Don't blame yourself: Conscious source monitoring modulates feedback control during speech production

Matthias K. Franken[1*], Robert J. Hartsuiker[1], Petter Johansson[2], Lars Hall[2], & Andreas Lind[2]

[1]Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium

[2]Department of Philosophy, Lund University Cognitive Science, Lund University, Box 192, 221 00 Lund, Sweden

Corresponding author: Matthias K. Franken, matthias.franken@mcgill.ca

[*]Currently at Department of Psychology, McGill University, Montreal, Quebec, Canada.

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

Abstract

Sensory feedback plays an important role in speech motor control. One of the main sources of evidence for this are studies where online auditory feedback is perturbed during ongoing speech. In motor control, it is therefore crucial to distinguish between sensory feedback and externally generated sensory events. This is called source monitoring. Previous altered feedback studies have taken non-conscious source monitoring for granted, as automatic responses to altered sensory feedback imply that the feedback changes are processed as self-caused. However, the role of *conscious* source monitoring is unclear. The current study investigated whether conscious source monitoring modulates responses to unexpected pitch changes in auditory feedback. During a first block, some participants spontaneously attributed the pitch shifts to themselves (self-blamers) while others attributed them to an external source (other-blamers). Before block 2, all participants were informed that the pitch shifts were experimentally induced. The self-blamers then showed a reduction in response magnitude in block 2 compared with block 1, while the other-blamers did not. This suggests that conscious source monitoring modulates responses to altered auditory feedback, such that consciously ascribing feedback to oneself leads to larger compensation responses. These results can be accounted for within the dominant comparator framework, where conscious source monitoring could modulate the gain on sensory feedback. Alternatively, the results can be naturally explained from an inferential framework, where conscious knowledge may bias the priors in a Bayesian process to determine the most likely source of a sensory event.

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

Introduction

It is well established that sensory feedback plays a major role in speech motor control. Speakers monitor the results of their articulation during speech production in order to avoid and correct speech errors (Guenther, 2016; Houde & Jordan, 1998; Tremblay et al., 2003). This process is known as feedback monitoring. In order for feedback monitoring to work, the auditory signal needs to be processed as self-produced, and thus distinguished from externally generated auditory signals. This is known as source monitoring, and usually occurs without conscious awareness. Previous investigations of feedback monitoring have mostly taken (non-conscious) source monitoring for granted. However, it is important to distinguish this non-conscious process from a conscious determination about whether one was the origin of a specific utterance or not, here called conscious source monitoring. While feedback monitoring logically depends on non-conscious source monitoring, it is unclear what the contribution of conscious source monitoring to feedback monitoring in speech production is. The current study therefore investigates whether conscious source monitoring modulates feedback monitoring in speech production.

Past research has suggested that speakers usually compensate for unexpectedly altered auditory feedback during speech production, by shifting their speech in the opposite direction (a so-called opposing response). Although in most studies, participants are typically aware of the unexpected pitch perturbations (Burnett et al., 1998a; Hain et al., 2000; Larson et al., 2001a), studies showing that these responses even occur in the absence of conscious awareness of the feedback alterations indicate that these responses are automatic (Franken, Eisner, et al., 2018; Hafke, 2008). For example, Hafke (2008) provided participants with pitch-shifted auditory feedback in varying magnitudes. Participants consciously perceived the larger pitch shifts, but not the smaller ones. Nevertheless, even the smallest pitch shifts, which participants did not consciously notice, led to opposing responses in the participants'

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

speech production. So conscious perception of the pitch shift is not necessary to generate a compensatory response. Furthermore, in many studies participants are made aware beforehand that they will hear alterations in the auditory feedback, and they are instructed to try to ignore them (Burnett et al., 1998b; Keough et al., 2013; Larson et al., 2001b; Scheerer & Jones, 2012). These studies show that participants usually still respond to pitch-shifted feedback. Together, these results suggest that pitch compensations are automatic, as consciousness is not required for the responses to occur, and the responses cannot easily be overridden by conscious effort. On the other hand, there is evidence that later parts of the responses can be modulated by task instructions (Hain et al., 2000), suggesting they are at least partly under voluntary control. Hain et al. instructed participants to change their pitch in the opposite direction of the pitch shift, in the same direction, or to ignore the shift. Their results suggested that, while earlier parts of the response were often made incorrectly, suggestive of automatic responses, later parts of the response (from about 300ms after shift onset) were almost always in the instructed direction, suggestive of voluntary control. Other studies using similar instructions to their participants (Zarate et al., 2010; Zarate & Zatorre, 2008) suggested that responses to pitch-shifted feedback are harder to suppress when the pitch shift is smaller or when participants had less vocal training. With less vocal training and/or smaller shifts, participants were less likely to suppress compensation responses in the 'ignore' condition. The effect of vocal training is suggestive of increased vocal control in musically trained participants, and thus that compensation responses can be controlled. The effect of shift size suggests that smaller shifts are more likely to be within the range of magnitudes of accidental pitch fluctuations, and thus are more likely to be processed as self-caused and therefore lead to compensatory responses.

These previous studies investigating the role of conscious awareness in feedback monitoring have collapsed two different dimensions with respect to speakers' conscious awareness of the feedback manipulations: (1) conscious awareness of a sudden unexpected change in auditory input and (2) conscious source monitoring of the unexpected feedback change. To date, the latter dimension is largely ignored in most studies. For example, Hafke (2008) compared a condition with very small pitch shifts that participants were not consciously aware of (because they did not consciously perceive the pitch change) with a condition with larger shifts that participants *were* consciously aware of. This comparison ignores the possibility that participants could be consciously aware of a pitch change, but still attribute it to their own production rather than to an experimental manipulation. If they do attribute it to themselves this way, then it could be expected that they still compensate for the shift, while if they do not attribute it to themselves, they could be expected to not compensate as much. In other words, if participants notice the feedback change, it is still a pertinent question whether or not they consider the change to be self-generated or manipulated by external factors (i.e., whether they blame themselves or external factors for the unexpected pitch change).

Most studies of pitch-shifted auditory feedback include a pitch shift of 100 cents, which is large enough for participants to consciously perceive (Bauer & Larson, 2003; Burnett et al., 1998a; Burnett & Larson, 2002; Chen et al., 2007; Greenlee et al., 2013; Hawco et al., 2009; Larson et al., 2001a; Liu et al., 2010; Scheerer et al., 2013). It is well established that very large pitch shifts (larger than 200-250 cents[2]) in auditory feedback lead to smaller responses, compared to smaller pitch shifts (Hawco et al., 2009; Korzyukov et al., 2012; Scheerer et al., 2013; Subramaniam et al., 2018). This is commonly interpreted as an indication that large pitch shifts are not considered to be self-generated, and thus do not lead to strong compensatory responses. In other words, this suggests that source monitoring modulates responses to pitch-shifted feedback. In these studies, however, source monitoring is not necessarily

[2] The unit 'cents' refers to a psychometric scale to express a frequency interval, where doubling the frequency in Hertz corresponds to a 1200 cents increase. An interval of 100 cents reflects the interval between, for example, C and C# in western music. See below (methods) for more information.

conscious, and is not isolated from the effects of pitch shift magnitude. The current study will try to disentangle conscious awareness of the pitch change and conscious source monitoring, and specifically investigate whether conscious source monitoring of the feedback change modulates responses to pitch-shifted auditory feedback during speech production. While, as we have seen, previous studies have shown that task instructions can lead to top-down modulation of responses to pitch-shifted feedback, these task instructions can be assumed to have led to conscious attribution of the pitch shifts to an external source. The current study instead compares the responses to pitch shifts when the feedback is demonstrably perceived as either self- or other-produced, thus allowing for the investigation of top-down modulation by conscious source monitoring.

The dominant theoretical framework with respect to source monitoring in speech production, the comparator model, suggests that the distinction between self-caused and externally generated sensory stimuli is mainly driven by a comparison between internal predictions of sensory feedback and perceived sensory input (Blakemore et al., 2002; Frith, 2005; Wolpert & Ghahramani, 2000). A match would lead to the recognition of sensory input as self-caused, while a (large enough) mismatch indicates that the sensory input is externally generated. Although left implicit, this framework seems to treat source monitoring as non-conscious. Non-conscious source monitoring is closely related to the concept of sense of agency, which is defined as the feeling that one is in control over one’s actions and their sensory consequences (Haggard, 2017). However, while sense of agency has been investigated and defined in detail, the role of non-conscious source monitoring has not. Whereas non-conscious source monitoring is usually only implied, it is clear that, according to the comparator model, it is a crucial prerequisite in the generation of a sense of agency, because this requires that the sensory stimuli in question are treated as self-generated. However, this framework does not provide a clear role for *conscious* source monitoring. The present study aims to extend previous research by investigating the possibility of top-down modulation of speech motor control by conscious source monitoring.

As discussed above, very large pitch shifts (i.e., larger than 200-250 cents) lead to a smaller response or no response (Hawco et al., 2009; Korzyukov et al., 2012; Subramaniam et al., 2018). Another example of the importance of shift magnitude is related to the response direction. Previous studies have not only shown opposing responses to altered feedback as discussed above, but also a minority of so-called following responses, where speakers follow the feedback shift instead of compensating for it (Behroozmand et al., 2012; Franken, Acheson, et al., 2018; Li et al., 2013). Some authors (Behroozmand et al., 2012; Burnett et al., 1998b) have shown that larger pitch shifts lead to more following responses or fewer opposing responses than smaller shifts. This has been argued to be associated with nonconscious source monitoring as well. It is usually assumed that smaller shifts are considered by the speaker to be self-generated, and that the speaker compensates for them by opposing the shift direction. As for following responses, it has been suggested that speakers are more likely to treat larger shifts as externally generated, which rather leads them to try to align their pitch with the feedback (much as matching the pitch of a tuning fork) (Hain et al., 2000; however, see Behroozmand et al., 2012 for an alternative view).

The current study used a preregistered procedure to investigate whether conscious source monitoring modulates the effect of unexpectedly pitch-shifted feedback on vocal control. Participants were exposed to pitch-shifted feedback (±25 cents or ±100 cents). After the first experimental block, it was determined through a semi-structured interview whether the participant considered the pitch shifts to be self-caused (in which case they were labelled a "self-blamer") or as an experimental manipulation (in which case they were labelled an "other-blamer"). After the interview, participants were informed of the feedback manipulations and a second block followed. This way, there was a between-participants source monitoring contrast (by comparing results in the first block between self-blamers and other-blamers) and a within-participants contrast for the self-blamers (by comparing the first and the second block for the self-blamers). If conscious source monitoring affects responses to pitch-shifted feedback in a way similar to how previous research has suggested that non-conscious source monitoring does, we expect larger compensation responses for the self-blamers in the first block, compared to the other-blamers' first block and the self-blamers' second block. Similarly, if response direction is also affected by conscious source monitoring, we expect a higher proportion of opposing responses when speakers consider the shifts to be self-generated.

Methods

All processed data and analyses are publicly available on the Open Science Framework via https://doi.org/10.17605/OSF.IO/Y2FNR .

Pre-registration

The study design, methods, analyses and hypotheses were preregistered at the Open Science Framework (https://doi.org/10.17605/OSF.IO/4KC9T). Any changes from or additions to the pre-registered report are explicitly labeled as such.

Participants

Sixty participants volunteered to take part in this study in exchange for course credit (9 male, 51 female, age M = 19.1, SD = 2.0). All participants signed an informed consent in accordance with the Declaration of Helsinki. All experimental procedures were approved by the ethical committee of the Faculty of Psychology and Educational Sciences at Ghent University. While the preregistered report estimated data acquisition to be finished by December 2019 at the latest, scheduling conflicts led data acquisition to continue until March 2020.

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

We used the software program G*Power (Faul et al., 2007) to conduct a power analysis to estimate the statistical power of an interaction of a within-subjects and a between-subjects factor in a mixed-design ANOVA. The effect size (partial $\eta^2 = 0.05$) and the correlation between repeated measures (0.42) were based on a pilot study. Our goal was to obtain 0.80 power to detect an effect of similar size at the standard 0.05 alpha error probability. This leads to a required sample size of at least 23 participants in each group. We decided to collect data until there were at least 30 participants in each group (self-blamers and other-blamers), resulting in (coincidentally) 60 participants. After the experiment, all participants passed a hearing screening (see below). Based on the audiogram, one participant was removed from further analyses due to a hearing threshold of >25 dB HL on 2 out of 5 frequencies for the left ear.

Procedure

The experiment consisted of two blocks of trials. Each block of trials consisted of 100 vocalizations. The appearance of the letters <EE> (pronounced as /e/ in Dutch) on the screen prompted participants to vocalize /e/, for as long as the letters were on the screen (4s). The screen also showed a progress bar indicating how far along in the current block they were. During vocalization, participants received auditory feedback through headphones. The pitch of the auditory feedback was shifted briefly up or down by either 25 or 100 cents, up to three times during each vocalization. Every pitch shift lasted for 200ms and there was 600-800ms (jittered) between successive pitch shifts. Participants were not told about the pitch shifts before the experiment.

After the first block, a semi-structured interview was conducted with the participants in order to categorize each participant as either a self-blamer or an other-blamer. After the interview, all participants were informed that the pitch shifts in the feedback were caused by experimental manipulation. Subsequently, the second block was carried out. After the second block, participants filled

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

out a questionnaire on locus of control (Brosschot et al., 1994), a questionnaire on their background with respect to language and music, and an audiogram was measured with a portable audiometer (Maico MA25).

Feedback manipulations

Speech was recorded using a DPA directional microphone (4088-B) positioned at about 2 cm from the participant's mouth. Auditory feedback was played back to the participant using a pair of home-built headphones, designed to provide extra passive sound isolation (Franken et al., 2019). The pitch shifts were administered using an Eventide Eclipse multi-effects processor (Eventide Inc., Little Ferry, NJ, USA), controlled via MIDI messages sent from a program written in Pure Data (Puckette, 1996), run on a laptop.

The applied pitch shifts were pseudo-randomized within a block, in such a way that in every two consecutive vocalizations, one of each perturbation type (-25, +25, -100 or +100 cents) was applied, as well as two null (0 cent) perturbations.

Semi-structured interview

The purpose of the semi-structured interview in between the two experimental blocks was to determine whether participants believed the pitch shifts in the auditory feedback during the first block were caused by themselves or not (see Appendix). The interview was designed to start with open questions in order to get participants to spontaneously mention the feedback manipulations if possible. Gradually, questions would be more specific. Theoretically, the interview would classify participants either as a "self-blamer", if they noticed pitch changes but attributed these to themselves, as an "other-blamer", if they noticed the pitch changes and attributed them to an experimental manipulation, or as “unaware of the pitch shifts”, if they did not notice any pitch changes. However, no participant was classified as being unaware of the pitch shifts.

Preregistered analyses

For every vocalization, the pitch of participants’ vocalizations was estimated using the autocorrelation method in Praat (Boersma & Weenink, 2017). For every pitch perturbation, the trial’s pitch contour was generated by extracting the participants’ pitch from -200ms to +700ms relative to perturbation onset. Pitch was converted to the cents scale with the following formula:

$$pitch_{cents} = 1200 \cdot \log_2\left(\frac{pitch_{Hz}}{pitch_{baseline}}\right)$$

Here, “$\text{pitch}_{\text{baseline}}$” is the mean pitch value in Hertz across the baseline window (-200 ms to 0 ms relative to perturbation onset). The sign of the pitch contours for upward perturbations (+25 cents and +100 cents) was flipped so feedback-opposing compensation responses have the same sign across perturbation directions. Missing parts in the pitch track that are shorter than 50 ms were interpolated linearly from neighboring time points.

Trials for which the pitch estimation failed were removed from the analysis. These pitch estimation fails were identified by visual inspection, looking for sharp discontinuities in the pitch contour, high variability in the pitch contour, and missing time points in the pitch track that last for over 50ms. This resulted in an average of 54.9 trials rejected out of 600 per person, with a median of 39 trials or 6.5%. All subjects had more than 20 remaining trials in each condition (perturbation by block). Note that Bauer et al. (Bauer & Larson, 2003) suggested a minimum of 15 trials is necessary to get a good estimation of the pitch response.

Based on the pitch contours, the response magnitude and the response type were estimated. With respect to the response magnitude, the maximal value of the pitch contour relative to the mean baseline pitch was averaged across trials (per block, perturbation, group, and participant) and taken as the compensation response magnitude. With respect to the response type, every single trial was classified as containing either an opposing or a following response, using a classification technique used in Franken et al. (2018). This method relies on determining the slope of the pitch contour shortly after perturbation onset to classify a response as opposing or following.

The main hypothesis was tested using a linear mixed effects model (carried out in R with the ‘lme4’ package (Bates et al., 2015)), with factors Source (Self, External), Block (Block 1, Block 2), and Perturbation Size (25 cents, 100 cents). The model included, in addition to main effects, all pairwise two-way and three-way interactions and the maximal random effects structure allowed by the data (with respect to model convergence). Degrees of freedom were estimated using Satterthwaite’s method. The same analysis was repeated with the response type (opposing or following) as a dependent variable. As this was a binary variable, a logistic mixed effects model was used here.

Exploratory analyses

In order to explore some of the factors that may determine whether participants were self-blamers or other-blamers, the results of the Locus of Control questionnaire, the music and language background questionnaire, and the audiograms were entered as independent variables in a stepwise logistic regression, with group (self-blamers vs. other-blamers) as the dependent variable. The Locus of Control questionnaire led to three variables, corresponding to the Self, Powerful Others, and Chance subscales (Brosschot et al., 1994). From the music and language questionnaire, the following variables were used: participants’ rating on whether they could easily imitate another accent or dialect, on whether they had a language talent, on whether they had a Dutch/Flemish accent when speaking a second language, whether they were talented in music, as well as a two-level factor indicating whether they had experience in playing a musical instrument. For the audiogram results, hearing thresholds were acquired for 5 different frequencies (500Hz, 1kHz, 2kHz, 4kHz, 8kHz) at each ear. The 10 resulting thresholds in dB HL were averaged across ears and frequencies. All numeric independent variables were transformed to a z-score for the stepwise regression. Stepwise regression was performed using the 'stepAIC' function in R, which added or removed variables from the model based on Akaike's Information Criterion (AIC).

Results

Preregistered analyses

Average Response Contour. Figure 1 shows the compensation responses per condition. Descriptively, it can be observed that participants in the self-blamers group tended to show a larger compensation response in Block 1 compared to Block 2. A similar trend is observable for the larger (±100 cents) pitch shifts in the other-blamers group, but not for the smaller (±25 cents) shifts. In addition, both groups show larger responses to larger shifts compared to smaller pitch shifts.

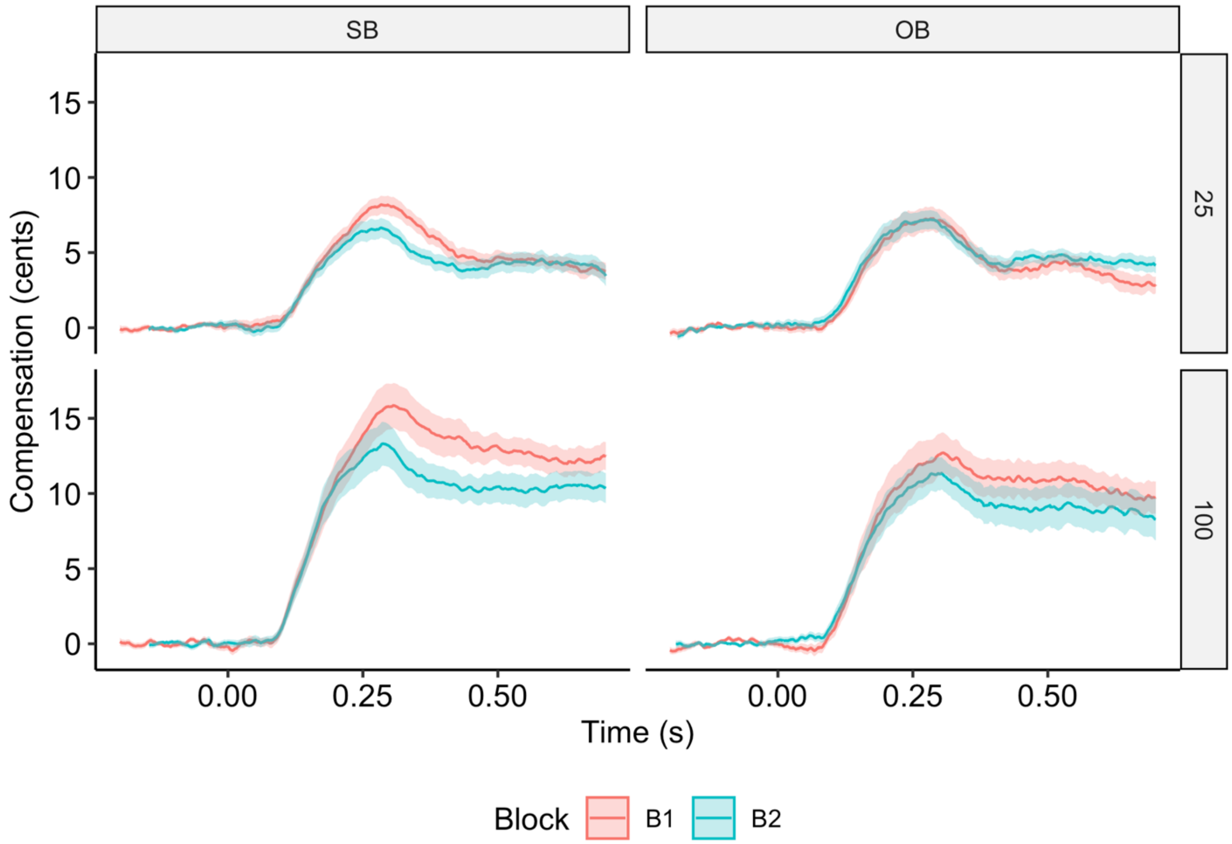


Figure 1. Compensation response pitch contours as a function of experimental block, group, and pitch shift magnitude. (SB = Self-blamers; OB = Other-blamers). Shading represents one standard error above and below the mean.

Response Magnitude. The same pattern is visible in Figure 2, where the compensation response magnitude per condition is visualized. From this figure, it can also be observed there is quite some individual variability with respect to response magnitude. The results of a linear mixed model on the response magnitudes are shown in Table 1. The model included random intercepts over participants, as well as random by-participant slopes for shift magnitude (but no random effects correlations). There were significant main effects of shift magnitude, block, as well as their two-way interaction. The main interaction of interest, block by group, was not significant, although there was a trend. There was no significant three-way interaction.

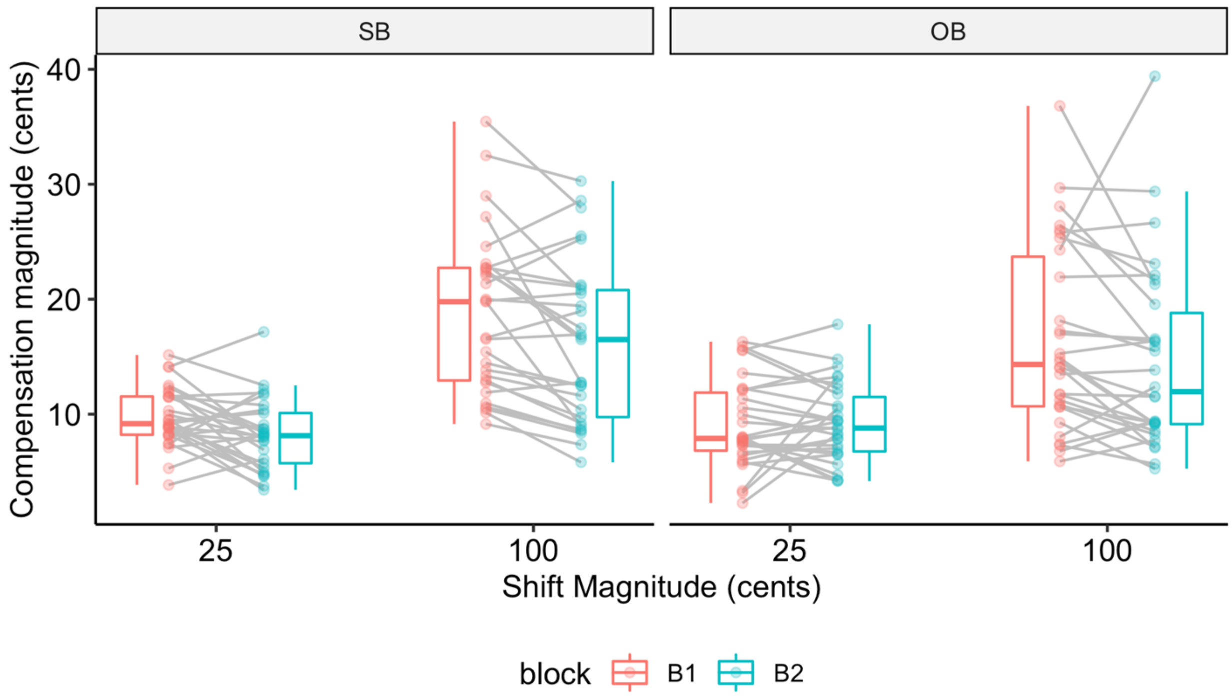


Figure 2. Compensation magnitude as a function of pitch shift magnitude, experimental block, and group. Box plots show the data spread across individuals, while circles represent individual participants' response magnitude.

Table 1. Type III AnoVa table of fixed effects in the linear mixed model ResponseMag ~ ShiftMag*Block*Group + (ShiftMag || Participant)

	SS	NumDF	DenDF	F	p
ShiftMag	**845.56**	**1**	**57**	**104.45**	**< .001***
Block	**130.47**	**1**	**114**	**16.12**	**< .001***
Group	4.20	1	57	0.52	.47
ShiftMag:Block	**37.37**	**1**	**114**	**4.62**	**.034***
ShiftMag:Group	15.42	1	57	1.90	.17
Block:Group	24.00	1	114	2.97	.088
ShiftMag:Block:Group	3.21	1	114	0.40	.53

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

A closer investigation of the two-way interaction between shift magnitude and block showed that response magnitude decreased from Block 1 to Block 2 for ±100 cents shifts (est. = 2.28, $\chi^2(1)$ = 19.00, *p* < .001), but not for ±25 cents (est. = 0.69, $\chi^2(1)$ = 1.74, *p* = .19). Although not strictly significant, the main interaction of interest was also investigated more closely as planned contrasts based on the hypotheses spelled out in the pre-registered analyses. With respect to the within-subject source monitoring contrast, the response magnitude decreased for the self-blamers group from Block 1 to Block 2 (est. = 2.13, $\chi^2(1)$ = 16.18, *p* < .001). This was not the case for the other-blamers, although there was a trend (est. = 0.85, $\chi^2(1)$ = 2.67, *p* = .10). As hypothesized, the difference in response magnitude between Block 1 and Block 2 was numerically larger for the self-blamers than for the other-blamers. The between-participants contrast did not reveal significant results. Both in Block 1 and in Block 2, there was no significant difference between the two groups, although the difference was numerically bigger for Block 1 (Block 1: est. = 1.51, $\chi^2(1)$ = 1.42, *p* = .47; Block 2: est. = 0.24, $\chi^2(1)$ = 0.035, *p* = .85).

Response Direction. Next, a similar analysis was performed on the response types. For every pitch shift, the response was as opposing or following based on the slope of the pitch response contour. Figure 3 shows the proportion of opposing responses per condition. From the figure, it can be observed that descriptively, there seems to be a higher proportion of opposing responses for larger pitch shifts, but there is little difference between blocks or groups. A logistic mixed model was run for statistical inference. The model included random intercepts for participants and random by-participant slopes for block and pitch shift magnitude. The results are shown in Table 2. There was a main effect of pitch shift magnitude, but none of the other factors were significant. For the sake of completeness, we still further investigated the main interaction of interest (Group by Block). The difference between blocks 1 and 2 was not significant in either group, and also numerically similar in size (self-blamers: est. = 0.51, $\chi^2(1)$ = 0.34, *p* > .99; other-blamers: est. = 0.50, $\chi^2(1)$ = 0.11, *p* > .99). The same pattern was observed for the between-groups comparisons per block (block1: est. = 0.49, $\chi^2(1) = 0.021$, $p > .99$; block2: est. = 0.49, $\chi^2(1) = 0.069$, $p > .99$).

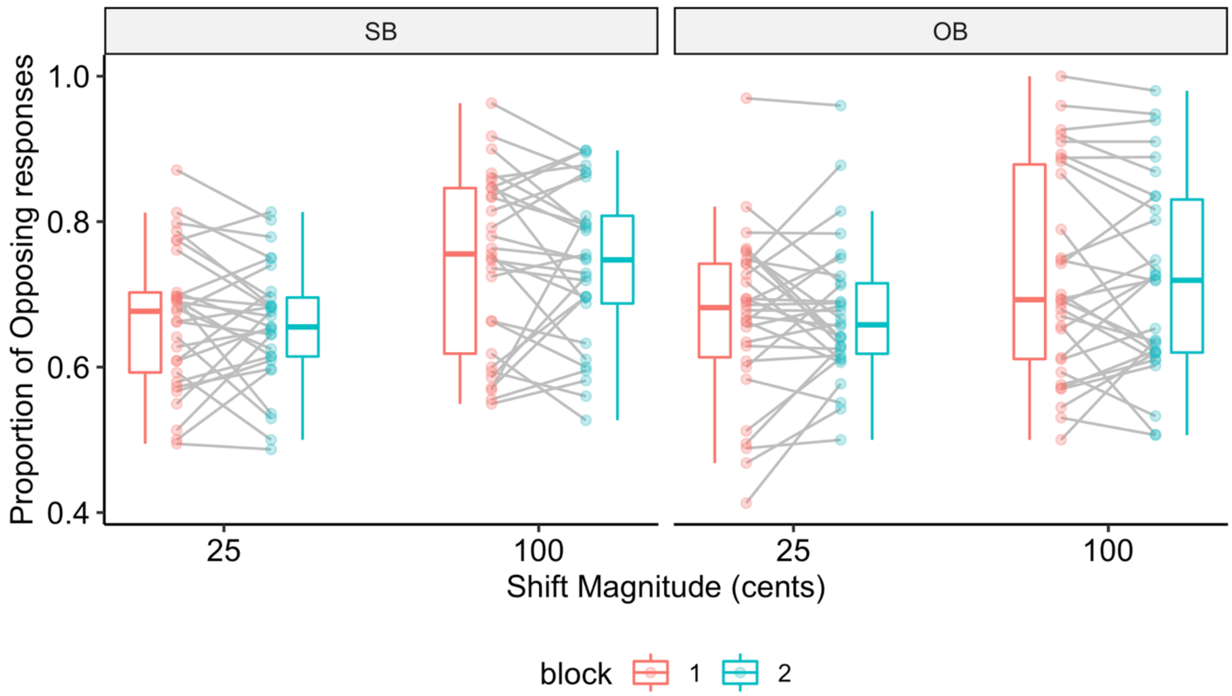


Figure 3. Proportion of opposing responses to pitch shifts, as a function of pitch shift magnitude, experimental block, and group.

Table 2. Type III AnoVa table of fixed effects in the (logistic) linear mixed model ResponseType ~ ShiftMag*Block*Group + (1 + ShiftMag + Block | Participant)

	DF	χ^2	p
ShiftMag	**1**	**38.29**	**< .001***
Block	1	0.41	.52
Group	1	0.044	.83
ShiftMag:Block	1	0.23	.63
ShiftMag:Group	1	0.28	.60
Block:Group	1	0.035	.85
ShiftMag:Block:Group	1	0.35	.55

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

Exploratory analyses

In addition to the results presented above, it is interesting to note that most if not all participants in the current study that were classified as other-blamers reported that even though they knew or suspected the feedback shift was experimentally manipulated, it still felt (sometimes) as if they caused the shift themselves. To determine what factors may play a role in whether a participant is a self- or an other-blamer, the results of the questionnaires and the audiogram were entered in a stepwise logistic regression. The resulting model only included participants' rating on whether they had musical talent and on whether they had a strong Flemish/Dutch accent when speaking another language, as well as an intercept. All of the other independent variables were removed from the model by the stepwise selection procedure. The final model's results are shown in Table 3. The final model's AIC was 79.73, while the intercept-only model had an AIC of 83.77. This suggests that self-blamers were more likely to rate themselves as having less musical talent, and a stronger accent when speaking another language. Group differences on these variables are also illustrated in Figure 4.

Table 3. The results of the final model of the stepwise logistic regression procedure (Group ~ 1 + music_talent + accent).

	Estimate	SE	*z*
(Intercept)	-0.066	0.28	-0.23
music_talent	-0.68	0.30	-2.25
accent	0.48	0.29	1.66

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

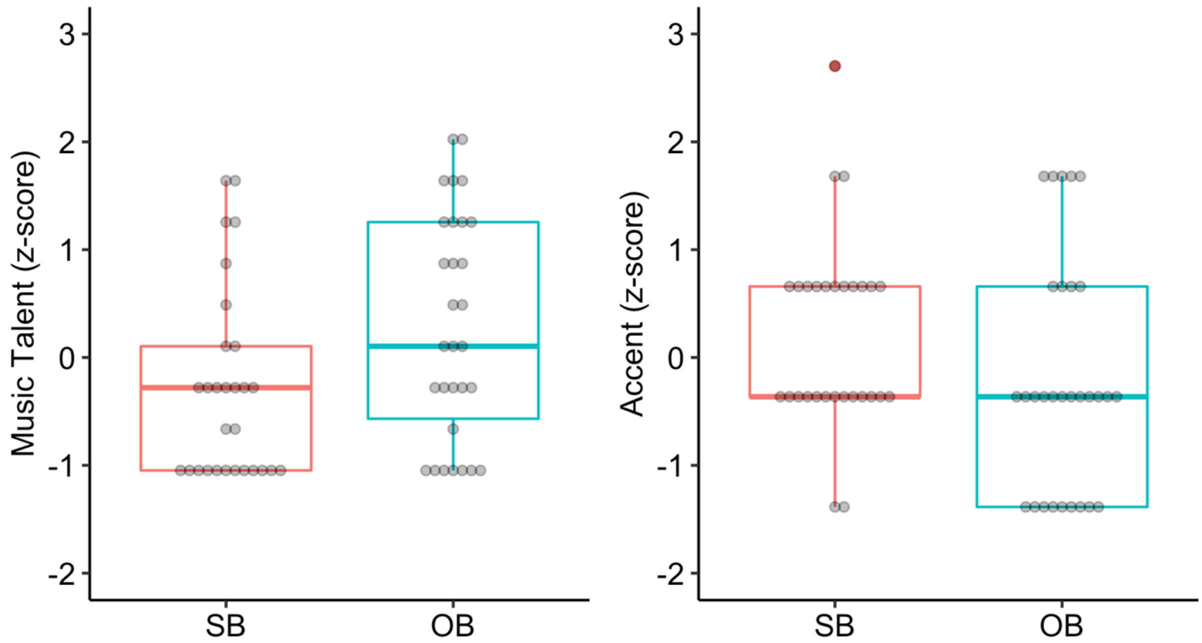


Figure 4. Ratings on musical talent (left) and accent in a foreign language (right) as a function of Group (SB = self-blamer, OB = other-blamer). Grey dots indicate individual data points.

Discussion

The current study investigated the role of conscious source monitoring in online speech motor control. Participants were split into two groups, depending on whether they considered the short pitch shifts in auditory feedback in Block 1 to be self-generated or not. Before Block 2, all participants were informed that the pitch manipulation was performed by the experimenter. There was a significant simple effect of the within-participant manipulation of conscious source monitoring, suggesting that informing participants of the manipulation led to smaller responses in Block 2 compared to Block 1 for the self-blamers. There was no such effect was for the other-blamers, although there was a smaller trend in the same direction. The lack of an interaction between Group and Block shows that there is no statistical evidence that the response magnitude difference between blocks was smaller in the other-

This manuscript has been accepted for publication in Quarterly Journal of Experimental Psychology on January 6, 2022. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without author's permission.

blamers compared to the self-blamers. While the simple effect in the self-blamers suggests that source monitoring leads to top-down modulation of online feedback control during speech production, the lack of an interaction makes the results for the other-blamers hard to interpret.

With respect to response direction, there was no indication that source monitoring affected the proportion of opposing responses. At first sight, this seems in contrast with previous studies linking response direction (following vs. opposing responses) to the considered source of the feedback shift (Hain et al., 2000), although other factors have been shown to influence response direction as well, such as ongoing pitch fluctuations right before feedback is shifted (Franken, Acheson, et al., 2018). It should be noted that previous studies did not focus on conscious source monitoring. It is possible therefore that associations of response direction with source monitoring relate to a non-conscious process rather than a conscious process that can be characterized through introspection. As mentioned in the results section, most if not all participants in the current study that were classified as other-blamers reported that it still felt (sometimes) as if they caused the shift themselves. These results are in line with a distinction between conscious (as revealed by introspection or explicit judgment) and non-conscious (as indicated by a ‘feeling’) source monitoring. This distinction may relate to a similar distinction between the judgment of agency and the feeling of agency proposed by Synofzik et al. (2008). Note that the main effect of pitch shift magnitude on response direction, although small, is in the opposite direction compared to previous studies. We find a higher proportion of opposing responses with 100 cents pitch shifts compared to 25 cents pitch shifts, while previous studies typically find lower proportion of opposing responses with larger pitch shift magnitudes (Behroozmand et al., 2012; Burnett et al., 1998b). These studies, however, show lower proportion of opposing responses from a shift magnitude of around 200 cents, and are thus not in contrast with the current findings.

Overall, then, the current results are partially in line with an association between conscious source monitoring and altered auditory feedback processing. The results suggest that, at least for the self-blamers, conscious source monitoring may lead to top-down modulation of online feedback monitoring. While the experimental design included both a within-participants and a between-participants contrast of source monitoring, the results only showed a within-participants effect of conscious source monitoring: Response magnitudes decreased significantly from Block 1 to Block 2 for the self-blamers, but not for the other-blamers. This suggests that, regardless of individual variability in responses to pitch-shifted feedback, the intervening instructions (i.e., revealing the manipulation to the participant in between blocks) affected these responses for the self-blamers. On the other hand, the between-participants comparison between self-blamers and other-blamers in Block 1 was not significant, and thus there was no interaction between group (self-blamers vs. other-blamers) and block. It is possible that the lack of an interaction effect reflects that there is no effect of group, or that the decreased response in Block 2 is simply a learning effect because the participants are experiencing the same perturbations for a second time. While we cannot completely rule out this possibility, this would not explain why the decrease in response magnitude is only significant for the self-blamers and not for the other-blamers. Alternatively, it is possible that the between-participants effect for Block 1 was obscured by individual variability. The current study assesses what participants considered to be the source of feedback shifts. This was measured through a semi-structured interview after the first experimental block. At this point, it is unclear how noisy the results of this semi-structured interview are, as this method is unable to characterize potential changes in source monitoring that may have occurred over the course of the first experimental block. Such changes in source monitoring over the course of the first block could be a reason that the interaction between Group and Block was not strictly significant, although a clear trend was visible. Specifically, many of the other-blamers could have started out as self-blamers, but became aware of the experimental manipulation over the course of the block. In fact, while other-blaming participants overall found it difficult to pinpoint when they first suspected the manipulation, several participants indicated they only became aware or suspicious of the external manipulation at some point during Block 1. If so, the Block 1 data for the other-blamers would contain a mix of self-blamers and other-blamers' responses, and thus would reflect a quite noisy average response for the other-blamers. This idea, if true, could potentially also explain the lack of a two-way interaction between Block and Group.

These results can be accounted for by adapting the comparator model framework by allowing for top-down modulation. The comparison between predicted sensory feedback and observed feedback, which is the core principle of the comparator framework, needs some threshold in order to determine what degree of mismatch counts as different enough to be labelled as caused by an external source. The current results are in line with a model where prior knowledge of the manipulation could bias this threshold, such that a smaller difference between prediction and feedback could lead to labeling the stimulus as externally generated, or with a model where top-down modulation affects the gain on feedback control. The current results do not allow to distinguish between these possibilities. Previous studies suggest that the threshold for unexpectedly pitch-shifted feedback may lie around 200-250 cents (Scheerer et al., 2013). Results in the current study showed an overall effect of Shift Magnitude, with larger responses to 100 cents shifts compared to 25 cents shifts. This is in line with increasing response magnitude with increasing shift magnitude. However, shifts above this 'threshold' of 200-250 cents show the reverse pattern, with smaller responses to the larger shifts (Franken et al., 2021; Korzyukov et al., 2017; Scheerer et al., 2013; Subramaniam et al., 2018). The smaller responses to these larger shifts have been suggested to be indicative of the speaker no longer recognizing the shift as self-produced. It is to date unclear whether task instructions such as in the present study could change this threshold value, and further studies would be necessary to further investigate this issue.

Alternatively, the results could be accounted for by a so-called *inferential* model. In contrast to the comparator model, the inferential model holds that source monitoring relies on the joint contribution of multiple sources of information, and not just the comparison between predicted and observed sensory feedback (Dennett, 1991; Lind et al., 2014; Wegner & Wheatley, 1999). In their theory of apparent mental causation, Wegner and Wheatley (1999) propose that three principles govern the sense of agency (and thus, presumably, source monitoring). These principles are *priority*, *consistency*, and *exclusivity* of the intention to act. In other words, there needs to be an intention before the action (priority), this intention needs to be compatible with the action (consistency), and this intention should be the only apparent cause of the action (exclusivity). This inferential account would allow for top-down modulation, as the conscious knowledge that the feedback was pitch-shifted violates the exclusivity principle. This rather “categorical” formulation could be naturally extended into a Bayesian account. From this point of view, revealing the nature of the experimental manipulation in between the experimental blocks could have led to a biased prior in a Bayesian process that estimates the most likely source of auditory stimuli during vocalization in Block 2. This idea is reminiscent of previous accounts of Bayesian error attribution in non-speech sensorimotor processing (Berniker & Kording, 2008; Wei & Körding, 2009).

The current results suggest that responses to unexpectedly pitch-shifted auditory feedback are sensitive to top-down control. At first sight, this seems in contrast with previous studies claiming that these responses are automatic (Burnett et al., 1998b; Hafke, 2008; Hain et al., 2000). However, automaticity does not necessarily preclude top-down control (Kiefer, 2008). As such, the current results are in line with recent work outside of the speech motor literature suggesting that automatic processes could be modulated by top-down control. It is useful in this regard to distinguish, in line with Ansorge and Horstmann (2007), between preemptive control and reactive control. While the processes involved in online feedback control may be automatic in the sense that they are not susceptible to top-down modulation or correction once the process has started (reactive control), task instructions or conscious knowledge about the experimental manipulation could lead to top-down changes in the system's configuration before the process has started, in such a way that some processing steps are enhanced or blocked (preemptive control). As such, informing the self-blamers about the experimental manipulation could have led to a preemptive change in the speech system's gain on auditory input during speech production, or to a change in bias for distinguishing between self-caused and externally generated stimuli.

The main research question of the current study was about how conscious source monitoring affected feedback processing during speech production. The main (pre-registered) analyses therefore were predicated on the assumption that without prior instruction, some participants would think the pitch-shifted feedback was self-produced (the self-blamers), while others would suspect pitch shifts to be actively manipulated by the researcher (the other-blamers). Additionally, we performed exploratory analyses to examine some of the factors that could be related to this individual variability between self-blamers and other-blamers. Of all the factors explored, only participants' self-ratings on their musical talent and their accentedness when speaking a foreign language were able to explain part of this variability. Self-blamers tended to rate themselves as having less musical talent and stronger accents, compared to other-blamers. This is in line with a pattern of vocal abilities (self-reports of more musical talent, or less foreign accent) being associated with smaller responses to pitch-shifted feedback. Several studies have previously shown that musicianship affects responses to pitch-shifted feedback (Behroozmand et al., 2014; Jones & Keough, 2008; Keough & Jones, 2009; Sturgeon et al., 2015; Zarate et al., 2010; Zarate & Zatorre, 2005, 2008). It has been suggested that higher vocal abilities may be associated with increased reliance on internal predictions and less on online auditory feedback (Jones &

Keough, 2008). For example, Zarate and Zatorre (2008) showed that experienced singers were more able to ignore pitch-shifted feedback, compared to non-musicians, indicating that singers relied less on auditory feedback for vocal control. Subramaniam et al. (2018) suggest that stronger reliance on internal predictions leads to a better distinction between self-generated and external stimuli. In addition, Scheerer and Jones (2012) showed that lower pitch variability during a baseline (indicating enhanced pitch control) is associated with smaller responses to pitch shifts. We are not aware of direct evidence linking self-perception of foreign accent to speech motor control, although some studies suggest that speech in a second language may be more vulnerable to delayed auditory feedback (Howell & Dworzynski, 2001; Mackay, 1970), suggesting that more vulnerable speech systems (such as when speaking a second language) may be more reliant on auditory feedback. Together with these findings, the current results indicate that participants with better musical skills or less foreign accents when speaking a second language may rely more on internal predictions than on online feedback, leading to smaller responses. The current study adds that these participants are also more likely to blame the pitch shift on the experiment(er) rather than on themselves (i.e., these are other-blamers).

The current study tried to categorize participants as self- or other-blamers based on a semi-structured interview. A limitation of this approach is that it is unclear what the reliability of the resulting participant classification is. Specifically, the results of pitch response analyses show that participants responded to unexpected pitch-shifted feedback by changing their pitch in the opposite direction. This means that during the interview, when participants refer to pitch changes or voice breaks, one could argue that it is not clear whether participants refer to the applied pitch shift, or to their response to the shift. In case of the latter, participants may be unjustly classified as self-blamers (their own response is of course self-caused, while the shift is not). However, the responses are much smaller in magnitude than the applied pitch shift, Furthermore, while responses include both auditory and somatosensory changes and the shifts are only auditory, the shift is abrupt and the response is gradual, which arguable makes the shift more salient compared to the response. Therefore, we maintain it is a fair assumption that if participants mention changes in their voice or pitch, they most likely refer to the pitch shifts in the auditory feedback. Nevertheless, we cannot completely rule out this possibility based on the current data.

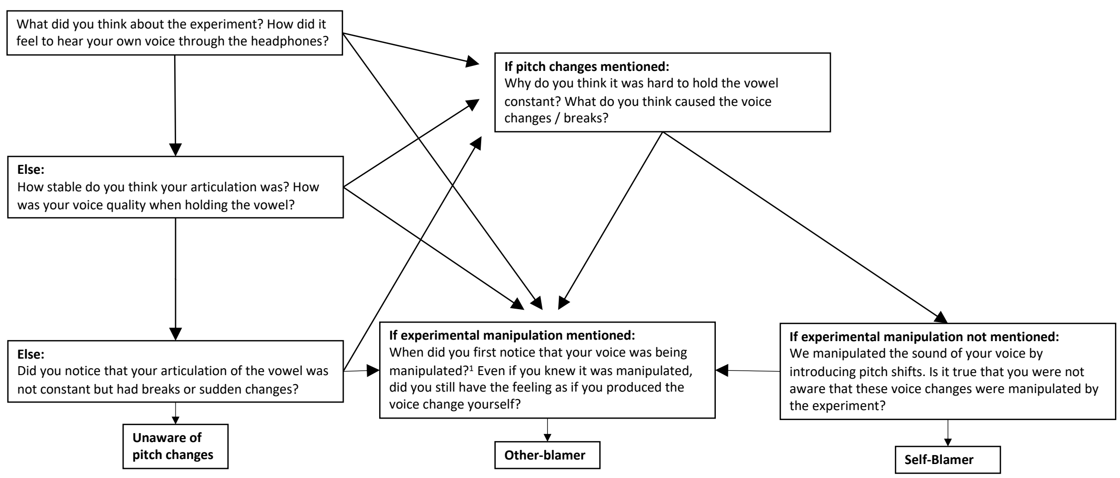
In conclusion, the current study aimed to investigate whether conscious source monitoring leads to top-down modulation of online feedback control during speech production. Although the hypothesized interaction was not significant, there were clear indications that for participants who blamed themselves for pitch errors, explicit instructions that pitch shifts were manipulated led to smaller responses to pitch-shifted feedback. This indicates that auditory-driven articulatory responses are at least partially modulated by conscious source monitoring. The results indicate that there is individual variability in how participants experience pitch-shifted feedback, which may be driven by variability in speakers' vocal abilities. The results are in line with recent views on top-down modulation of automatic perceptual processing.

Acknowledgement

We are grateful to Pascal Mestdagh for his help with acoustic measurements and experimental setup, and to Tania Selleslags and Urgent FM for the use of their lab. M.K.F. was supported by the Special Research Fund (BOF) at Ghent University. A.L. was supported by the Swedish Research Council (Grant No. 2015-00446) and the Crafoord Foundation (Grant No. 20170734), L.H. and P.J. were supported by the Swedish Research Council (Grant No. 2014-1371) and the Crafoord Foundation. The authors declare that there is no conflict of interest.

Appendix: Semi-structured Interview

The flow chart below illustrates the questions during the semi-structured interview. The first question was always the top left, and succeeding questions were determined based on the participant's responses (specifically, whether the participant mentioned pitch changes in the auditory feedback, or the fact that there was an experimental manipulation, or neither).



[1]The participant was asked to indicate this on a visual progress bar that was similar to the progress bar that was visible on the screen during the experimental block.

References

Ansorge, U., & Horstmann, G. (2007). Preemptive control of attentional capture by colour: Evidence from trial-by-trial analyses and orderings of onsets of capture effects in reaction time distributions. *Quarterly Journal of Experimental Psychology*, *60*(7), 952–975. https://doi.org/10.1080/17470210600822795

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

Bauer, J. J., & Larson, C. R. (2003). Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: Improvements in methodology for the pitch-shifting technique. *The Journal of the Acoustical Society of America*, *114*(2), 1048. https://doi.org/10.1121/1.1592161

Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A., & Larson, C. R. (2014). Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain and Cognition*, *84*(1), 97–108. https://doi.org/10.1016/j.bandc.2013.11.007

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. *The Journal of the Acoustical Society of America*, *132*(4), 2468–2477. https://doi.org/10.1121/1.4746984

Berniker, M., & Kording, K. (2008). Estimating the sources of motor errors for adaptation and generalization. *Nature Neuroscience*, *11*(12), 1454–1461. https://doi.org/10.1038/nn.2229

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242. https://doi.org/10.1016/S1364-6613(02)01907-1

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer [Computer Program]* (6.0.33).

Brosschot, J. F., Gebhardt, W. A., & Godaert, G. L. R. (1994). Internal, powerful others and chance locus of control: Relationships with personality, coping, stress and health. *Personality and Individual Differences*, *16*(6), 839–852. https://doi.org/10.1016/0191-8869(94)90228-3

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998a). 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., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998b). 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., & Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America*, *112*(3), 1058. https://doi.org/10.1121/1.1487844

Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F[sub 0] responses to pitch-shifted voice feedback during English speech. *The Journal of the Acoustical Society of America*, *121*(2), 1157–1163. https://doi.org/10.1121/1.2404624

Dennett, D. Clement. (1991). *Consciousness explained*. Penguin Books.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

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

Franken, M. K., Eisner, F., Acheson, D. J., McQueen, J. M., Hagoort, P., & Schoffelen, J. (2018). Self-monitoring in the cerebral cortex: Neural responses to small pitch shifts in auditory feedback during speech production. *NeuroImage*, *179*(June), 326–336. https://doi.org/10.1016/j.neuroimage.2018.06.061

Franken, M. K., Hartsuiker, R. J., Johansson, P., Hall, L., & Lind, A. (2021). Speaking with an alien voice: Flexible sense of agency during vocal production. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(4), 479–494. https://doi.org/10.1037/xhp0000799

Franken, M. K., Hartsuiker, R. J., Johansson, P., Hall, L., Wartenberg, T., & Lind, A. (2019). Does passive sound attenuation affect responses to pitch-shifted auditory feedback? *The Journal of the Acoustical Society of America*, *146*(6), 4108–4121. https://doi.org/10.1121/1.5134449

Frith, C. (2005). The self in action: Lessons from delusions of control. *Consciousness and Cognition*, *14*(4), 752–770. https://doi.org/10.1016/J.CONCOG.2005.04.002

Greenlee, J. D. W., Behroozmand, R., Larson, C. R., Jackson, A. W., Chen, F., Hansen, D. R., Oya, H., Kawasaki, H., & Howard, M. A. (2013). Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PloS One*, *8*(4), e60783. https://doi.org/10.1371/journal.pone.0060783

Guenther, F. H. (2016). *Neural Control of Speech*. The MIT Press.

Hafke, H. Z. (2008). Nonconscious control of fundamental voice frequency. *The Journal of the Acoustical Society of America*, *123*(1), 273–278. https://doi.org/10.1121/1.2817357

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207. https://doi.org/10.1038/nrn.2017.14

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*(2), 133–141. https://doi.org/10.1007/s002219900237

Hawco, C. S., Jones, J. A., Ferretti, T. R., & Keough, D. (2009). ERP correlates of online monitoring of auditory feedback during vocalization. *Psychophysiology*, *46*(6), 1216–1225. https://doi.org/10.1111/j.1469-8986.2009.00875.x

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216.

Howell, P., & Dworzynski, K. (2001). Strength of German accent under altered auditory feedback. *Perception & Psychophysics*, *63*(3), 501–513. https://doi.org/10.3758/BF03194416

Jones, J. A., & Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Experimental Brain Research*, *190*(3), 279–287. https://doi.org/10.1007/s00221-008-1473-y

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

Keough, D., & Jones, J. A. (2009). The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *The Journal of the Acoustical Society of America*, *126*(2), 837–846. https://doi.org/10.1121/1.3158600

Kiefer, M. (2008). Top-down modulation of unconscious “automatic” processes: A gating framework. *Advances in Cognitive Psychology*, *3*(1–2), 289–306. https://doi.org/10.2478/v10053-008-0032-2

Korzyukov, O., Bronder, A., Lee, Y., Patel, S., & Larson, C. R. (2017). Bioelectrical brain effects of one’s own voice identification in pitch of voice auditory feedback. *Neuropsychologia*, *101*, 106–114. https://doi.org/10.1016/j.neuropsychologia.2017.04.035

Korzyukov, O., Karvelis, L., Behroozmand, R., & Larson, C. R. (2012). ERP correlates of auditory processing during automatic correction of unexpected perturbations in voice auditory feedback. *International Journal of Psychophysiology*, *83*(1), 71–78. https://doi.org/10.1016/j.ijpsycho.2011.10.006

Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., & Hain, T. C. (2001a). Comparison of voice F[sub 0] responses to pitch-shift onset and offset conditions. *The Journal of the Acoustical Society of America*, *110*(6), 2845. https://doi.org/10.1121/1.1417527

Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., & Hain, T. C. (2001b). Comparison of voice F[sub 0] responses to pitch-shift onset and offset conditions. *The Journal of the Acoustical Society of America*, *110*(6), 2845. https://doi.org/10.1121/1.1417527

Li, W., Chen, Z., Liu, P., Zhang, B., Huang, D., & Liu, H. (2013). Neurophysiological evidence of differential mechanisms involved in producing opposing and following responses to altered auditory feedback. *Clinical Neurophysiology*, *124*(11), 2161–2171. https://doi.org/10.1016/j.clinph.2013.04.340

Lind, A., Hall, L., Breidegard, B., Balkenius, C., & Johansson, P. (2014). Auditory feedback of one’s own voice is used for high-level semantic monitoring: The “self-comprehension” hypothesis. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00166

Liu, H., Wang, E. Q., Chen, Z., Liu, P., Larson, C. R., & Huang, D. (2010). Effect of tonal native language on voice fundamental frequency responses to pitch feedback perturbations during sustained vocalizations. *The Journal of the Acoustical Society of America*, *128*(6), 3739–3746. https://doi.org/10.1121/1.3500675

Mackay, D. G. (1970). How Does Language Familiarity Influence Stuttering under Delayed Auditory Feedback? *Perceptual and Motor Skills*, *30*, 655–669.

Puckette, M. (1996). Pure data. *Proceedings of the International Computer Music Conference*, 269–272.

Scheerer, N. E., Behich, J., Liu, H., & Jones, J. A. (2013). ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience*, *240*, 176–185. https://doi.org/10.1016/j.neuroscience.2013.02.054

Scheerer, N. E., & Jones, J. A. (2012). The relationship between vocal accuracy and variability to the level of compensation to altered auditory feedback. *Neuroscience Letters*, *529*(2), 128–132. https://doi.org/10.1016/j.neulet.2012.09.012

Sturgeon, B. A., Hubbard, R. J., Schmidt, S. A., & Loucks, T. M. (2015). High F0 and musicianship make a difference: Pitch-shift responses across the vocal range. *Journal of Phonetics*, *51*, 70–81. https://doi.org/10.1016/j.wocn.2014.12.001

Subramaniam, K., Kothare, H., Mizuiri, D., Nagarajan, S. S., & Houde, J. F. (2018). Reality Monitoring and Feedback Control of Speech Production Are Related Through Self-Agency. *Frontiers in Human Neuroscience*, *12*, 82. https://doi.org/10.3389/fnhum.2018.00082

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219–239. https://doi.org/10.1016/j.concog.2007.03.010

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, *423*(6942), 866–869. https://doi.org/10.1038/nature01710

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*(7), 480–492. https://doi.org/10.1037/0003-066X.54.7.480

Wei, K., & Körding, K. (2009). Relevance of error: What drives motor adaptation? *Journal of Neurophysiology*, *101*(2), 655–664. https://doi.org/10.1152/jn.90545.2008

Wolpert, D., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat Neurosci*, *3*(Suppl), 1212–1217.

Zarate, J. M., Wood, S., & Zatorre, R. J. (2010). Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*, *48*(2), 607–618. https://doi.org/10.1016/j.neuropsychologia.2009.10.025

Zarate, J. M., & Zatorre, R. J. (2005). Neural substrates governing audiovocal integration for vocal pitch regulation in singing. *Annals of the New York Academy of Sciences*, *1060*, 404–408. https://doi.org/10.1196/annals.1360.058

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage*, *40*(4), 1871–1887. https://doi.org/10.1016/j.neuroimage.2008.01.026