localizing epileptogenic foci. In each patient, the grid is implanted across a wide area of the cortical hemisphere suspected of harboring the epileptogenic source, and cortical activity is recorded from the grid over an extended period of time, in order to record activity during a seizure event. The period of recording time can often take up to a week, during which time the patient is often willing to participate in psychophysical experiments. In this situation, we have been able to run several patients in the pitch feedback perturbation experiment we had previously run with MEG.

Relationship between SIS and SPRE

Figure 4 shows results for a single subject. Four channels' activity are highlighted to illustrate several findings. First, the spectrogram plots are averages across the spectrograms of each trial, and they show that there are distinctly different signatures of neural activity change in different frequency bands. For example, electrode 22's response to speech onset shows up as a power increase in the high gamma (50-150 Hz) band and a power decrease in the beta (13-30 Hz) band. As indicated by the line plots next to the spectrograms, we chose to focus our analysis on high gamma band power changes. The plots for this band show that responses to speech onset varied across electrodes, with auditory electrodes (21, 22, 23) showing varying degrees of speaking-induced suppression (SIS), and the premotor electrode (45) showing "anti-SIS", with greater activity in the speak condition than the listen condition, especially prior to speech onset. Motor activity preceding actual speech onset is as expected, and finding varying degrees of SIS across auditory electrodes is consistent with SFC model predictions if one assumes some auditory regions are more feedback-related, and compare auditory input with efference-copy-derived predictions, while other regions are not involved in processing feedback. As discussed above, this was also found in the MEG data.

In considering these same electrodes' responses to the pitch perturbation, it is again not surprising to see a SPRE response in the premotor electrode (i.e., an enhanced response to the perturbation during speaking as compared with passive listening), since this would be the expected activity of a motor region generating production changes to compensate for the feedback perturbation. Nor is it surprising to see SPRE responses in the auditory electrodes, although it is interesting to see increasing SPRE as one moves dorsally from electrode 21 to 23. However, comparison of the auditory electrodes' SIS and SPRE responses, taken together, pose a challenge to the SFC model,

as it is described above and shown in Figure 1. Electrodes 21 and 23 have both SIS and SPRE responses, which the model would account for if these represented high-order auditory areas controlling the gain $\mathbf{K}_t(\tilde{\mathbf{y}})$ on feedback prediction errors (the mechanism responsible for SPRE), with the feedback prediction error signals conveying the SIS effect from more primary areas where feedback is compared. By this account, electrode 21 would seem to be over a more primary area, since it has a strong SIS response but very little SPRE. Electrodes 21 and 23 illustrate two possibilities explained by the model: areas showing SIS but not SPRE, and areas showing SIS and SPRE. Electrode 22, on the other hand, demonstrates response properties not predicted by the model: a SPRE response with no SIS response. In the model, high-order auditory areas showing SPRE are all connected to lower areas showing SIS. In order to show only SPRE, a high-order area would have to be connected to a lower area that merely reports changes in auditory input, but does not compare them with a prediction.

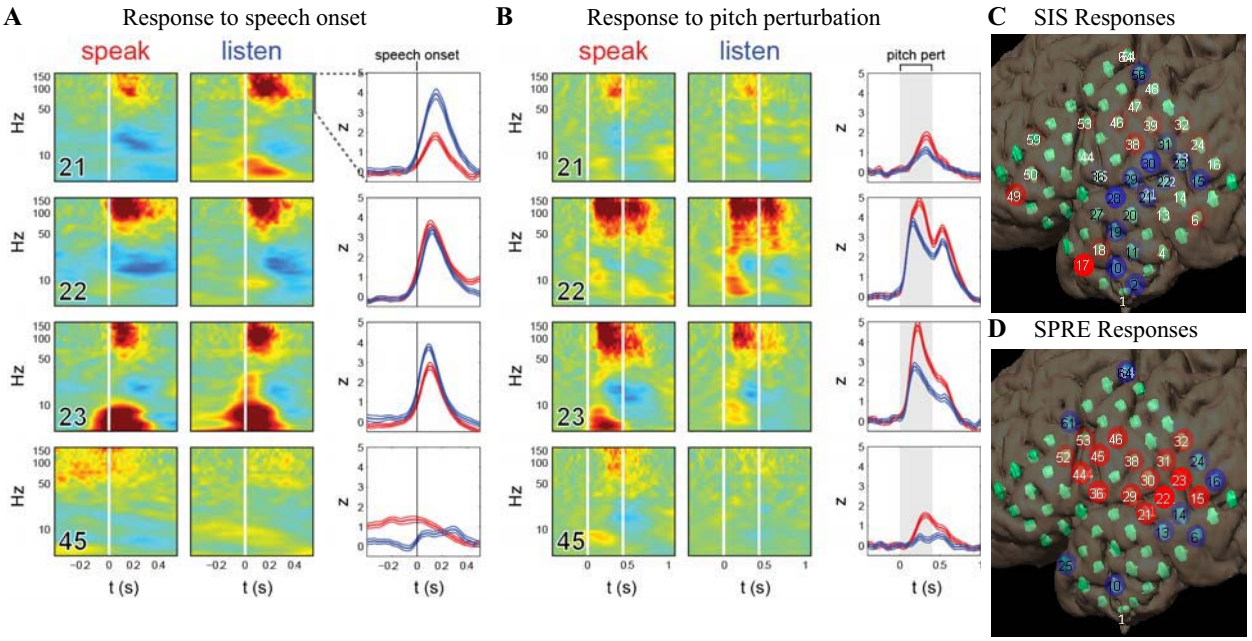

FIGURE 4. ECoG study results for a single subject with a left hemisphere grid. **A**: Responses to speech onset at three auditory electrodes (21, 22, 23) that ascend the dorsal auditory pathway, and one ventral premotor electrode (45) (see grid maps in parts C and D for electrode locations). First two columns show spectrograms of responses around speech onset (at 0 ms) in the speak and listen conditions, while third column focuses on power changes over time in the high gamma (50-150 Hz) band in the speak (**red**) and listen (**blue**) conditions. **B**: A similar set of plots for these same electrodes responses around onset (at 0 ms) of the pitch perturbation. **C**: Grid layout (1cm electrode spacing) showing pattern of speaking-induced suppression (SIS) responses at speech onset. **Blue** electrodes show SIS (speak response < listen response), while **red** electrodes show anti-SIS (speak response > listen response), with transparency indicating relative intensity of SIS or anti-SIS. **D**: Grid layout showing pattern of SPRE (**red**) and anti-SPRE (**blue**) responses.

The grid layouts in figures 4C and D show the there were also other electrodes that, like electrode 21, show predominantly SPRE responses and minimal or no SIS responses. This can be seen not only at motor/frontal sites, where the lack of a response in the listen condition is expected, but also at other sites in the temporal lobe and the IPL. This pattern was also seen across all subjects. A histogram analysis found 57 electrodes exhibiting SPRE and of these, only 10 also exhibited SIS. It is true that those electrodes showing both SIS and SPRE were all in the temporal lobes and IPL, but there were many more electrodes in these non-motor regions that showed only SPRE. The results suggest that the model should be modified to allow high-order auditory areas to receive input from a mix of low order areas, some conveying feedback prediction errors, and others conveying only auditory input changes. A strong prediction of the model that remains, however, is that SPRE electrodes not showing SIS should only be lacking SIS because the response to speech onset is equal in both speaking and listening conditions. In other words, the model predicts a cortical region cannot have SPRE without a response to speech onset, and indeed the data bear out this prediction.

Relationship between Compensation and SPRE

A final prediction of the model to be considered here concerns how different electrodes' response characteristics should indicate the directness of their role in generating compensatory responses to feedback alterations. The model postulates that premotor areas, which update the state estimate $\hat{\mathbf{x}}_t$ passed to the control law $\mathbf{U}_t(\hat{\mathbf{x}}_t)$ in motor cortex, are the most proximal to generation of the compensatory motor actions. Premotor area activations, therefore, should be highly correlated with compensatory output. In auditory regions, high-order areas are postulated to contain the

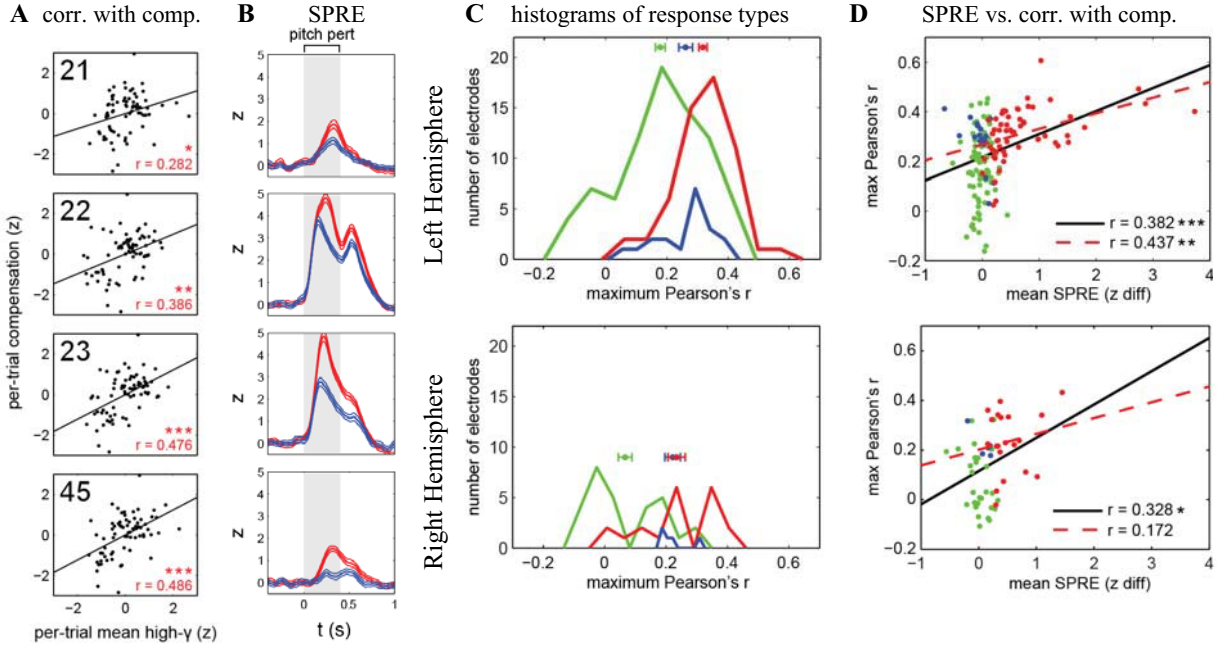

FIGURE 5. Relationship between SPRE and correlation with compensation. **A**: Correlation with compensation. Column show scatterplots of compensation vs. high gamma power for each of the electrodes of the subject shown in Figure 4. Each black dot represents a trial in the speak condition. For each trial in each plot, the peak compensation for the pitch perturbation and mean high gamma response power are measured in z-scores. In each scatterplot, a fit line shows the correlation with compensation, with the r-value of the correlation, as well as its significance, shown in **red** at the bottom right of each plot (* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$). **B**: For reference, high gamma power of responses to the pitch perturbation in the speak and listen conditions for the electrodes of column A (i.e., a repeat of Figure 4B) are also shown. **C**: Histograms of temporal electrode response types across all subjects' data (**green** = no response to perturbation; **blue** = same response to perturbation in both speak and listen conditions; **red** = significant SPRE response), with each electrode binned by how correlated its mean high gamma response in the speak condition was with compensation across trials. Mean compensation (+/- standard error) for each response type are shown by the colored intervals above the histograms. Upper and lower histogram show left and right hemisphere electrodes, respectively. D: Regression analysis of (SPRE) vs. (correlation with compensation). Plots show scatterplots of SPRE versus correlation with compensation for the temporal electrode data from all subjects. Each colored dot shows an electrode's mean SPRE (difference in z-scores between speak and listen responses to perturbation) (horizontal axis) versus the correlation between its high gamma response and compensation across trials in the speak condition (vertical axis). Dot color corresponds to the same response types shown in column C. Two fit lines are shown in each plot: black line shows the correlation of (SPRE) vs. (correlation with compensation) for all temporal electrodes; red line shows same, but just considering the electrodes with significant SPRE (red dots). R-values and significance of each of these correlations are also listed in the plots (again, * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$, and a lack of a * indicating lack of significance). As with column C, upper and lower plots show left and right hemisphere electrodes, respectively.

Kalman gain function $\mathbf{K}_t(\tilde{\mathbf{y}})$ that maps the feedback prediction errors $\tilde{\mathbf{y}}_{t\text{-}\hat{N}}$ determined in lower areas to the state corrections $\hat{\mathbf{e}}_t$ used by premotor cortex to update the state estimate. This results in two predictions: (1) higher auditory areas exhibit SPRE while lower areas exhibit SIS, and (2) higher auditory areas are more proximal to the generation of compensatory motor actions than lower areas, meaning that higher, SPRE-showing areas' activations

should be more correlated with compensatory output than lower, SIS-showing areas' activations would be. Taken together, these last two predictions suggest that the degree to which a temporal electrode exhibits SPRE should be related to how much its activation predicts compensation. As shown in Figure 5, this relationship was indeed found to be the case in the ECoG data.

Figures 5A and B show further analysis of the electrodes shown in Figure 4. Figure 5A shows how well each electrode's mean high gamma response to the pitch perturbation is correlated with the compensation produced by the subject on each trial. Figure 5B repeats the plots from Figure 4 that show each electrode's SPRE response. Taken together, the plots suggest a positive relationship between the trial-to-trial correlation with compensation for the pitch perturbation and its degree of SPRE: as one moves down the rows of plots in A and B, both correlation with compensation and SPRE are increasing. This result is also seen in the data across all subjects. Figure 5C shows that, across all subjects' data, left temporal electrodes showing SPRE had significantly higher mean correlation with compensation than did electrodes with other response types. A similar trend was evident in the right temporal lobe. Figure 5D presents this result in more detail. For each electrode, mean SPRE was calculated as difference between mean speak and listen responses (both measured in z-scores) and plotted against that electrode's correlation with compensation. Fit lines through the resulting scatter data showed a highly significant relationship between SPRE and correlation with compensation in the left temporal lobe, and a significant relationship in the right temporal lobe. (Lack of strong statistical significance in the right hemisphere data is likely due to the smaller number of electrodes' data analyzed in this hemisphere.)

SUMMARY AND CONCLUSIONS

The concept of state feedback control (SFC) is a powerful and flexible model of motor control, and many current models of speech motor control can be described as examples of SFC. Here, we have considered an SFC model of speech motor control with a very general form, and found it can account for many of the known characteristics of the role of auditory feedback in the control of speech. This SFC model can also explain many of the phenomena observed in studies of the neural processing of auditory feedback. Our own recent experiments' results suggest ways the model should be extended, but nevertheless largely confirm further predictions from the model.

ACKNOWLEDGMENTS

This work was supported by grants from NSF (grant BCS-0926196) and NIH (grant R01-DC010145).

REFERENCES

1. Tomblin, J.B., et al., *Long-term trajectories of the development of speech sound production in pediatric cochlear implant recipients.* J Speech Lang Hear Res, 2008. **51**(5): p. 1353-68.
2. Cowie, R. and E. Douglas-Cowie, *Postlingually acquired deafness: Speech deterioration and the wider consequences*1992, Hawthorne, NY, US: Mouton de Gruyter. x, 304.
3. Lane, H. and B. Tranel, *The Lombard sign and the role of hearing in speech.* Journal of Speech and Hearing Research, 1971. **14**(4): p. 677-709.
4. Chang-Yit, R., J. Pick, Herbert L., and G.M. Siegel, *Reliability of sidetone amplification effect in vocal intensity.* Journal of Communication Disorders, 1975. **8**(4): p. 317-324.
5. Yates, A.J., *Delayed auditory feedback.* Psychological Bulletin, 1963. **60**(3): p. 213-232.
6. Burnett, T.A., et al., *Voice F0 responses to manipulations in pitch feedback.* Journal of the Acoustical Society of America, 1998. **103**(6): p. 3153-61.
7. Jones, J.A. and K.G. Munhall, *Perceptual calibration of F0 production: evidence from feedback perturbation.* Journal of the Acoustical Society of America, 2000. **108**(3 Pt 1): p. 1246-51.
8. Heinks-Maldonado, T.H. and J.F. Houde, *Compensatory responses to brief perturbations of speech amplitude.* Acoustics Research Letters Online, 2005. **6**(3): p. 131-7.
9. Houde, J.F. and M.I. Jordan, *Sensorimotor adaptation in speech production.* Science, 1998. **279**(5354): p. 1213-6.
10. Purcell, D.W. and K.G. Munhall, *Compensation following real-time manipulation of formants in isolated vowels.* Journal of the Acoustical Society of America, 2006. **119**(4): p. 2288-97.
11. Shiller, D.M., et al., *Perceptual recalibration of speech sounds following speech motor learning.* Journal of the Acoustical Society of America, 2009. **125**(2): p. 1103-13.

12. Guenther, F.H. and T. Vladusich, *A Neural Theory of Speech Acquisition and Production.* J Neurolinguistics, 2012. **25**(5): p. 408-422.
13. Tian, X. and D. Poeppel, *Mental imagery of speech and movement implicates the dynamics of internal forward models.* Front Psychol, 2010. **1**: p. 166.
14. Price, C.J., J.T. Crinion, and M. Macsweeney, *A Generative Model of Speech Production in Broca's and Wernicke's Areas.* Frontiers in Psychology, 2011. **2**: p. 237.
15. Stengel, R.F., *Optimal Control and Estimation*1994, Mineola, NY: Dover Publications, Inc.
16. Todorov, E., *Optimal control theory*, in *Bayesian Brain: Probabilistic Approaches to Neural Coding*, K. Doya, et al., Editors. 2006, MIT Press: Cambridge, MA. p. 269-98.
17. Jacobs, O.L.R., *Introduction to Control Theory*. 2nd ed1993, Oxford, UK: Oxford University Press. 390.
18. Todorov, E. and M.I. Jordan, *Optimal feedback control as a theory of motor coordination.* Nat Neurosci, 2002. **5**(11): p. 1226-35.
19. Shadmehr, R. and J.W. Krakauer, *A computational neuroanatomy for motor control.* Experimental Brain Research, 2008. **185**(3): p. 359-81.
20. Guigon, E., P. Baraduc, and M. Desmurget, *Optimality, stochasticity, and variability in motor behavior.* J Comput Neurosci, 2008. **24**(1): p. 57-68.
21. Houde, J.F. and S.S. Nagarajan, *Speech production as state feedback control.* Frontiers in Human Neuroscience, 2011. **5**: p. 82.
22. Hickok, G., J.F. Houde, and F. Rong, *Sensorimotor integration in speech processing: computational basis and neural organization.* Neuron, 2011. **69**(3): p. 407-22.
23. Hirano, S., et al., *Cortical activation by monaural speech sound stimulation demonstrated by positron emission tomography.* Exp Brain Res, 1997. **113**(1): p. 75-80.
24. Hirano, S., et al., *Cortical processing mechanism for vocalization with auditory verbal feedback.* Neuroreport, 1997. **8**(9-10): p. 2379-82.
25. Hirano, S., et al., *Cortical speech processing mechanisms while vocalizing visually presented languages.* Neuroreport, 1996. **8**(1): p. 363-7.
26. Ford, J.M. and D.H. Mathalon, *Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking.* Journal Of Psychiatric Research, 2004. **38**(1): p. 37-46.
27. Ford, J.M., et al., *Neurophysiological evidence of corollary discharge dysfunction in schizophrenia.* American Journal of Psychiatry, 2001. **158**(12): p. 2069-2071.
28. Numminen, J. and G. Curio, *Differential effects of overt, covert and replayed speech on vowel- evoked responses of the human auditory cortex.* Neuroscience Letters, 1999. **272**(1): p. 29-32.
29. Numminen, J., R. Salmelin, and R. Hari, *Subject's own speech reduces reactivity of the human auditory cortex.* Neuroscience Letters, 1999. **265**(2): p. 119-22.
30. Curio, G., et al., *Speaking modifies voice-evoked activity in the human auditory cortex.* Human Brain Mapping, 2000. **9**(4): p. 183-91.
31. Houde, J.F., et al., *Modulation of the auditory cortex during speech: An MEG study.* J Cogn Neurosci, 2002. **14**(8): p. 1125-38.
32. Heinks-Maldonado, T.H., S.S. Nagarajan, and J.F. Houde, *Magnetoencephalographic evidence for a precise forward model in speech production.* Neuroreport, 2006. **17**(13): p. 1375-9.
33. Ventura, M.I., S.S. Nagarajan, and J.F. Houde, *Speech target modulates speaking induced suppression in auditory cortex.* BMC Neuroscience, 2009. **10**: p. 58.
34. Greenlee, J.D., et al., *Human Auditory Cortical Activation during Self-Vocalization.* PLoS One, 2011. **6**(3): p. e14744.
35. Behroozmand, R., et al., *Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation.* Clin Neurophysiol, 2009. **120**(7): p. 1303-12.
36. Zheng, Z.Z., K.G. Munhall, and I.S. Johnsrude, *Functional Overlap between Regions Involved in Speech Perception and in Monitoring One's Own Voice during Speech Production.* Journal of Cognitive Neuroscience, 2010. **22**(8): p. 1770-81.
37. Katseff, S., J. Houde, and K. Johnson, *Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback?* Language and Speech, 2012. **55**(2): p. 295-308.
38. Chang, E.F., et al., *Human cortical sensorimotor network underlying feedback control of vocal pitch.* Proceedings of the National Academy of Sciences, 2013. **In Press.**
39. Owen, J.P., et al., *Performance evaluation of the Champagne source reconstruction algorithm on simulated and real M/EEG data.* Neuroimage, 2012. **60**(1): p. 305-23.
40. Robinson, S. and J. Vrba, *Functional neuroimaging by synthetic aperture magnetometry (SAM)*, in *Recent advances in Biomagnetism*, Yoshimoto and e. al., Editors. 1999, Tokyo University Press: Sendai. p. 302-305.
41. Rauschecker, J.P. and S.K. Scott, *Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing.* Nat Neurosci, 2009. **12**(6): p. 718-24.
42. Hickok, G. and D. Poeppel, *The cortical organization of speech processing.* Nature Reviews Neuroscience, 2007. **8**(5): p. 393-402.