Consistency influences altered auditory feedback processing

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD-18-248.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	14-Feb-2019
Complete List of Authors:	Franken, Matthias; Ghent University, Experimental Psychology; Radboud University, Donders Institute for Brain, Cognition, and Behaviour; Max-Planck-Institut fur Psycholinguistik Acheson, Daniel; Max-Planck-Institut fur Psycholinguistik; Radboud University, Donders Institute for Brain, Cognition, and Behaviour McQueen, James; Radboud University, Hagoort, Peter; Max Planck Institute for Psycholinguistics; Radboud University, Donders Institute for Brain, Cognition, and Behaviour Eisner, Frank; Radboud University, Donders Institute for Brain, Cognition, and Behaviour
Keywords:	speech production, formants, speech motor control, sensorimototr integration

SCHOLARONE™
Manuscripts

Consistency influences altered auditory feedback processing

Matthias K. Franken[1,2,3], Daniel J. Acheson[1,2], James M. McQueen[1], Peter Hagoort[1,2] & Frank Eisner[1]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

[2]Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

[3]Department of Experimental Psychology, Ghent University, Belgium

Corresponding author: Matthias K. Franken, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium, +3292649427, matthias.franken@ugent.be

Abstract

Previous research on the effect of perturbed auditory feedback in speech production has focused on two types of responses. In the short term, speakers generate compensatory motor commands in response to unexpected perturbations. In the longer term, speakers adapt feedforward motor programs in response to feedback perturbations, in order to avoid future errors. The current study investigated the relation between these two types of responses to altered auditory feedback. Specifically, it was hypothesized that consistency in previous feedback perturbations would influence whether speakers adapt their feedforward motor programs. In an altered auditory feedback paradigm, formant perturbations were applied either across all trials (the consistent condition) or only to some trials while the others remained unperturbed (the inconsistent condition). The results showed that speakers' responses were affected by feedback consistency, with stronger speech changes in the consistent condition compared to the inconsistent condition. Current models of speech-motor control can explain this consistency effect. However, the data also suggest compensation and adaptation are distinct processes, which is not in line with all current models.

Introduction

Speaking is a complex

motor skill, and requires speakers to continuously monitor their own output to ensure accurate

performance (Elman, 1981). To investigate the role of auditory feedback in speech-motor

control, researchers have artificially altered what speakers hear by manipulating auditory

feedback in real time (Houde & Jordan, 1998; Purcell & Munhall, 2006b). These studies show

that the speech production system uses auditory input to control the production process. Acoustic

parameters altered in real-time have included pitch (Burnett, Freedland, Larson, & Hain, 1998),

formant frequencies (Houde & Jordan, 1998) and fricative noise (Shiller, Sato, Gracco, & Baum,

2009). Broadly, research has focused on auditory feedback being used on two different time-

scales. In the short term, unexpected auditory feedback leads to immediate corrective responses

(Burnett, Senner, & Larson, 1997; Franken, Acheson, McQueen, Hagoort, & Eisner, 2018;

Purcell & Munhall, 2006b). This line of research has shown that speakers on average quickly

compensate for sudden auditory feedback perturbations. In the longer term, speakers show

evidence of adaptation to consistent feedback (Houde & Jordan, 1998; Jones & Munhall, 2000).

Over time, speakers adapt to the new sensorimotor environment by changing their feedforward

speech motor commands accordingly (Purcell & Munhall, 2006a). This adaptation is typically

seen in the persistence of changes in speech output even after normal feedback has been restored.

Both the short-term compensation and the longer-term adaptation are almost always partial

responses and therefore do not fully undo the effects of the introduced feedback manipulation.

Theoretical frameworks have been developed to account for feedback-based speech

adjustments. For example, in the DIVA model (Guenther, 2006), a distinction is made between

the feedforward and feedback control systems. The feedback control system compares the

expected sensory consequences to the observed sensory input (i.e., the feedback). A mismatch leads to corrective behavioral adjustments. Over time, these adjustments are incorporated in the feedforward system in order to avoid future errors. Thus, according to DIVA, adaptive learning is dependent on short-term error-driven adjustments. A different framework is the state feedback control (SFC) model (Houde, Kort, Niziolek, Chang, & Nagarajan, 2013; Houde & Nagarajan, 2011), where auditory feedback is used to control and update an internal estimate of the dynamic state of the speech production system. As in DIVA, unexpected auditory feedback perturbations in the SFC model lead to compensatory behavioral adjustments and/or longer-term changes to the internal forward model. In contrast to DIVA, however, SFC assumes a single controller that operates on an internal state estimate. This estimate is based in part on auditory feedback and in part on the internal forward model. Adjustments of the forward model reflect adaptive learning, and thus are not directly dependent on error-driven adjustments, as they are in DIVA.

An open question is how the short- and long-term adjustments relate to each other. In the remainder of this article, we will refer to the immediate response to altered auditory feedback as compensation, and to longer-term changes in feedforward commands as adaptation. In Bayesian approaches to sensorimotor learning and adaptation (Franklin & Wolpert, 2011), the usefulness of adaptation varies across contexts. If a mismatch between expected and observed feedback is consistent time after time, it makes sense to adapt, in order to avoid future errors. If, however, the unexpected feedback is an isolated event, adaptation would lead to an additional error on the next attempt to speak. In the current study, we directly test whether feedback consistency affects adaptation. If auditory feedback is consistent over time, speakers should adapt, whereas there should be no adaptation when the feedback is inconsistent.

In addition, while many studies have investigated compensation or adaptation, only few have looked at them together and hence been able to examine their interaction. Although a number of studies quantify both compensation and adaptation (Houde & Jordan, 2002; Keough, Hawco, & Jones, 2013; Purcell & Munhall, 2006a), "compensation" is often measured as the change in speech during altered feedback, which may reflect both compensation and adaptation effects. Recently, Parrell et al. (2017), did disentangle compensation and adaptation using separate experimental paradigms. Furthermore, both within and outside the speech domain, there is disagreement about whether compensation and adaptation are distinct processes (Albert & Shadmehr, 2016; Shadmehr & Krakauer, 2008). In the speech domain, the leading models hold opposite views on this issue (the processes are distinct in SFC but not in DIVA).

In the current study, we tried to distinguish between compensation and adaptation using an altered auditory feedback paradigm with online adjustments to the first formant of a vowel that participants produced. We thus asked whether compensation and adaptation are distinct processes. We compared an inconsistent condition, where auditory feedback was perturbed on some trials but was left unperturbed on other trials, with a consistent condition, where the feedback was perturbed on all trials. We expected that speakers would show compensation by changing their speech output in the perturbed trials in both conditions (vs. the unperturbed baseline trials). In addition, we expected speakers to show adaptation of their feedforward motor commands in the consistent condition, but not (or less so) in the inconsistent condition. We expected more change in the perturbed trials in the consistent than in the inconsistent condition, as the former should show both compensation and adaptation.

Materials & Methods

Participants

Twenty-six healthy volunteers (age: M = 22, SD = 2.7; 17 females) participated after providing written informed consent in accordance with the Ethics Committee of the social sciences faculty of Radboud University. All participants had normal hearing, were native speakers of Dutch and had no history of speech and/or language pathology. Two participants were excluded because their speech production was too quiet to trigger feedback perturbation. We decided to obtain useable data from 24 participants based on informal comparisons with prior studies reporting effects of perturbed auditory feedback (e.g., 20, 10 and 29 participants in the experiments reported in Burnett, Freedland, Larson & Hain (1998), 28 in Purcell & Munhall (2006b), and 18 in Jones & Munhall (2000)).

Paradigm

The experiment consisted of four blocks of 80 trials each (see Figure 1). The blocks differed in consistency and in the direction of the perturbation. On every trial, participants were instructed to vocalize /e/ as soon as the letters <ee> (the appropriate orthography in Dutch) appeared on a computer screen and to keep doing so until the letters disappeared (3 seconds later). They were asked to vocalize at a comfortable loudness. The trials were relatively long (compared to previous formant-adaptation studies) in order to be able to assess within-trial compensation. During vocalization, participants' speech was recorded and played back through headphones. The loudness gain was set so that it was rather loud but comfortable, and the gain was kept the same across participants. In some trials (perturbation trials), the auditory feedback was manipulated by shifting the first formant (F1) by 6.7% up (in two blocks) or down (in the other two blocks). These formant shifts were smaller than in previous F1 perturbation studies, to make

sure that participants were not consciously aware of the perturbation. There were two consistency conditions. In the consistent condition, each block started with 20 start trials (no perturbation), followed by 40 perturbed trials (shifting F1 in the same direction in all trials) and finally 20 end trials (no perturbation). In the inconsistent condition, the blocks also started with 20 start trials and ended with 20 end trials, but the 40 trials in the middle were randomly assigned to be either non-perturbed or perturbed. F1 was shifted in the same direction in all perturbed trials in each block in the inconsistent condition and there were 20 perturbed trials in total. Trial 21 (i.e., the first trial after the 20 start trials) in all four blocks was always perturbed. Each participant was presented with all four blocks (consistent and inconsistent crossed with direction of F1 shift; see Figure 1). The order of the blocks was counterbalanced while making sure consistency alternated between blocks.

All voice recordings were made on one channel using a Sennheiser ME64 cardioid microphone, which was set up in a sound-attenuated booth and connected to a dedicated soundcard Motu MicroBook II outside the booth, connected in turn to a Windows laptop. Auditory feedback was delivered through the same soundcard which was also connected to Sennheiser HD 2801-13 headphones. Stimulus presentation and sound recording times were controlled by Audapter (Cai, Boucek, Ghosh, Guenther, & Perkell, 2008; Tourville, Cai, & Guenther, 2013) and MathWorks Matlab (R2013b). None of the participants noticed any delay in the auditory feedback due to online processing. No additional filters (besides the ones implemented in Audapter) were used for online audio processing.

Analysis

For every trial, the participants' speech recordings were marked for speech onset and offset by visual inspection, and F1 was estimated in Mels (Stevens, Volkmann, & Newman, 1937). Two

time windows of interest were identified for further analyses. Trial F1 contours were extracted from 50 to 1500ms after speech onset. Trials where formant estimation failed within this window were rejected (across participants, on average 1.5 (sd = 4.6) trials were removed). Most analyses were conducted on either a late time window (1000-1500ms after speech onset), or an early time window (50-150ms). The late window may reflect the added effects of both cross-trial adaptation as well as the results of online compensation during the current trial. The early window will reflect only cross-trial adaptation, as previous studies suggested that compensation responses may start from as early as 150ms after speech onset (Cai et al., 2012; Purcell & Munhall, 2006b).

Further analyses were carried out with R (R Core Team, 2013). For every subject and every block, F1 values were normalized in the following way:

$$F1_{norm} = 100 \cdot (-d) \cdot \frac{F1 - F1_{start}}{F1_{start}}$$

$F1_{start}$ is the average F1 value of the start trials for that subject in that block, and d is the sign of the direction of perturbation in that particular block. This leads to $F1_{norm}$ values being expressed as the percentage change in the opposite direction to the perturbation direction, relative to the average F1 in the start phase. In other words, if participants compensate for the F1 perturbation, we expect $F1_{norm}$ to be positive, irrespective of the perturbation direction.

Statistical testing was done by means of linear mixed effects models as implemented by the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015). In a first step, an appropriate model was selected by means of Akaike's Information Criterion (AIC). A series of models that differed with respect to their random effects structure were compared. Subsequently, the most appropriate random effects structure was selected and models with varying fixed effect structures were compared. The reported p values were calculated using a Satterthwaite approximation of the degrees of freedom.

In order to examine within-trial changes in the F1 time course, a non-parametric permutation test was performed with a clustering method to correct for multiple comparisons (Maris & Oostenveld, 2007), as implemented in the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). This was done for the data across the entire trial (50ms-1500ms) to determine whether there was a difference between the perturbed and unperturbed trials within the middle section of the inconsistent blocks and when such a difference might arise. The advantage of the cluster-based permutation test is that it is a data-driven way to detect differences between the conditions across the trial's time course, without the need of specifying a priori a time window of interest. Samples for which the Perturbed - Unperturbed contrast exceeded an uncorrected α level of .05 were temporally clustered. Cluster-level statistics were calculated by summing the t-statistics. Next, a permutation distribution was calculated by randomly exchanging data between the two trial types, and calculating the maximal positive and negative cluster-level statistics for all 10,000 permutations. The observed cluster-level statistic was tested against the permutation distribution, thus correcting for multiple comparisons.

Results

Three sets of analyses were conducted. The first set focused on the late time window (1000-1500ms), which would reflect both online compensation and adaptation effects. The second set of analyses focused only on the beginning of the trial (50-150ms) to isolate longer-term adaptation effects (i.e., before within-trial compensation could occur), and the third set of analyses (based on data from the late window, 1000-1500ms) attempted to identify whether feedback consistency affected adaptation or compensation.

In the first set of analyses, the average normalized F1 values in the late time window were compared across consistency conditions. An overview is presented in Figure 2, where

normalized F1 is shown as a function of Trial Type in both Conditions (random vs. consistent). As the figure suggests, participants altered the F1 in their speech output as a function of Condition and Trial Type.

The data from the perturbed and end trials were entered in linear mixed effects models. Model selection was done in two steps by first varying the random effects structure, and subsequently varying the fixed effects structure on the best-fitting model. Table I shows the fixed effects output of the best-fitting model, which included main effects of Trial Type and Consistency, but not their interaction. Without a fixed intercept in the model, a significant main effect indicates that F1 differed from the average start F1 (given the normalization). Table I shows that F1 was significantly altered in perturbed trials in the consistent condition, but not in the inconsistent condition. In addition, F1 in the end trials was lower than in the perturbed trials ($p = .022$). Under altered auditory feedback, participants responded by shifting their F1 in the opposite direction compared to the perturbation, but only under consistent perturbation across trials. After auditory feedback returned to normal, F1 returned to baseline.

Due to the design of the experiment, the number of perturbation trials differed between conditions. To make sure that the difference between the conditions was not due solely to the different number of trials, the same linear mixed effects model was run on data including all perturbation trials in the inconsistent conditions and only the first 20 perturbation trials in the consistent condition. As in the initial model, participants adjusted F1 in the consistent condition (estimate = .0049, SE = .00091, $t(24.01) = 5.40$, $p < .001$) but not in the inconsistent condition (est. = .0013, SE = .0011, $t(24.10) = 1.15$, $p = .26$). Again, F1 lowered F1 back to baseline in the end trials (est. = -0.0026, SE = 0.0011, $t(24.07) = -2.27$, $p = 0.033$).

In order to compare the perturbed and unperturbed trials in the middle section of the inconsistent blocks, a similar analysis was carried out on the inconsistent condition only. The best-fitting model included a significant main effect of Trial Type (perturbed vs. unperturbed), suggesting that in this late time window, there was more F1 change in the perturbed trials compared to the unperturbed ($t(1383) = 2.42$, $p = .016$). This difference between perturbed and unperturbed trials in the inconsistent blocks is evidence of online compensation, given that the current feedback is the only difference between these trials. Importantly, because these analyses have focused on the average F1 at the end of each trial, the results could reflect within-trial compensation as well as adaptation effects. In order to have a clearer view on how these effects develop, Figure 3 shows the F1 contour (or, more precisely, the difference in F1 contour from the average contour in the start trials) as a function of condition and direction of perturbation.

Interestingly, the F1 contour for perturbed and unperturbed trials in the inconsistent condition are similar initially, but the two time-courses diverge towards the end of the trial. A cluster-based permutation test was carried out on the data from the entire trial (i.e., 50-1500ms) to determine whether the perturbed-trial data differed from the unperturbed-trial data. This analysis allows us to isolate the effect of compensation without the need to specify a time window of interest a priori. Effects of adaptation should be the same in perturbed and unperturbed trials because they are the result of previous trials (there was only a very small difference in the average amount of perturbed trials in the five trials preceding a perturbed trial (2.36) or an unperturbed trial (2.56)), while effects of compensation are dependent on differences in online auditory feedback between the perturbed and unperturbed trials. For the blocks with downward F1 perturbation, F1 in the unperturbed trials was lower than in the perturbed trials (one-sided test; $p = 0.035$, CI = [0.031 0.039]). It tended to be higher in the positive F1 perturbation blocks, but this was not significant

(one-sided test; p = 0.27, CI = [0.26 0.28]). The effect for the downward F1 shift was driven by a single large cluster starting from 734ms after speech onset. The time-course of the uncorrected t statistics in Figure 4 suggests that F1 change increased with time for both F1 shift directions, suggesting that this effect is indeed due to within-trial compensation. Note that this difference between perturbed and unperturbed trials could be due to (1) compensation in the perturbed trials, or (2) compensation in the unperturbed trials (if there is adaptation, normal auditory feedback in unperturbed would constitute a feedback perturbation in the opposite direction, possibly leading to compensation in the opposite direction). We suggest these two options are not mutually exclusive, and that both reflect speakers' sensitivity to online auditory feedback.

In the second set of analyses, we attempted to isolate the longer-term adaptation effect from any within-trial compensation. This was done by focusing on the initial portion of the trials (50-150ms). Figure 5 shows the average over the early time windows for both conditions and both perturbation directions. It can be seen especially in the consistent condition that the perturbed trials show an adaptive change compared to the start trials (increased F1 when the perturbation was a downward F1 shift, and decreased F1 when perturbation was an increased F1 shift). A linear mixed effects model that included fixed main effects of Perturbation Direction and Consistency as well as their interaction (with random slopes by participant for both main effects and their interaction) showed a significant main effect of Perturbation Direction (est. = -.0036, SE = .0017, t(23.02) = -2.18, p = .040). This suggests that the F1 change at trial beginnings was dependent of the feedback perturbation direction, and thus evidence of adaptation. Interestingly, this is a main effect, suggesting there was adaptation in both consistent and inconsistent trials. In addition, the interaction between Perturbation Direction and Consistency was marginally significant (est. = -.0038, SE = .0022, t(23.06) = -1.75, p = .093). Specifically, this interaction

was driven by the fact that the effect of perturbation direction on F1 was larger for the consistent (est. = .0074, $\chi^2(1)$ = 19.43, p < .001) compared to the inconsistent condition (est. = .0036, $\chi^2(1)$ = 4.72, p = .030). Overall, despite the lack of clear aftereffects, there was adaptation in both conditions, and most adaptation occurred in the consistent condition.

Finally, in the third set of analyses, we examined whether the consistency-related F1 adjustment for the perturbation trials was due to compensation or adaptation. A linear regression model was run where the consistency-related adjustment (i.e., the overall effect of consistency between 1000ms and 1500ms after speech onset: the difference between consistent perturbed trials and inconsistent perturbed trials) was regressed against the amount of compensation and the amount of adaptation for each participant. Within-trial compensation was quantified by taking the difference in F1 adjustments between inconsistent perturbation and non-perturbation trials for that participant (between 1000ms and 1500ms after speech onset). While this may reflect compensation in perturbed and unperturbed trials, both are online compensation effects and thus reflect speakers' direct response to online feedback perturbations. Longer-term adaptation was quantified by taking the average F1 adjustment for that participants' last 10 trials in the mid phases of the consistent condition (between 50ms and 150ms after speech onset; i.e., before compensation could occur). Table II shows the results of the linear regression model. A marginally significant effect of Adaptation suggests that part of the consistency difference can be explained by adaptation. In other words, the consistent perturbation leads to increased adaptation. There was no significant relation with compensation. In addition, there was no significant association between adaptation and compensation ($r(46)$ = 0.093, p = 0.53). This result, though null, suggests that compensation and adaptation are two distinct processes, with consistency-related response differences mainly due to adaptation.

Discussion

The current study used an altered auditory feedback paradigm to investigate how feedback-related speech responses are affected by the consistency of feedback perturbations. The results indicate that consistency affects how speakers respond to altered auditory feedback, suggesting that more consistent feedback-based prediction errors lead to stronger behavioral adjustments.

Specifically, there was a difference between inconsistent and consistent conditions in the perturbed trials. The perturbed trials in both conditions were exactly the same, so any response difference must be due to the consistency difference. This suggests that speakers' motor adjustments are affected by the history of previous trials' perturbations, as hypothesized. The stronger build-up of adaptation in the consistent condition could be because every intervening unperturbed trial in the inconsistent condition leads to adaptation in the opposite direction (because after adaptation, no perturbation is essentially a perturbation in the opposite direction). This latter view is in line with what both DIVA and SFC models would predict. Alternatively, the results may be explained by an adjustment of the gain of feedback-driven adaptation mechanisms, leading to stronger adaptation in the consistent condition. Inconsistent auditory feedback could be considered less reliable, leading to a reduction in gain. In the consistent condition, in contrast, increased feedback reliability leads to a higher gain in feedback-related processing. Gain modulation of feedback processing is consistent with findings from the broader (non-speech) motor control literature (Gonzalez Castro, Hadjiosif, Hemphill, & Smith, 2014). In an arm-reaching task with force field perturbations, the rate of trial-to-trial adaptations was associated with consistency of the environment, showing quicker/stronger adaptation with more consistent perturbations. While Castro et al. disentangled environmental consistency and environmental variability, showing that the former affects adaptation rate while the latter affects

same-trial feedback responses, the current study is not able to disentangle these two factors. The current results therefore do not allow us to distinguish between the possibility of increased adaptation in the consistent condition or differential build-up of adaptation across trials in the consistent as opposed to the inconsistent condition.

In the current study, there was no aftereffect in the overall analysis, although Figure 2 suggests a trend in the expected direction. Previously, the presence of an aftereffect has commonly been taken as evidence in favor of adaptation of feedforward commands. The lack of evidence for aftereffects in the current study may be due to the rather small magnitude of our perturbation, and/or to the long trial length. Longer trials may lead to quicker dissipation of the adaptation effect in the end trials. An alternative explanation is the number of trials: the relatively low number of perturbation trials may have been too low to induce after-effects. The presence or absence of after effects, however, has no bearing on the present demonstration of a consistency effect. In particular, the F1 adjustment at the beginning of trials shows that participants adapted their feedforward speech commands. This change occurred early in the trial, before compensation responses take effect (Purcell & Munhall, 2006b), and must thus reflect longer-term adaptation. The association between participants' adaptation and the consistency-related difference between conditions on the perturbation trials also suggests that consistency leads to stronger adaptation. Although this is expected based on existing theoretical frameworks and previous work in motor control, this has not been demonstrated for speech production to date.

With respect to compensation, an analysis of the F1 time course revealed a difference between perturbed and non-perturbed trials in the inconsistent condition, suggesting that speakers showed within-trial feedback responses, at least in the blocks with downward F1 shifts. The analysis of

the F1 time courses suggested in addition (1) that compensation increases over time, as expected, and (2) that speakers are more sensitive to downward F1 shifts compared to upward F1 shifts. We speculate that this effect of perturbation direction may be due to asymmetry in Dutch vowel space: starting from /e/, there are more close-by vowel phonemes with decreasing F1 compared to increasing F1. As Niziolek & Guenther (2013) have shown, when auditory feedback is manipulated by shifting a vowel towards a close-by phoneme boundary, speakers tend to respond more strongly in order to avoid misinterpretation by the listener.

How can current theoretical frameworks account for the current results? The DIVA model (Guenther, Ghosh, & Tourville, 2006) suggests adaptation of the feedforward control subsystem occurs by integrating feedback-based corrections to avoid similar errors on subsequent trials. We speculate that DIVA might account for the present results by modulating the weights on the feedback control signal, but the model is not specific about how this modulation could be done. Even without this modulation, however, the DIVA model could account for the results by taking into account that trials with unperturbed feedback in the inconsistent condition could cancel out part of the adaptation built up so far, and thus lead to an overall difference in adaptation between the consistent and inconsistent conditions. Nevertheless, the DIVA model predicts that adaptation and compensation are associated: feedforward motor control is adapted by integrating a weighted version of the compensation response of the previous trial. This hypothesis was not borne out by the current results, as there was no correlation between compensation and adaptation. Although this is a null result, it suggests that compensation and adaptation are two separate processes, with mainly adaptation being affected by feedback consistency.

A somewhat different model is the SFC model (Houde & Nagarajan, 2011). This model assumes a Kalman gain function on the feedback prediction error. The Kalman gain depends on

the variability in the observed feedback, and thus can upregulate the influence of feedback when it is reliable (low variability), or downregulate when it is not. However, this gain controls the influence of feedback on the state estimate, and therefore influences the online responses. While Houde & Nagarajan (2011) state that adaptation is linked to updating the internal forward model (cfr. Shadmehr, Smith, & Krakauer, 2010), it is unclear whether a similar gain or even the same gain is applied to the feedback's influence on adaptation of the forward model. Note that the SFC model does argue for separate mechanisms that support compensation and adaptation. In addition, a recent study shows a dissociation between affected compensatory behavior and affected forward-model adaptation in clinical disorders (Parrell et al., 2017). Overall, it seems that the SFC model is better equipped to account for the present results, although it is unclear how the feedback gain on adaptation of the forward model is regulated. While the DIVA model does specify explicitly how the feedforward commands are adapted, it assumes a close association between compensation and adaptation, which is not in line with the present results.

In summary, the present report suggests that speakers' feedback-based speech adjustments depend on the consistency of past feedback errors. This can be implemented in the speech system by assuming that perturbed and unperturbed trials cancel each other out in the inconsistent condition, or by keeping track of the feedback error history. If the mismatch between expected and observed auditory input is consistent, feedforward control is adapted to the new environment. If the mismatch is sporadic, strong adaptation may in fact cause additional errors, and is therefore not warranted. In addition, the current data suggest that short-term compensation and forward-model adaptation are two distinct processes. This can be accounted for only by some models (such as the SFC model), but not by others (such as the DIVA model).

References

Albert, S. T., & Shadmehr, R. (2016). The Neural Feedback Response to Error As a Teaching Signal for the Motor Learning System. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *36*(17), 4832–45. https://doi.org/10.1523/JNEUROSCI.0159-16.2016

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*(6), 3153–3161. https://doi.org/10.1121/1.423073

Burnett, T. A., Senner, J. E., & Larson, C. R. (1997). Voice F0 responses to pitch-shifted auditory feedback: a preliminary study. *Journal of Voice*, *11*(2), 202–211. https://doi.org/10.1016/S0892-1997(97)80079-3

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE*, *7*. https://doi.org/10.1371/journal.pone.0041830

Cai, S., Boucek, M., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2008). A System for Online Dynamic Perturbation of Formant Trajectories and Results from Perturbations of the

Mandarin Triphthon /iau/. In *International Seminar on Speech Production 2008* (pp. 65–68). Strasbourg, France.

Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *The Journal of the Acoustical Society of America*, *70*(1), 45. https://doi.org/10.1121/1.386580

Franken, M. K., Acheson, D. J., McQueen, J. M., Hagoort, P., & Eisner, F. (2018). Opposing and following responses in sensorimotor speech control: Why responses go both ways. *Psychonomic Bulletin & Review*, *25*(4), 1458–1467. https://doi.org/10.3758/s13423-018-1494-x

Franklin, D. W., & Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, *72*(3), 425–442. https://doi.org/10.1016/j.neuron.2011.10.006

Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A., & Smith, M. A. (2014). Environmental consistency determines the rate of motor adaptation. *Current Biology : CB*, *24*(10), 1050–61. https://doi.org/10.1016/j.cub.2014.03.049

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, *39*(5), 350–365. https://doi.org/10.1016/j.jcomdis.2006.06.013

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*(3), 280–301. https://doi.org/10.1016/j.bandl.2005.06.001

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9469813

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech Language and Hearing Research*, *45*(2), 295–310.

https://doi.org/Doi 10.1044/1092-4388(2002/023)

Houde, J. F., Kort, N. S., Niziolek, C. A., Chang, E. F., & Nagarajan, S. S. (2013). Neural

evidence for state feedback control of speaking. In *Proceedings of Meetings on Acoustics*

(Vol. 19, pp. 060178–060178). Acoustical Society of America.

https://doi.org/10.1121/1.4799495

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers*

in Human Neuroscience, *5*(28). https://doi.org/10.3389/fnhum.2011.00082

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from

feedback perturbation. *The Journal of the Acoustical Society of America*, *108*(3), 1246.

https://doi.org/10.1121/1.1288414

Keough, D., Hawco, C., & Jones, J. a. (2013). Auditory-motor adaptation to frequency-altered

auditory feedback occurs when participants ignore feedback. *BMC Neuroscience*, *14*, 25.

https://doi.org/10.1186/1471-2202-14-25

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data.

Journal of Neuroscience Methods, *164*(1), 177–190. https://doi.org/DOI

10.1016/j.jneumeth.2007.03.024

Niziolek, C. A., & Guenther, F. H. (2013). Vowel Category Boundaries Enhance Cortical and

Behavioral Responses to Speech Feedback Alterations. *Journal of Neuroscience*, *33*(29),

12090–12098. https://doi.org/Doi 10.1523/Jneurosci.1008-13.2013

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). Fieldtrip: Open Source Software

for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data.

Computational Intelligence and Neuroscience, *2011*(2011).

https://doi.org/doi:10.1155/2011/156869

Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired Feedforward

Control and Enhanced Feedback Control of Speech in Patients with Cerebellar

Degeneration. *The Journal of Neuroscience*, *37*(38), 9249–9258.

https://doi.org/10.1523/JNEUROSCI.3363-16.2017

Purcell, D. W., & Munhall, K. G. (2006a). Adaptive control of vowel formant frequency:

Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of

America*, *120*(2), 966–977. https://doi.org/10.1121/1.2217714

Purcell, D. W., & Munhall, K. G. (2006b). Compensation following real-time manipulation of

formants in isolated vowels. *The Journal of the Acoustical Society of America*, *119*(4), 2288.

https://doi.org/10.1121/1.2173514

R Core Team. (2013). R: a language and environment for statistical computing. Vienna, Austria:

R Foundation for Statictical Computing. Retrieved from http://www.r-project.org

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control.

Experimental Brain Research, *185*(3), 359–381. https://doi.org/10.1007/s00221-008-1280-5

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and

adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108.

https://doi.org/10.1146/annurev-neuro-060909-153135

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech

sounds following speech motor learning. *J Acoust Soc Am*, *125*(2), 1103–1113.

https://doi.org/10.1121/1.3058638

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A Scale for the Measurement of the

Psychological Magnitude Pitch. *The Journal of the Acoustical Society of America*, *8*(3),

185–190. https://doi.org/10.1121/1.1915893

Tourville, J. A., Cai, S., & Guenther, F. (2013). Exploring auditory-motor interactions in normal

and disordered speech (pp. 060180–060180). https://doi.org/10.1121/1.4800684

Figure Captions

Fig. 1. Overview of the experimental blocks, in a 2-by-2 design (consistency by perturbation direction). Order of the blocks was counterbalanced across participants, so that two blocks of the same consistency never followed each other. The sequence of perturbed and unperturbed trials in the inconsistent condition is an example; trial-type order was randomized across participants (see main text for constraints on randomization).

Fig. 2. F1 adjustments as a function of trial type and consistency. The F1 adjustment values express change in F1 as a percentage of the average F1 value in the start trials (see methods), averaged across a late time window (1000 ms – 1500 ms after speech onset). Error bars reflect standard errors across participants (none given for the start trials, because these data are normalized relative to the start trial baseline). (start = unperturbed trials at the beginning of a block; perturb. = perturbation trials; no perturb. = unperturbed trials in the middle portion of an inconsistent block, end = unperturbed trials at the end of a block)

Fig. 3. F1 time courses from 50 ms to 1500 ms after speech onset, as a function of trial type (color), consistency (column), and perturbation direction (row). Semi-transparent shading indicates standard error across participants.

Fig. 4. T statistic (uncorrected) for the inconsistent perturbed vs. non-perturbed trials across time for each perturbation direction.

Fig. 5. Average normalized F1 values at the beginning of the trial (50-150ms after speech onset), as a function of condition (columns), perturbation direction (rows), trial type (color) and trial number (x axis). Note that here, perturbed and unperturbed trials in the inconsistent condition are

collapsed in one category, as a given trial number in this condition could reflect either type of

trial depending on the participant and perturbation direction.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Fig. 1. Overview of the experimental blocks, in a 2-by-2 design (consistency by perturbation direction). Order of the blocks was counterbalanced across participants, so that two blocks of the same consistency never followed each other. The sequence of perturbed and unperturbed trials in the inconsistent condition is an example; trial-type order was randomized across participants (see main text for constraints on randomization).

146x102mm (300 x 300 DPI)
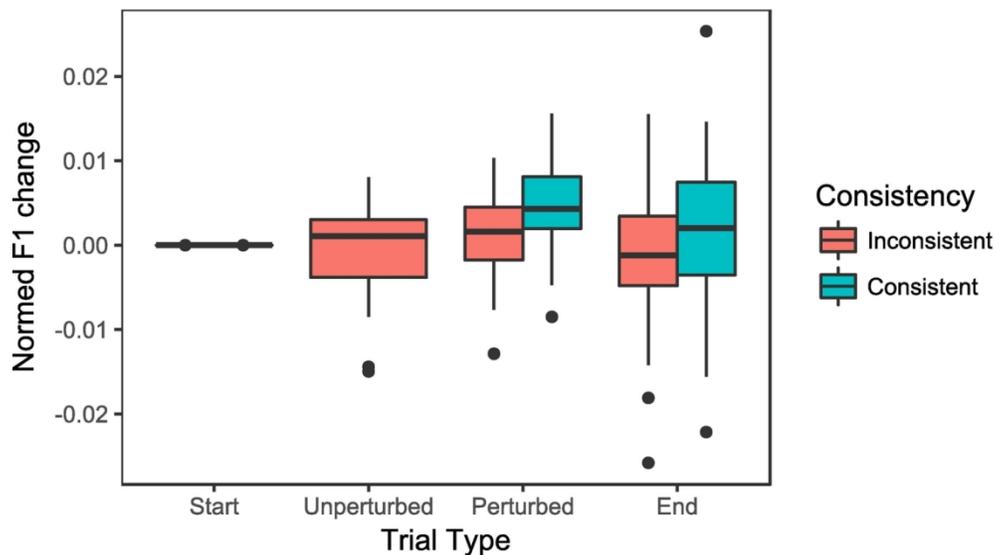
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Fig. 2. F1 adjustments as a function of trial type and consistency. The F1 adjustment values express change in F1 as a percentage of the average F1 value in the start trials (see methods), averaged across a late time window (1000 ms – 1500 ms after speech onset). Error bars reflect standard errors across participants (none given for the start trials, because these data are normalized relative to the start trial baseline). (start = unperturbed trials at the beginning of a block; perturb. = perturbation trials; no perturb. = unperturbed trials in the middle portion of an inconsistent block, end = unperturbed trials at the end of a block)

131x72mm (300 x 300 DPI)
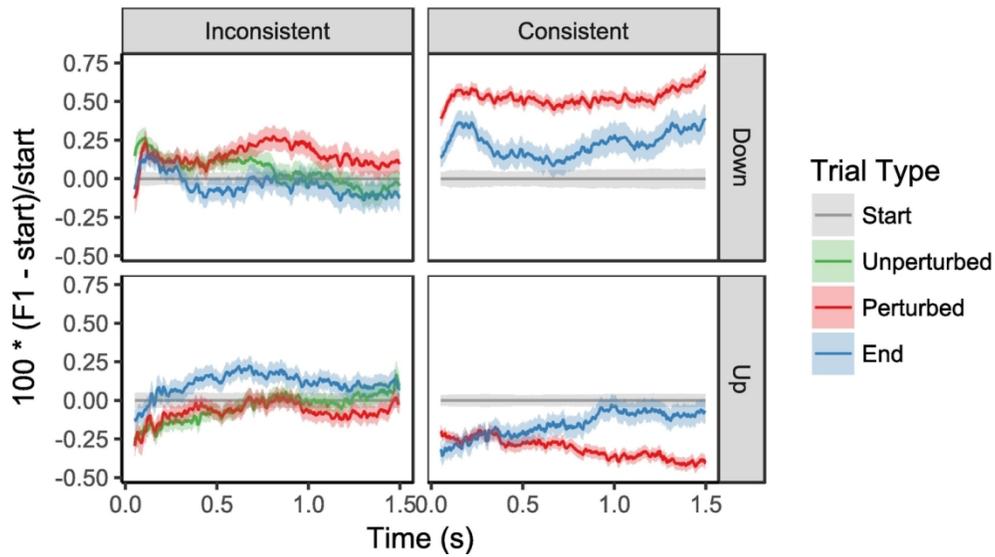
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Fig. 3. F1 time courses from 50 ms to 1500 ms after speech onset, as a function of trial type (color), consistency (column), and perturbation direction (row). Semi-transparent shading indicates standard error across participants.

131x72mm (300 x 300 DPI)
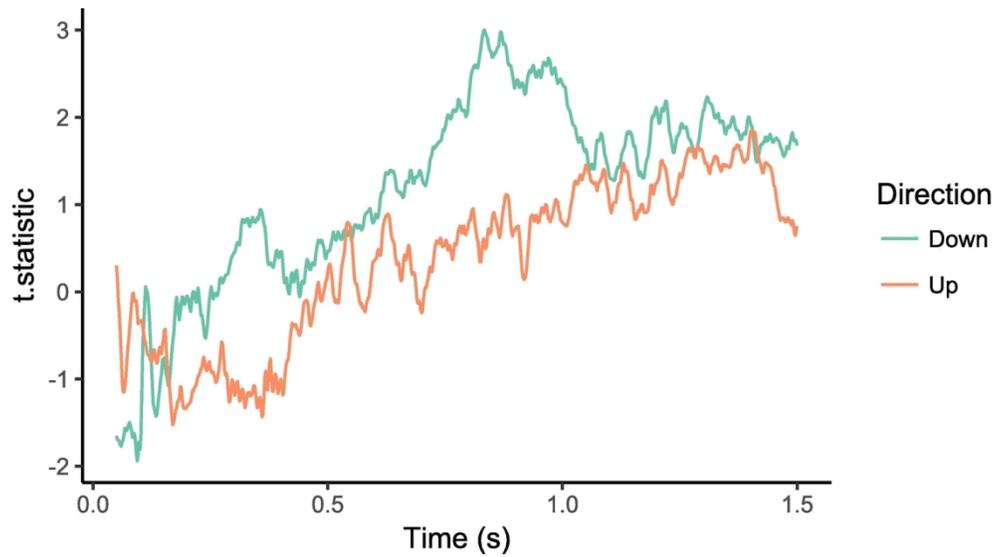
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Fig. 4. T statistic (uncorrected) for the inconsistent perturbed vs. non-perturbed trials across time for each perturbation direction.

131x72mm (300 x 300 DPI)
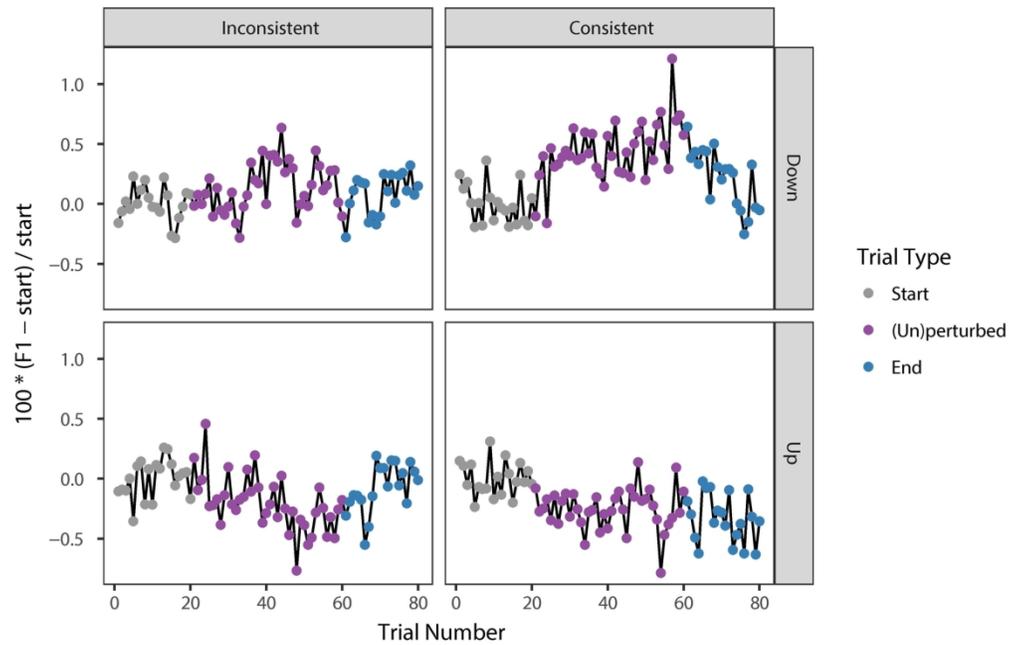
Fig. 5. Average normalized F1 values at the beginning of the trial (50-150ms after speech onset), as a function of condition (columns), perturbation direction (rows), trial type (color) and trial number (x axis). Note that here, perturbed and unperturbed trials in the inconsistent condition are collapsed in one category, as a given trial number in this condition could reflect either type of trial depending on the participant and perturbation direction.

164x105mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table I. Model: F1 ~ 0 + Condition + TrialType + (1 + Condition*TrialType | Participant)

Fixed effects	Estimate	SE	t	p (Satt.)
Condition.Inconsistent	.0012	.0011	1.09	.29
Condition.Consistent	.0048	.0011	4.48	**< .001***
TrialType.End	-.0026	.0011	-2.46	***.022***

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table II. Consistency-related speech adjustment as a function of Adaptation and Compensation.

	Estimate	SE	t	p
(Intercept)	.0024	.0018	1.33	.19
Adaptation	.29	.16	1.85	.07[(*)]
Compensation	-.060	.081	-.75	.46
Adaptation:Compensation	4.22	7.39	.57	.57