



Opposing and following responses in sensorimotor speech control: Why responses go both ways

Matthias K. Franken^{1,2,3}  · Daniel J. Acheson^{1,2} · James M. McQueen^{2,4} · Peter Hagoort^{1,2} · Frank Eisner⁴

Published online: 4 June 2018
© Psychonomic Society, Inc. 2018

Abstract

When talking, speakers continuously monitor and use the auditory feedback of their own voice to control and inform speech production processes. When speakers are provided with auditory feedback that is perturbed in real time, most of them compensate for this by opposing the feedback perturbation. But some responses follow the perturbation. In the present study, we investigated whether the state of the speech production system at perturbation onset may determine what type of response (opposing or following) is made. The results suggest that whether a perturbation-related response is opposing or following depends on ongoing fluctuations of the production system: The system initially responds by doing the opposite of what it was doing. This effect and the nontrivial proportion of following responses suggest that current production models are inadequate: They need to account for why responses to unexpected sensory feedback depend on the production system's state at the time of perturbation.

Keywords Speech production · Auditory feedback · Speech perception · Pitch

An important aspect of action control is performance monitoring through sensory feedback. Such control allows us to confirm an appropriate action plan, adapt to a changing environment, or learn from our mistakes (Wolpert & Ghahramani, 2000). For example, when throwing a ball, feedback could show us that the throw was successful, or it could indicate the need for adaptation to new conditions (e.g., if the wind suddenly shifts). Similar processes are at play in speech and musical production.

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13423-018-1494-x>) contains supplementary material, which is available to authorized users.

✉ Matthias K. Franken
matthias.franken@ugent.be

- ¹ Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands
- ² Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
- ³ Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium
- ⁴ Donders Centre for Cognition, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

The importance of auditory feedback during speech production has been established in altered-auditory-feedback experiments (Burnett, Freedland, Larson, & Hain, 1998; Houde & Jordan, 1998). For example, Burnett et al. manipulated the pitch frequency of speakers' auditory feedback in real time. Typically, speakers respond by adjusting the pitch in their output in the opposite direction from the pitch shift (Burnett et al., 1998; Liu & Larson, 2007). Speakers thus compensate for unexpected changes in auditory feedback so that their actual output more closely matches their intended output. We argue here that there is more to sensorimotor control than these opposing responses.

Vocal motor control is a noisy process and needs constant feedback monitoring. In both speech (Akagi, Iwaki, & Minakawa, 1998) and singing (Akagi & Kitakaze, 2000), pitch production is not constant at the target pitch level, but fluctuates around the target. Pitch fluctuations are in fact an identifying feature of human pitch control, and removing them makes synthesized speech or song sound robotic (Akagi et al., 1998). These fluctuations are maintained by constant feedback monitoring and subsequent updating of the vocal motor commands. Using a vowel production task, Niziolek, Nagarajan, and Houde (2013) showed that the auditory feedback control system is indeed sensitive to small deviations in vocal production.

Feedback monitoring and subsequent compensation is accounted for in several theoretical frameworks (Guenther, Ghosh, & Tourville, 2006; Hickok, 2012; Houde & Nagarajan, 2011). These theories hypothesize internal forward models, which predict the sensory (e.g., auditory) consequences of actions in real time. These predictions are compared with the incoming auditory feedback. A mismatch causes the speech motor system to initiate corrective (i.e., compensating) motor commands.

In contrast to these model-based predictions, however, several studies have reported that sometimes, instead of feedback compensation, responses are observed that follow the direction of the altered feedback (Burnett et al., 1998; Hain et al., 2000; Larson, Sun, & Hain, 2007). These following responses are less frequent than opposing (i.e., compensating) responses, and are usually reported at the subject level. However, looking at single trials, Behroozmand, Korzyukov, Sattler, and Larson (2012) showed that even subjects who show an opposing response on average may show following responses on some trials. Therefore, the focus on the average response may have obscured the field's view of the nature of following responses (Behroozmand et al., 2012).

Following responses have led some authors to suggest that voice pitch control has two feedback modes: one for tracking an external referent (eliciting following responses) and another for correcting for internal disturbances (eliciting opposing responses; Burnett et al., 1998; Hain et al., 2000). For example, in the context of choir singing, a singer might follow the pitch of another as an external referent (e.g., go flatter if the fellow singer is singing flat). However, the feedback signal of one's own voice should activate the feedback mode for internal disturbances and therefore lead to an opposing response (e.g., go sharper when you're singing flat). Both feedback modes may thus be simultaneously active.

The small number of studies that have examined following responses suggest that such responses occur more often when the pitch manipulation is larger and that they have a shorter duration than opposing responses (Burnett et al., 1998). Behroozmand et al. (2012) showed that predictable altered feedback may encourage a tendency to follow the feedback.

Here we investigated what factors play a role in feedback-based pitch control. Participants tried to match a pitch target while vocalizing. They received auditory feedback through headphones, which sometimes was unexpectedly pitch-shifted for 500 ms. None of the participants were aware of the pitch shift. We expected participants on average to compensate for the feedback, but at the single-trial level to sometimes follow and sometimes oppose the shift. The opposing/following balance might depend not only on whether the perturbation is considered a self-generated speech error or on how large the perturbation is, but also on the state of the system at the time of the perturbation. In Experiment 1, we therefore explored whether system-

internal constraints limit how it can respond to a perturbation. If so, the system's ongoing pre-perturbation pitch fluctuations should be predictive of the response.

Experiment 1

Method and materials

All of the raw data as well as major analysis scripts can be accessed online (raw data: <https://osf.io/pgrbe/>; analysis: doi: 10.17605/OSF.IO/DAR54).

Participants Thirty-nine healthy volunteers (age: $M = 22$ years, $SD = 3.6$; 27 females, 12 males) participated after providing written informed consent as approved by the local ethics committee (CMO region Arnhem/Nijmegen). All participants had normal hearing, were native speakers of Dutch, and had no history of speech and/or language pathology. The sample size was based on a power analysis of the magnetoencephalography (MEG) connectivity effects in Ford, Gray, Faustman, Heink, and Mathalon (2005), indicating that one would need about 33 participants ($d_z = 0.506$, power = 80%). We therefore planned to test at least 36 participants. This sample size exceeded those in most related previous studies (e.g., Behroozmand et al., 2012).

Paradigm Participants performed a tone-matching task while their brain activity was measured using MEG. The MEG data will not be presented here. A trial started with a short tone (duration of 300 ms). Subsequently, a visual cue ("EE") instructed the participants to start vocalizing the Dutch vowel /e/ for the duration of the cue (3 s), while trying to match the pitch of the tone they had just heard.

The participants were recorded, and the recorded voice signal was used to provide online auditory feedback. In half of the trials, participants received normal auditory feedback (control trials). In the others (perturbation trials), the auditory feedback was normal at first, but starting 500–1,500 ms after speech onset, the feedback pitch was increased by 25 cents for a duration of 500 ms before it returned to normal feedback for the remainder of the trial. Most previous studies had used shorter pitch shifts, to avoid volitional responses. For the purposes of the MEG analyses, we used longer perturbations, but they were smaller in pitch so that participants would not be aware of them. Participants received 99 perturbation and 99 control trials, randomly mixed across two blocks of 99 trials.

Stimuli The stimuli were pure tones with one of three frequencies. The frequencies were individually tailored to be four, eight, and 11 semitones above the participant's average pitch, as determined through five practice vocalizations.

The auditory feedback shifts were performed using the phase vocoder implemented in the Audapter software (Cai, Boucek, Ghosh, Guenther, & Perkell, 2008).

All recordings were made using a Sennheiser ME64 cardioid microphone, which was set up in a magnetically shielded room and connected through an audio mixer to a dedicated Motu MicroBook II soundcard outside the room. Auditory feedback was delivered through the same soundcard, which was connected to CTF audio air tubes. Stimulus presentation was controlled by a Windows computer running Audapter¹ and The MathWorks' Matlab environment.

Analysis For every trial, the pitch of the participant's vocalization was determined using the autocorrelation method in Praat (Boersma & Weenink, 2013). The pitch contours were exported to Matlab for further processing.

The pitch contours were epoched from 500 ms before to 1,000 ms after perturbation onset. The data were detrended and converted from hertz to the cent scale using the following formula:

$$F0 [\text{cents}] = 1,200 * \log_2 \left(\frac{F}{F_{\text{baseline}}} \right).$$

Here, F is the original pitch frequency in hertz, whereas F_{baseline} is the average pitch frequency in hertz across a baseline window (−200 to 0 ms before perturbation onset). Trials that contained artifacts (sharp discontinuities or a failure of the algorithm to estimate a pitch contour) were removed from the analysis.

Pitch contours in the control and perturbation trials were compared using a cluster-based permutation test (Maris & Oostenveld, 2007). Samples for which the contrast perturbation–control exceeded an uncorrected α level of .05 were temporally clustered. Cluster-level statistics were calculated by summing the t statistics. Next, a permutation distribution of statistics was calculated by randomly exchanging trials between the conditions and calculating the maximal positive and negative cluster-level statistics for each of 1,000 permutations. The observed cluster-level statistic was tested against the permutation distribution.

Each trial was classified as having an opposing or a following response (as in Behroozmand et al., 2012). Two different methods were used for classification. In the first, the point was determined within the time window of 60–400 ms after perturbation onset at which the average of the *rectified* pitch contours was maximal (*point of maximal deviation*). Linear regression was performed on the single-trial data between 60 ms after perturbation onset and the point of maximal

deviation. If the slope of the linear fit was positive, the response was classified as “following”; if the slope was negative, the response was classified as “opposing.” For the distribution of the slopes of the linear fits, see the [supplementary materials](#). No threshold was applied for the slope to be significantly different from 0, and additional analyses leaving out the trials with slopes near 0 yielded similar results (see the [supplementary materials](#)).

The second classification method was based on the Castellan change-point test (Siegel & Castellan, 1988). This change-point test yields the K statistic:

$$K = |2W_j - j(N + 1)|, j = 1, 2, \dots, N - 1.$$

Here, W_j is the cumulative sum of the ranks at a sample number j , and N is the total number of samples. We calculated K for every trial over the time window of 0–300 ms after perturbation onset. The point at which K is maximal is the *change point*. If $2W_j - j(N + 1)$ at that point was positive, the trial was classified as “opposing”; if it was negative, the trial was classified as “following.”

If the two methods did not yield the same classification for a particular trial, classification was determined through visual inspection of the pitch contour (this occurred on average in 23.0% of a participant's data, range = 7.2%–43.3%). If there was no clear response, the trial was excluded from further analysis (across participants, 7.9% (range = 2%–16.3%) of trials). The classification procedure was performed on both the perturbation and control trials. Because there was no perturbation onset in the control trials, random time points were chosen, while making sure that their distribution across trials was equal to the distribution of perturbation onsets within the same participant.

To examine how participants' responses depended on the state of their voice motor system at the moment of perturbation, the slope and the average F0 value over the 100 ms before perturbation onset were determined for each trial type.

Another way to identify differences between opposing and following trials would be to compare the magnitude and latency of the responses. The peak response was identified for each response type in each participant by subtracting the average control response from the average opposing or following response. The response latency was then quantified as the point in time between 50 and 500 ms at which the difference was largest.

Results

Overall, participants compensated for the pitch increase in the perturbation trials by lowering their pitch (Fig. 1a). The pitch contour in the perturbation trials differed from that is the control trials ($p = .002$). This difference was mainly driven by a component lasting from 144 to 765 ms after perturbation onset.

¹ In the Audapter software, a pitch shift is sometimes accompanied by an unintended small intensity change. These intensity changes do not have bearing on the present results, given their small magnitude and the fact that they did not always occur. See the supplementary materials for a detailed analysis.

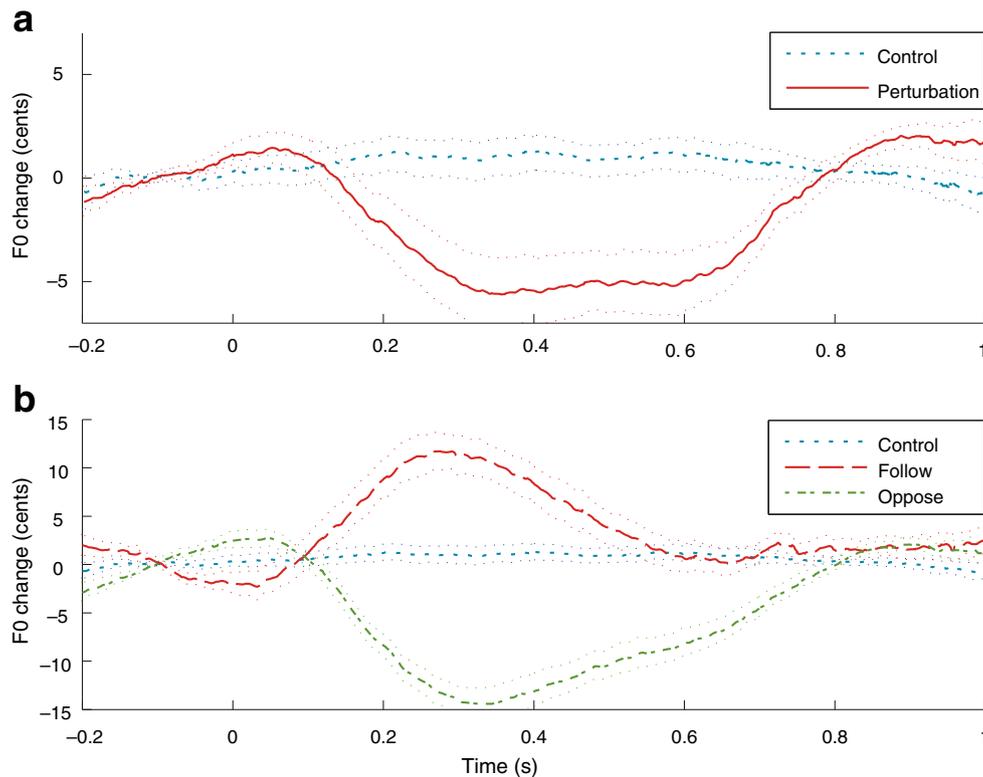


Fig. 1 Grand averages comparing changes in F0 in perturbation and control trials in Experiment 1, time-locked to perturbation onset (a). The perturbation trial contour is collapsed over opposing and following responses. In the perturbation trials, perturbation starts at 0 s and lasts

until 0.5 s. (b) Grand averages for the perturbation trials, split up into following and opposing trials. Thin dotted lines represent the 95% confidence intervals

We then classified each perturbation trial as either following or opposing (or neither). The same classification was performed on the control trials. The distribution of opposing and following trials (Fig. 2) shows a clear effect of perturbation: In the control trials, the proportion of trials classified as opposing is about 50%, reflecting random fluctuations of the pitch

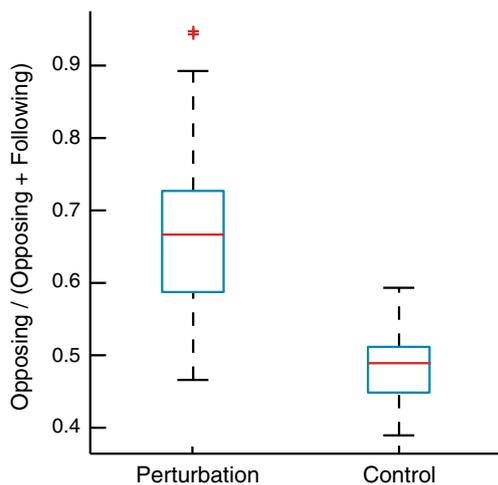


Fig. 2 Proportions of trials in Experiment 1 classified as opposing in the perturbation and control conditions. The y-axis shows the number of trials classified as opposing divided by the total sum of trials that were classified as either opposing or following

contour, whereas in the perturbation trials, the proportion of opposing trials is larger [$t(38) = 8.16, p < .001, CI = [0.14, 0.23]$, Cohen’s $d = 1.96$], ranging from just under 50% to over 90%. Participants thus followed the feedback perturbation in a nontrivial number of trials (10%–50%).

As expected, the pitch contour in opposing trials differed from that in following trials (Fig. 1b, $p = .002$). This was mainly driven by a component (with opposing responses having a smaller magnitude than following responses) from 108 to 812 ms after perturbation onset, but also by a smaller difference in the opposite direction (following < opposing), from 91 ms before perturbation onset until 77 ms after perturbation onset. This suggests that before perturbation onset, the pitch contours in opposing and following trials already differed.

The pitch contour in the opposing perturbation trials differed from that in the control trials classified as opposing (Fig. 3a, $p = .002$). This effect was driven by a component (arrows in Fig. 3c, 213–712 ms) in which pitch was lower in the perturbation trials, and by a later component (from 791 ms) in which pitch was higher in the perturbation than in the control trials. A similar pattern was found for the following trials (Figs. 3b and 3d, $p = .002$), in which the perturbation trials had lower pitch than the control trials from 338 until 723 ms after perturbation onset (arrows in Fig. 3d). So, even for the following trials, the pitch was lower than in the similarly classified

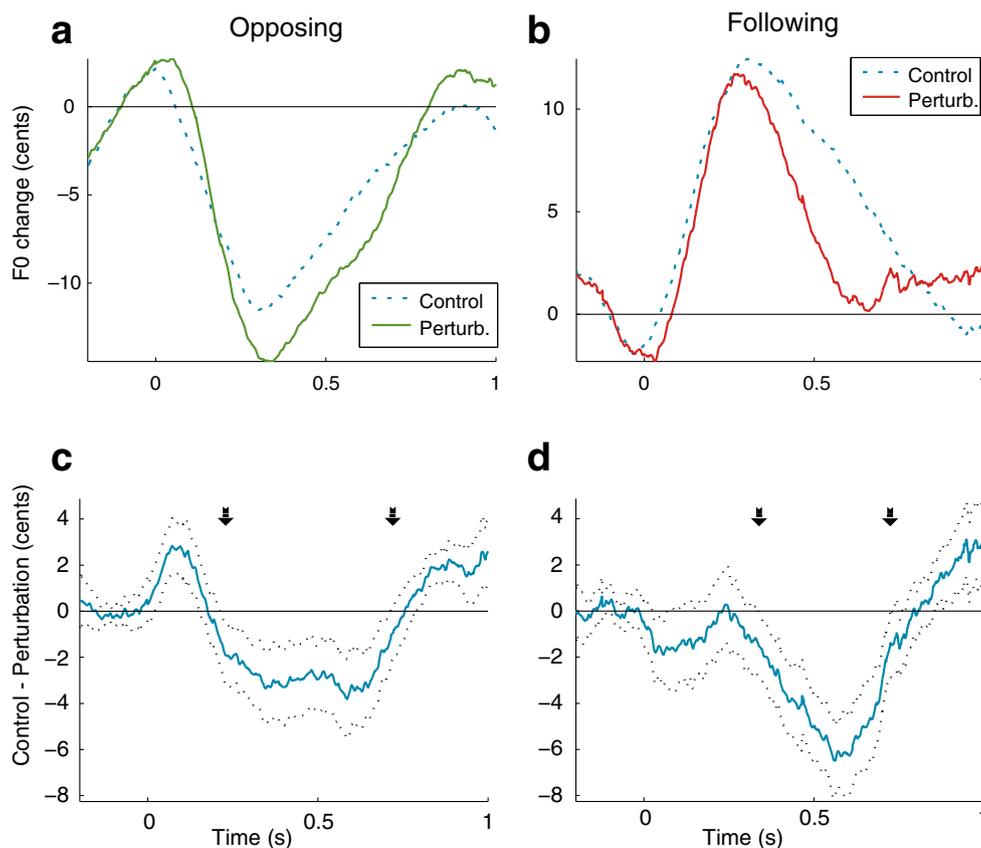


Fig. 3 Comparison of F0 changes between the perturbation and control condition for similarly classified trials in Experiment 1. (a) F0 changes for opposing perturbation trials and control trials classified as opposing. (b) F0 changes for following perturbation trials and control trials classified as

control trials. This means that the pitch increase (Fig. 3b) may not entirely be indicative of a following response, but may also (or instead) reflect an ongoing F0 fluctuation with an additional, smaller opposing response.

The small early difference between following and opposing perturbation trials (Fig. 1b) suggested a difference even before perturbation onset. The results of Wilcoxon signed rank tests showed that both the pitch slope ($z = -4.24, p < .001, r = .48$) and the average value ($z = -5.25, p < .001, r = .59$) over the 100-ms time window before perturbation onset differed between following and opposing trials (Fig. 4). This effect was also found continuously across the data, as well as for the trials within each response type (see the [supplementary materials](#)). So the pitch contour before perturbation onset was predictive of the response type that the perturbation then generated, suggesting a dynamic interaction between ongoing pitch production and the feedback perturbation.

With respect to response peaks, following responses on average peaked earlier [$t(38) = 3.66, p < .001, CI = [0.02, 0.08]$, Cohen's $d = 0.74$] and were smaller [$t(38) = 17.11, p < .001, CI = [24.28, 30.80]$, Cohen's $d = 3.91$] than the opposing responses. An earlier and/or smaller response in following trials can be explained as a result of the added effects of the

following. (c, d) Difference waves corresponding to the comparisons in panels a and b; dotted black lines represent 95% confidence intervals. Arrows indicate the onset and offset of the major component driving the statistical difference

perturbation and the following response, resulting in an even larger pitch deviation.

Experiment 2

This experiment was carried out (1) to replicate the findings of Experiment 1 and (2) to extend those findings to a condition in which the perturbations went in the opposite direction. We predicted that the direction of the responses to perturbations would again depend on ongoing fluctuations in the production system at perturbation onset.

Method and materials

Twenty-four new volunteers (age: $M = 23$ years, $SD = 2.8$; 18 females, six males) participated. The sample size was based on a power analysis of the Wilcoxon signed-rank tests between pre-perturbation slope and response type in Experiment 1, indicating that one would need at least 19 subjects ($d_z = 0.9, \alpha = .05, \text{power} = 95\%$). One participant did not speak loudly enough to trigger the perturbations and so was excluded from the analysis.

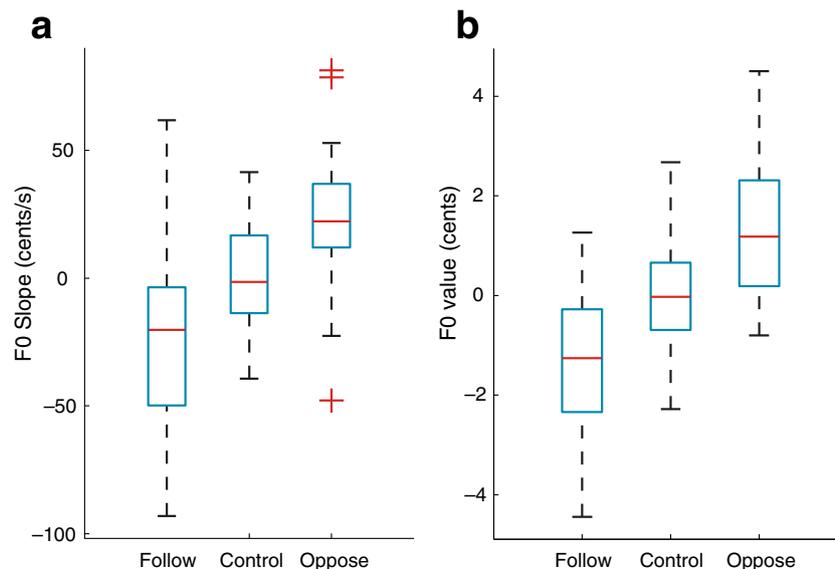


Fig. 4 F0 slope (a) and average F0 value (b) over a 100-ms time window before perturbation onset as a function of trial classification in Experiment 1

The procedure was the same as in Experiment 1, except for the following. The experiment took place in a sound-attenuated booth, and no MEG activity was recorded. Stimuli were delivered through Sennheiser HD 280 PRO headphones. There were two experimental blocks, the order of which was counterbalanced across participants. The upward block was the same as in Experiment 1 (+ 25-cent pitch shifts). In the downward block, all perturbations were – 25-cent shifts.

Results

The results show the same pattern as in Experiment 1. Participants compensated for the pitch increase in the perturbation trials by lowering their pitch in the upward condition and raising their pitch in the downward condition (Fig. 5). The pitch contours in the perturbation trials differed from those in the control trials (upward, $p = .002$; downward, $p = .002$).

Next, individual trials were classified. Figure 6 shows the average pitch contours for opposing and following trials for the downward (Fig. 6a) and upward (Fig. 6b) conditions. The proportion of opposing trials was larger in the perturbation trials than in the control trials (Fig. 7), in both the upward condition [$t(22) = 6.75$, $p < .001$, $CI = [0.12\ 0.23]$, Cohen's $d = 1.41$] and the downward condition [$t(22) = 9.33$, $p < .001$, $CI = [0.18\ 0.28]$, Cohen's $d = 1.95$].

The pitch contour in the opposing perturbation trials differed from that in the control trials classified as opposing (upward, $p = .002$; downward, $p = .006$), with the pitch in perturbation trials being lower than in control trials in the upward condition, and higher than in control trials in the downward condition (Figs. 8a and 8c). The same pattern was observed for the following trials

(upward, $p = .012$; downward, $p = .002$; Figs. 8b and 8d). So, regardless of perturbation direction, both following and opposing trials reflected an ongoing F0 fluctuation with an added opposing response.

In addition, both the pitch slope (upward, $z = -3.77$, $p < .001$, $r = .56$; downward, $z = 3.98$, $p < .001$, $r = .59$) and average value (upward, $z = -3.95$, $p < .001$, $r = .58$; downward, $z = 4.05$, $p < .001$, $r = .60$) preceding perturbation onset differed between following and opposing trials (Fig. 9). This was also observed continuously across the data, and for trials within each response type (see the [supplementary materials](#)).

There were no response-peak latency differences between following and opposing responses [upward, $t(22) = 0.59$, n.s.; downward, $t(22) = 1.80$, n.s.], in contrast to Experiment 1. However, Fig. 8 does suggest that the following responses tend to return to baseline quicker than the opposing responses.

Discussion

In the present study we investigated speakers' responses to unexpected shifts in sensory feedback. Using an altered-auditory-feedback paradigm, we investigated whether responses were dependent on the state of the speech production system at shift onset. Overall, participants compensated for the pitch-shifted feedback by opposing the direction of the pitch shift, in line with previous research. This result is consistent with an internal forward model that compares the incoming auditory signal with the predicted auditory feedback (Wolpert & Ghahramani, 2000). Interestingly, all participants also followed the feedback shift on some trials. The proportions of following trials varied across participants (range = 10%–50%).

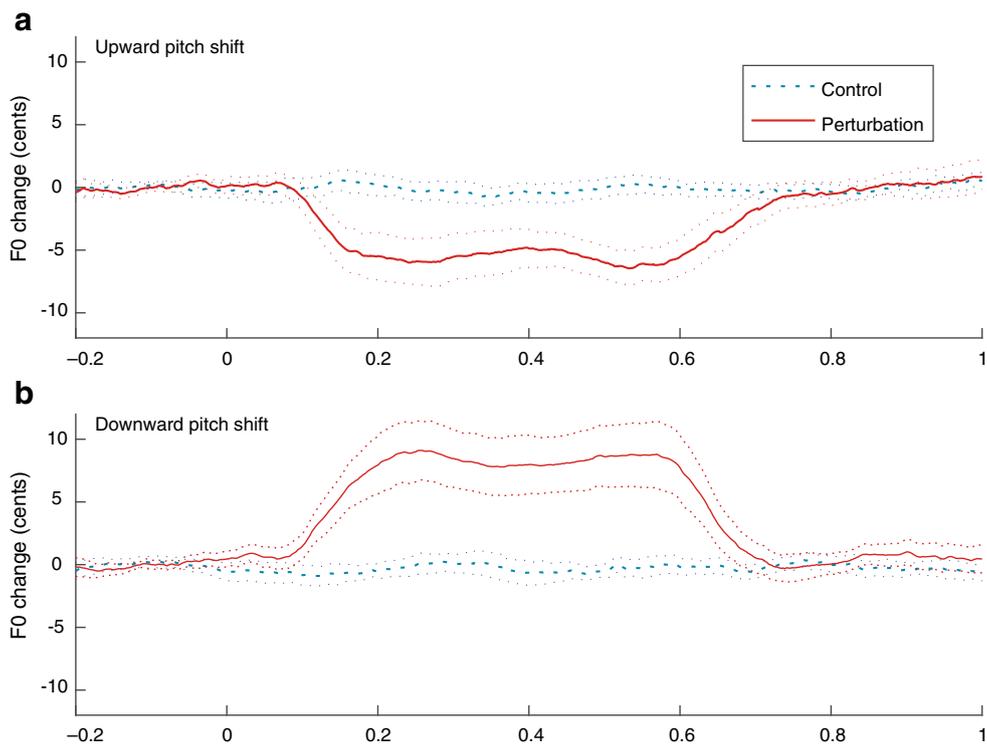


Fig. 5 Grand averages comparing changes in F0 in the perturbation and control trials in Experiment 2, time-locked to perturbation onset. In the perturbation trials, perturbation starts at 0 s and lasts until 0.5 s. (top) Perturbation trial contour, collapsed over opposing and following

responses for the upward perturbation block. (bottom) Grand average for the downward perturbation block. Thin dotted lines represent the 95% confidence intervals

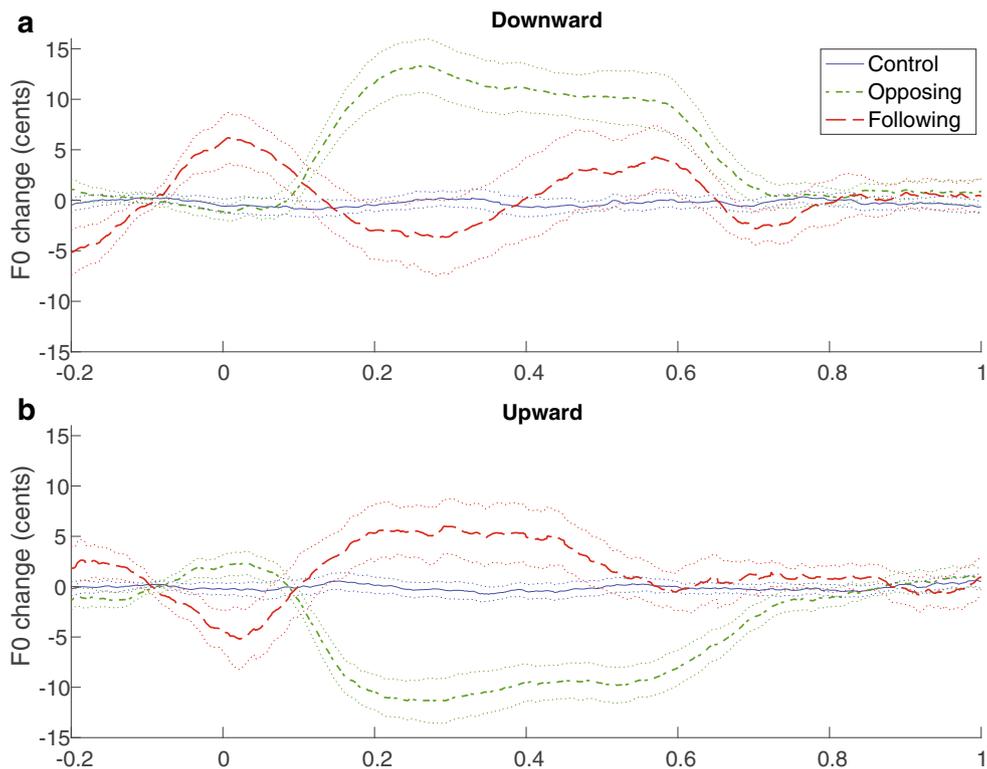


Fig. 6 Grand averages comparing changes in F0 in control, opposing, and following trials in Experiment 2, time-locked to perturbation onset. In the perturbation trials, perturbation starts at 0 s and lasts until 0.5 s. (a)

Data for the downward perturbation condition. (b) Data for the upward perturbation condition. Thin dotted lines represent the 95% confidence intervals

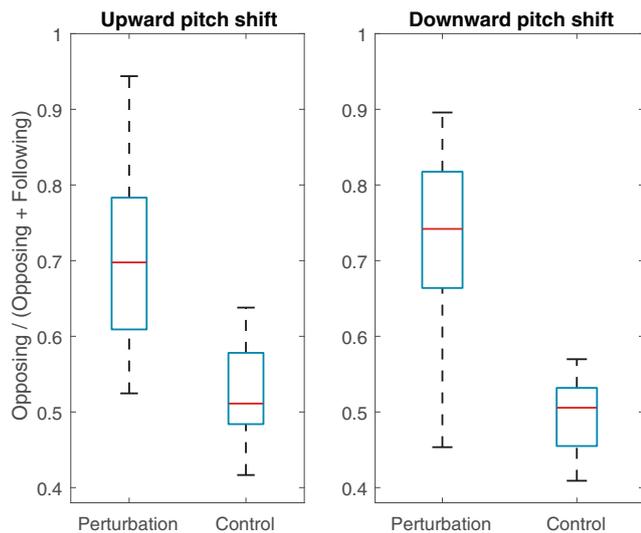


Fig. 7 Proportions of trials in Experiment 2 (upward left, downward right) classified as opposing in the perturbation and control conditions. The y-axis shows the number of trials classified as opposing divided by the total sum of trials that were classified as either opposing or following

Following responses are not in line with many models of sensory feedback processing for motor control (Houde & Nagarajan, 2011; Wolpert & Ghahramani, 2000). These models hypothesize that the goal of the motor system is to

minimize the discrepancy between the predicted sensory representation and the sensory feedback. However, when participants follow the direction of feedback perturbations in their vocal output, the discrepancy between prediction and sensory input increases.

These models could account for following responses in two ways. One possibility is that participants may have difficulty determining the direction of the pitch shift. This difficulty may have been due to the small magnitude of the pitch perturbations in this study; some listeners may have been able to detect the pitch change (unconsciously) without being able to correctly identify its direction (Neuhoff, Knight, & Wayand, 2002). Misidentifying the direction may then lead to following the pitch shift by mistake. However, because following responses are less common with smaller feedback perturbations (Burnett et al., 1998), and given that following is more common when the pitch change direction is predictable (Behroozmand et al., 2012), it is unlikely that misidentification of the perturbation direction was the sole cause of following responses.

The other possible account is that in some cases the auditory input is considered by the speaker to be externally driven rather than self-generated. An externally driven perturbation should be followed, whereas a self-generated shift should be

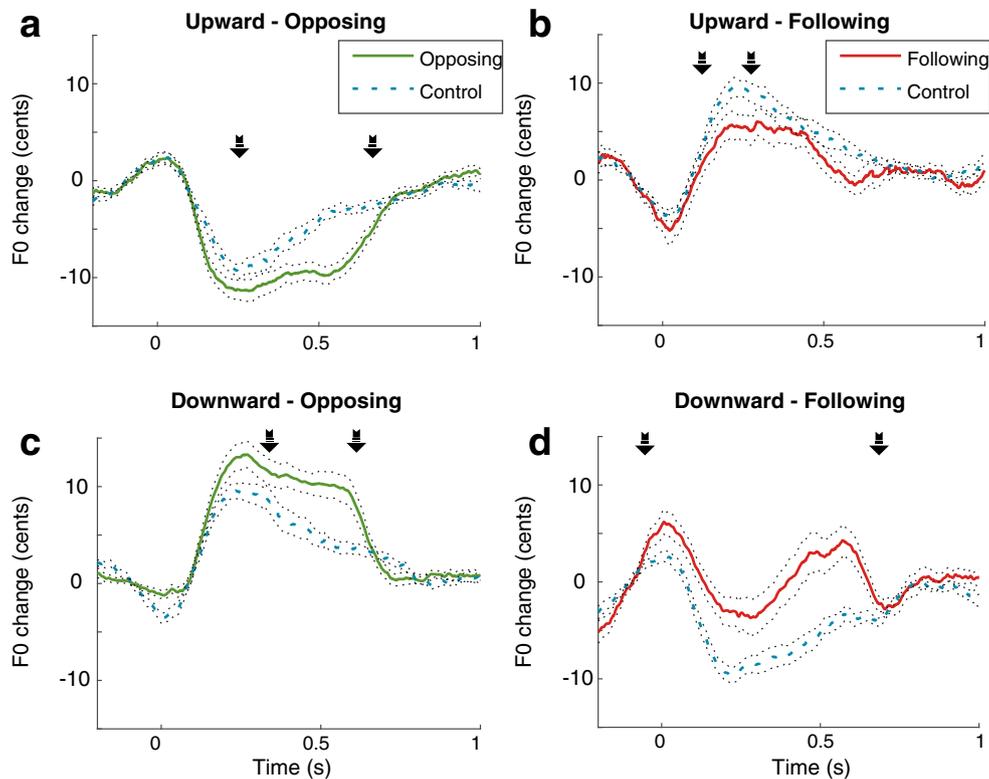


Fig. 8 Comparison of F0 changes between the perturbation and control conditions for similarly classified trials in Experiment 2. (a) F0 changes for opposing perturbation trials and control trials classified as opposing in the upward perturbation block. (b) F0 changes for following perturbation trials and control trials classified as following in the upward perturbation

block. (c, d) The same data for the downward perturbation block. Arrows indicate the onset and offset of the major component driving the statistical differences between the perturbation (either opposing or following) and control trials

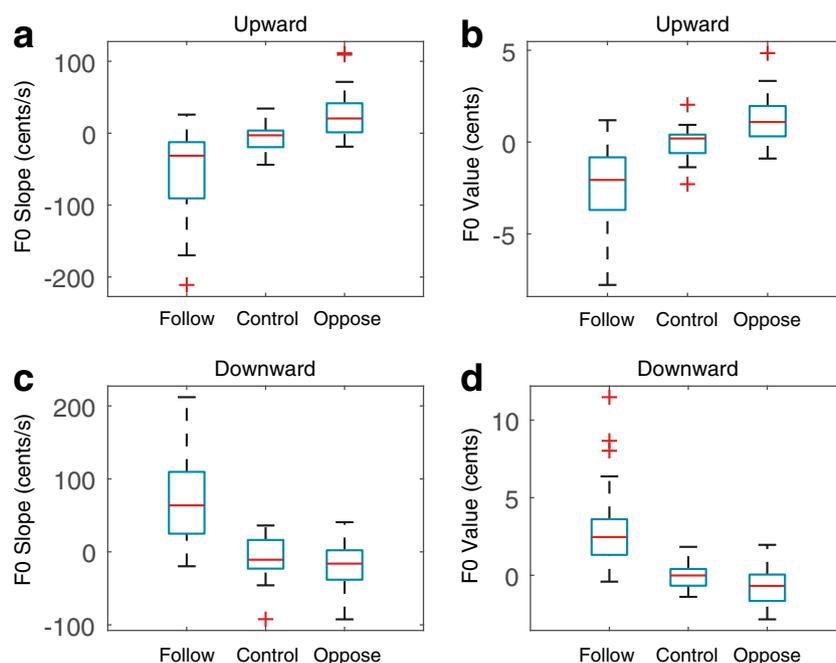


Fig. 9 F0 slope (a, c) and average F0 value (b, d) over a 100-ms time window before perturbation onset as a function of trial classification in Experiment 2. Panels a and b show the data for the upward block, and panels c and d for the downward block

opposed, in line with the model proposed by Hain et al. (2000). In the present study, however, it is unclear why sometimes the feedback shift would be considered to be self-generated (leading to opposing responses) and sometimes to originate from an external source (leading to following responses). In addition, the fact that no participant reported being aware of the perturbations suggests there were no clear changes in the perceived source of the auditory feedback.

We suggest that the following responses observed in both experiments may instead reflect system-internal constraints on the operation of the speech-production system. These constraints would need to be added to current models of sensorimotor control. Pitch is known to show fluctuations around a target pitch level (Akagi et al., 1998). These fluctuations could be driven by continuous feedback monitoring and (over)compensation. The present results indicate three interactions between ongoing fluctuations and the perturbation response.

First, the direction of the response was related to the F0 contour before perturbation onset. This suggests that the response is dependent on the current state of the system. Specifically, when the participants' pitch was decreasing or lower than average, the response tended to increase the pitch, and when pitch was increasing, the response tended to decrease it. This suggests that the system initially reacts to a pitch perturbation by doing the opposite of what it was doing, irrespective of the perturbation direction. This mechanism would be useful in a natural context. If a pitch mismatch is detected in the feedback signal, it is likely that ongoing compensatory articulations are going in the wrong direction. Thus, simply changing pitch in the opposite direction would be a good strategy.

The second interaction involved the comparison of the perturbation trials to the similarly classified control trials. Both the so-called “following” trials and the opposing trials showed opposing trends, suggesting that for the “following” trials, an opposing response may have been added to an ongoing pitch fluctuation.

Third, Experiment 1 suggested that the response peak for opposing trials was larger and occurred later than the peak on following trials. A following response would be detected as increasing the prediction error even more, leading to a quicker readjustment, and thus to an overall smaller response with an earlier peak.

Following responses so far have not been documented in nonspeech motor control (e.g., reaching movements). This may be because investigators have mainly applied perturbations from the start of the movement and because averaging across trials could wash out the following responses. The present results indicate the importance of investigating unexpected perturbations in ongoing movements at the single-trial level. Future work should examine whether these findings generalize to action domains beyond speech motor control.

Together, the present findings show evidence of a dynamic interplay between the state of the motor system and incoming sensory feedback, in line with a dynamic systems approach to cognitive processing (van Gelder, 1998). More generally, this study indicates that looking beyond the average response can lead to a more complete view of the nature of feedback processing in motor control. It also leads to the prediction that the direction of sensory feedback responses in domains outside speech production will also be conditional on the state of the motor system at the time of the perturbation.

Author note F.E. was funded by the Gravitation program Language in Interaction from the Dutch Science Foundation (NWO). M.K.F. was funded by the Research Talent program of the NWO.

References

- Akagi, M., Iwaki, M., & Minakawa, T. (1998). Fundamental frequency in continuous vowel utterance and its perception. In R. H. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP 98)* (pp. 1519–1522). Sydney, Australia: Australian Speech Science and Technology Association.
- Akagi, M., & Kitakaze, H. (2000). Perception of synthesized singing voices with fine fluctuations in their fundamental frequency contours. In *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP 2000)* (pp. 458–461). Beijing, China: China Military Friendship Publish.
- Behroozmand, R., Korzyukov, O., Sattler, L., & Larson, C. R. (2012). Opposing and following vocal responses to pitch-shifted auditory feedback: evidence for different mechanisms of voice pitch control. *Journal of the Acoustical Society of America*, *132*, 2468–2477. doi:<https://doi.org/10.1121/1.4746984>
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Retrieved from www.praat.org
- Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *Journal of the Acoustical Society of America*, *103*, 3153–3161. doi:<https://doi.org/10.1121/1.423073>
- Cai, S., Boucek, M., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2008). A system for online dynamic perturbation of formant frequencies and results from perturbation of the Mandarin triphthong /iau/. In R. Sock, S. Fuchs, & Y. Laprie (Eds.), *Proceedings of the 8th International Seminar on Speech Production* (pp. 65–68). Strasbourg, France: INRIA.
- Ford, J. M., Gray, M., Faustman, W. O., Heink, T. H., & Mathalon, D. H. (2005). Reduced gamma-band coherence to distorted feedback during speech when what you say is not what you hear. *International Journal of Psychophysiology*, *57*, 143–150. doi:<https://doi.org/10.1016/j.ijpsycho.2005.03.002>
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, *96*, 280–301. doi:<https://doi.org/10.1016/j.bandl.2005.06.001>
- Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research*, *130*, 133–141. doi:<https://doi.org/10.1007/s002219900237>
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145. doi:10.1038/Nrn2158
- Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*, 1213–1216.
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, *5*, 82. doi:<https://doi.org/10.3389/fnhum.2011.00082>
- Larson, C. R., Sun, J., & Hain, T. C. (2007). Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F0 and amplitude control. *Journal of the Acoustical Society of America*, *121*, 2862. doi:<https://doi.org/10.1121/1.2715657>
- Liu, H., & Larson, C. R. (2007). Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *Journal of the Acoustical Society of America*, *122*, 3671–3677. doi:<https://doi.org/10.1121/1.2800254>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190. doi:<https://doi.org/10.1016/j.neumeth.2007.03.024>
- Neuhoff, J. G., Knight, R., & Wayand, J. (2002). Pitch change, sonification, and musical expertise: Which way is up? In *Proceedings of the 2002 International Conference on Auditory Display* (pp. ICAD02-1–ICAD02-6). Kyoto, Japan.
- Niziolek, C., Nagarajan, S., & Houde, J. (2013). What does motor efference copy represent? Evidence from speech production. *Journal of Neuroscience*, *33*, 16110–16116.
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York, NY: McGraw-Hill.
- van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*, 615–628, disc. 629–665. doi:<https://doi.org/10.1017/S0140525X98001733>
- Wolpert, D., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*(Suppl), 1212–1217.