4aSCa5. The relationship between vocal pitch feedback error and event-related brain potentials.

Jeffery A. Jones*, Nichole Scheerer and Anupreet Tumber

*Corresponding author's address: Psychology Department, Wilfrid Laurier University, Waterloo, N2L 3C5, Ontario, Canada, jjones@wlu.ca

Understanding the neural processing of auditory feedback during speech is essential to the development of a comprehensive model of speech motor control. Currently the relationship between the magnitude of errors detected in feedback and the evoked neural responses is unclear. We exposed speakers to sudden changes in vocal pitch that ranged from 0 to 400 cents in magnitude. Vocal responses and auditory event-related potentials (ERPs: P1-N1-P2-N2 components) were measured. Results showed that vocal response magnitudes were relatively consistent when speakers were exposed to small feedback perturbations (<250 cents). Larger perturbations (>300 cents) caused decreased vocal response magnitudes. P1 amplitudes showed a non-specific increase when feedback was perturbed. N1 amplitudes demonstrated more specificity: smaller feedback perturbations evoked one size of response, while larger feedback perturbations elicited a larger response. P2 amplitudes increased with increases in the feedback perturbation magnitude. Moreover, a reliable relationship existed between vocal response magnitude and P2 amplitude: vocal response magnitude and P2 amplitude increased in response to perturbations between 50 and 250 cents, and then decreased in response to larger perturbations. ERPs allow us to hypothesize the stages of processing. Results will be discussed with respect to perceptual and production thresholds, and implications for speech motor control.

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

Speakers rely on sensory feedback during speech production to learn to produce novel sounds and to detect errors in their productions. Auditory feedback is particularly important to ensure that the sounds produced and then heard by listeners are as the speaker intended. Numerous studies have shown that speakers respond to feedback alterations by changing their speech output in ways that oppose the feedback alteration (Burnett, Freedland, Larson, & Hain, 1998; Jones & Munhall, 2002; Munhall, MacDonald, Byrne, & Johnsrude, 2009; Scheerer, Behich, Liu, & Jones, 2013). For example, when speakers hear their vocal pitch (F0) suddenly shifted up or down, they respond by adjusting their F0 in the opposite direction (Burnett et al., 1998). This pitch shift reflex (PSR) appears to be optimized for correcting for small deviations in auditory feedback, while larger feedback perturbations are only partially compensated. As implied by its ‘reflexive’ description, the PSR occurs in a highly automatic fashion. For example, Munhall and colleagues (2009) found that speakers compensated for alterations of their first and second formants in the word “head”, even if they were explicitly instructed to ignore the auditory stimulus or avoid compensation. Our lab has found similar results when trained and untrained vocalists were asked to ignore vocal pitch perturbations they heard (Keough, Hawco, & Jones, 2013). Together these results suggest that the PSR functions in a highly automated manner to stabilize voice fundamental frequency around a desired target.

Recently, a number of investigators have asked whether there is a relationship between compensation to perturbed auditory feedback and a speaker’s awareness of the change in auditory stimuli. There are numerous reasons to investigate the relationship between the thresholds for perception and production. For example, there are significant individual differences observed when speakers are exposed to altered auditory feedback. It is possible that when speakers become aware of errors, they choose different strategies to negate these errors online and in future productions. These different strategies may underlie some of the variability across individuals.

Hafke (2008) investigated the threshold for compensation to pitch perturbations and the threshold at which trained singers reported hearing the perturbations. In Hafke’s experiment, the singers produced vocalizations while hearing varying magnitudes of pitch shifted auditory feedback. After each vocalization participants reported whether or not they perceived a change in their pitch during the vocalization. The results of the study suggested that on average, participants’ perceptual thresholds for detection of pitch alterations were around 21 cents (where 100 cents is equivalent to a semitone). On the other hand, participants compensated to pitch shifts as small as 9 cents, suggesting that the motor system is more sensitive to pitch changes in auditory feedback. Hafke (2008) suggested that the different magnitudes of feedback alterations required perceptual awareness and a motor response, providing support for a perception-action dissociation and dual stream processing of auditory information. Hickok and Poeppel (2004) proposed two separate auditory processing streams (analogous to the two processing streams in vision; Ungerleider & Mishkin, 1982; Milner & Goodale, 1995): a dorsal stream that projects from the auditory cortex to the parietal-temporal boundary of the sylvian fissure (area Spt), which plays a role in auditory-motor transformations, and a ventral stream that projects from the auditory cortex towards the inferior posterior temporal cortex, which plays a role in linking auditory information to conceptual representations. However, a simpler explanation for the differences between perception and production observed in these threshold studies is that the perceptual responses and error induced compensatory responses are based on different signal detection criteria (Green & Swets, 1966).

Like all biological systems, variability in neural processing and speech output means that the system generating vocal compensations and the system underlying awareness of vocal mistakes will only designate deviations that exceed the normal variability in an individual’s voice as errors. It is possible that the criteria for triggering a compensation response or a conscious report of a vocal perturbation are different. Thus, previous reports of a motor and perceptual processing separation may simply reflect a difference in the magnitude of feedback alterations required to elicit a motor response and for an individual to report a pitch change, rather than the early separation of processing of externally generated and self-produced sounds.

Previous studies have determined perceptual thresholds by interpolating responses based on measured responses to induced feedback alterations of a small number of magnitudes. For the present study, participants were presented a variety of brief pitch perturbations between 0 and 40 cents in magnitude to obtain a more precise estimate of the degree of compensation to small perturbations. In addition, participants reported whether or not they detected a change in their pitch following each vocalization to determine the threshold at which participants reported errors in their auditory feedback. Based on previous work, we expected that the magnitude of feedback perturbation required to elicit awareness of pitch perturbations would be larger than the magnitude required to elicit a motor response. In addition, event-related potentials (ERPs) were recorded in attempt to identify relationships between the vocal and behavioural responses reported to the feedback perturbations, as well as elicited neural responses.
METHODS

Participants

Twenty-five participants between the ages of 18-20 years (mean=18.63, SD=0.74; 15 female and 10 male) participated in this study. Participants were right-handed native Canadian English speakers and they did not speak a tonal language. Prior to testing, the participants’ hearing was assessed using a diagnostic audiometer to ensure hearing thresholds ≤35dB HL at 250, 500, 1000, 2000, 4000 and 8000 Hz. One participant demonstrated a 40 dB hearing threshold for frequencies above 1000 Hz. Participants also completed a music experience questionnaire, and one participant reported having formal vocal training. Informed consent was obtained from all participants according to the ethical policies at Wilfrid Laurier University. All participants received course credit for their participation in this study.

Procedure

Participants vocalized the vowel /a/ 130 times over five blocks, while exposed to normal and frequency altered feedback (FAF). After each vocalization participants pressed designated buttons on a keyboard to indicate whether or not they believed the pitch of their voice was altered. Participants were cued to start vocalizing by a green box on screen, while a red box signified they should stop vocalizing, at which time they were to indicate whether or not they heard their voice shifted in pitch. The first block of the experiment was a practice block of 10 trials. The participants were informed that during the first five trials they would hear their normal, unaltered voice, while the last five trials would be altered. During the five altered trials in the practice block, participants’ auditory feedback was perturbed 40, 20, 30, 5, and then 25 cents, six times per vocalization in order to expose participants to the range of perturbations they would hear during the experimental blocks. Following the practice block, participants performed four experimental blocks. Each experimental block contained 30 trials where the participants’ auditory feedback was perturbed 5, 10, 15, 20, 25, 30, or 40 cents downwards, six times per vocalization, or left unaltered. Each perturbation had a fixed duration of 200 ms and occurred with an inter-stimulus interval of 700-900 ms, resulting in vocalizations that were approximately seven seconds in length. The perturbation magnitude remained consistent for the six perturbations per vocalization, but was pseudo-randomly presented throughout the experiment. This resulted in 90 trials per perturbation magnitude.

Apparatus

Participants wore a HydroCel GSN 64 1.0 electrode cap (Electrical Geodesics, Inc., Eugene, OR), Etymotic ER-3 insert headphones (Etymotic Research, Elk Grove Village, IL), and a headset microphone (Countryman Isomax E6 Omnidirectional Microphone), and were tested in an electrically shielded booth. Voice signals were sent to a mixer (Mackie Oynx 1220, Loud Technologies, Woodinville, WA), followed by a digital signal processor (DSP; VoiceOne, T.C. Hellicon, Victoria, BC), which altered the fundamental frequency of the voice signals. The altered voice signals were then presented back to the participant through headphones as altered auditory feedback. The unaltered voice signal was digitally recorded (TASCAM HD-P2, Montebello, CA) at a sampling rate of 44.1 KHz for later analysis.

Behavioural Recording and Analysis

The unaltered voice signal was segmented into separate vocalizations and F0 values were calculated for each vocalization using the SWIPE’ algorithm (Camacho & Harris, 2008). Each vocalization was then segmented based on the onset of the six perturbations. F0 values for each of the six perturbed segments were normalized to the baseline period, which was the portion of the segment 200 ms prior to the onset of the perturbation, by converting Hertz values to cents using the following formula:

\[ \text{Cents} = 100(12 \log_2 \frac{F}{B}) \]

In the formula, F is the F0 value in Hertz and B is the mean frequency of the baseline period.

Cents values were calculated for the 200 ms prior to the perturbation (the baseline period), and 1000 ms after the perturbation. An averaged F0 trace was constructed for each perturbation magnitude, 0 (unaltered), 5, 10, 15, 20, 25, 30, and 40, for each participant.
Vocal responses were quantified by examining the response magnitude and latency. The magnitude of compensation was determined by finding the point at which each participant’s averaged F0 trace deviated maximally from its baseline mean, and the latency was calculated as the time at which this maximal deviation occurred. Behavioural responses were quantified by calculating the hit rate, defined as the proportion of detected perturbations, relative to total perturbations, at each perturbation magnitude.

**ERP Recording and Analysis**

EEG signals were recorded from 64 scalp electrodes and referenced online to the vertex (Cz) electrode. Data were band-pass filtered (1-30Hz) and digitized (12-bit precision) at 1000 samples per second. Electrode impedances were maintained below 50 kΩ throughout the duration of the experiment (Ferree et al., 2001). After data acquisition, EEG voltage values were re-referenced to the average voltage across all sites. The data were then epoched into segments from 100 ms before the onset of the perturbation to 500 ms after perturbation onset. Data were analyzed offline for movement artifacts and any segment with voltage values exceeding 55 µV of the moving average over an 80 ms span were rejected. In addition, a visual inspection of all data was completed to ensure artifacts were being adequately detected. Two subjects were eliminated from further analyses, as they had less than 50% of their trials retained across all perturbation magnitudes. Across all other participants, on average, between 69-72 trials (77-80%) were retained for each perturbation magnitude.

Six electrodes were included in the analysis: Fz, Cz, F3, C3, C4, and F4. These electrodes were then grouped into left: average of F3 and C3, medial: average of Fz and Cz, and right: average of F4 and C4. These electrodes were chosen based on visual inspection of the regions demonstrating the most robust P1-N1-P2 components, as well as previous research suggesting that fronto-medial and centro-frontal regions are the maximal generators of these components in response to feedback perturbations (Behroozmand et al., 2009; Chen et al., 2012; Korzyukov et al., 2012, Scheerer et al., 2013).

For each participant, averaged waveforms were created for the unaltered and the FAF conditions for each electrode. Grand averaged waveforms were created for all conditions by averaging the data from all participants for each electrode, followed by baseline correction. For all average files for each participant, the maximum amplitude and latency were calculated for the ERP components of the P1-N1-P2 complex. Based on visual inspection of the latency of the most prominent ERP peaks, these components were extracted at time windows from 50-100ms, 100-200ms, and 200-300ms, respectively.

**Statistical Analysis**

Separate 8 (perturbation magnitude) x 3 (electrode group) repeated-measures analysis of variances (RM-ANOVAs) were conducted on the amplitudes of the P1-N1-P2 complex as well as vocal response magnitudes. Similarly, 7 (perturbation magnitude) x 3 (electrode group) RM-ANOVAs were conducted on the latencies of the P1-N1-P2 complex, vocal response magnitudes, as well as hit rates. The unaltered (0 cent) condition was not analyzed with regards to ERP and vocal latencies, as stimuli were not presented during the unaltered trials, thus data were randomly sampled with no true reference, rendering latency information meaningless. Similarly, hit rates were not calculated for the unaltered (0 cent) trials, as there were no stimuli to detect. Preplanned polynomial contrasts were performed in order to investigate any significant response trends as a result of the increasing feedback perturbation magnitudes. In addition, post-hoc least significant difference (LSD) tests were conducted to tease apart the effects of increasing feedback perturbation magnitudes on recorded responses. Correlational analyses were performed, in order to investigate the relationships between the vocal, behavioural, and ERP data. The Greenhouse-Geisser (Greenhouse & Geisser, 1959) correction was used in instances where Mauchley’s Assumption of Sphericity was violated. However, original degrees of freedom were reported for ease of interpretation.

**RESULTS**

**Vocal Responses**

A RM-ANOVA conducted to investigate the effect of perturbation magnitude on vocal response magnitude showed a significant main effect of perturbation magnitude (F(7,154)=24.301, p<.001). In addition, polynomial contrasts highlighted a significant linear trend (F(1,22)=49.186, p<.001) (see Figure 1). Pairwise comparisons indicated that with exception of the 5-cent perturbation, all perturbation magnitudes elicited larger vocal responses...
than the unaltered condition (p<.05). The 40-cent perturbation resulted in significantly larger responses than the 30, 25, 20, 15, 10 and 5-cent perturbations (p<.05). This trend continued as the 30-cent perturbation elicited larger vocal responses than the 20, 15, 10, and 5-cent perturbations (p<.05). The 25-cent perturbation resulted in larger vocal response magnitudes than the 15, 10, and 5-cent perturbations (p<.01). The 20-cent perturbation elicited larger vocal responses than the 10 and 5-cent perturbations (p<.05), while the 15 and 10-cent perturbations resulted in larger vocal responses than the 5-cent perturbation (p<.05).

The RM-ANOVA investigating the influence of perturbation magnitude on vocal response latency failed to reach significance (F(7,154)=.778, p=.564).

![Averaged F0 traces for the eight feedback perturbation magnitudes.](image)

**FIGURE 1.** Averaged F0 traces for the eight feedback perturbation magnitudes.

### Behavioural Results

A RM-ANOVA was conducted to investigate the influence of perturbation magnitude on the proportion of feedback perturbations detected, or the hit rate. A significant main effect of perturbation magnitude was found (F(6,132)=18.213, p<.001) (see Figure 2). Pairwise comparisons indicate that all perturbations magnitudes were detected more readily than the 5 and 10-cent perturbations (p<.001), while the 5 and 10-cent perturbations did not result in significantly different detection rates (p=.09). In addition, the 40-cent perturbation was detected more readily than all other perturbation magnitudes (p<.01).

### ERP Results

**P1 Amplitudes and Latencies**

A two-way RM-ANOVA was conducted to investigate the influence of perturbation magnitude and electrode site on P1 amplitudes. Significant main effects of perturbation magnitude (F(7,154)=3.732, p=.006) and electrode site (F(2,44)=8.264, p=.003) were found. However, the interaction between perturbation magnitude and electrode site failed to reach significance (F(14,308)=1.111, p=.359). Pairwise comparisons indicate that all perturbation magnitudes elicited larger P1 amplitudes than the 0-cent (unaltered) condition (see Figure 3). In addition, the 40-cent perturbation elicited significantly larger P1 amplitudes than the 15-cent perturbation. Pairwise comparisons also indicated that P1 amplitudes were largest at the medial electrode sites, relative to the left (p<.001), or right (p=.045) sites.

The two-way RM-ANOVA investigating the influence of perturbation magnitude and electrode site on P1 latency failed to find a significant effect of perturbation magnitude (F(6,132)=.937, p=.471), site (F(2,44)=.297, p=.648), or an interaction between perturbation magnitude and site (F(12,264)=1.138, p=.330).
FIGURE 2. Hit rate (percentage of detected feedback perturbations) across each perturbation magnitude.

N1 Amplitudes and Latencies

A two-way RM-ANOVA was conducted to investigate the influence of perturbation magnitude and electrode site on N1 amplitudes. A significant main effect of perturbation magnitude (F(7,154)=2.768, p=.010) was found. However, the main effect of electrode site (F(2,44)=.183, p=.833) and the perturbation magnitude by electrode site interaction (F(14,208)=.854, p=.534), failed to reach significance. Polynomial contrasts indicated that N1 amplitudes were modulated by perturbation magnitude in a linear manner (F(1,22)=12.681, p=.002) (see Figure 3). Accordingly, pairwise comparisons indicated that the 40, 30, 25, and 20-cent perturbations elicited significantly larger (more negative) N1 amplitudes than the 0-cent (unaltered) condition, while the 40, 30, and 20-cent perturbations also elicited significantly larger (more negative) N1 amplitudes than the 5-cent perturbations.

The two-way RM-ANOVA investigating the influence of perturbation magnitude and electrode site on N1 latency failed to find a significant effect of perturbation magnitude (F(6,132)=1.671, p=.133), site (F(2,44)=1.011, p=.372), or an interaction between perturbation magnitude and site (F(12,264)=.785, p=.595).

P2 Amplitude and Latencies

A two-way RM-ANOVA was conducted to investigate the influence of perturbation magnitude and electrode site on P2 amplitudes. A significant main effect of perturbation magnitude (F(7,154)=4.894, p<.001) and electrode site (F(2,44)=8.445, p=.001) were found. However, the perturbation by electrode site interaction failed to reach significance (F(14,308)=1.509, p=.173). Polynomial contrasts indicate that P2 amplitudes were modulated linearly by perturbation magnitude (F(1,22)=13.114, p=.002). Pairwise comparisons indicated that with exception of the 10-cent perturbation, all perturbation magnitudes elicited significantly larger P2 amplitude than the 0-cent (unaltered) condition. In addition, the 40-cent perturbation elicited larger P2 amplitudes than the 5, 10, and 20-cent perturbations, while the 30-cent perturbation elicited larger P2 amplitudes than the 10-cent perturbations (see Figure 3). Pairwise comparisons also indicated that P2 amplitudes were largest at medial electrode sites, relative to left (p=.001) and right (p=.004) lateralized sites.

The two-way RM-ANOVA investigating the influence of perturbation magnitude and electrode site on P2 latency failed to find a significant effect of perturbation magnitude (F(6,132)=.218, p=.971), site (F(2,44)=1.150, p=.317), or an interaction between perturbation magnitude and site (F(12,264)=.851, p=.598).
Correlational Analyses

Correlational analyses were completed to determine if the proportion of feedback perturbations detected, or hit rate, was related to vocal response magnitudes, as well as P1-N1-P2 amplitudes. A significant correlation was found between number of hits and vocal response magnitude ($r=.224$, $p(2\text{-tailed})=.004$, $n=161$), P1 amplitude ($r=.184$, $p(2\text{-tailed})=.020$, $n=161$), and P2 amplitude ($r=.216$, $p(2\text{-tailed})=.006$, $n=161$), but not N1 amplitude ($r=-.142$, $p(2\text{-tailed})=.073$, $n=161$). In addition, a significant correlation was found between response magnitude and P2 amplitude ($r=.160$, $p(2\text{-tailed})=.043$, $n=161$).

DISCUSSION

In this study, speakers were exposed to small feedback perturbations and asked to report whether or not their feedback was altered. Vocal, behavioural, and ERP responses were recorded in attempt to estimate the magnitude of feedback alteration required to elicit a compensatory response, as well as the threshold at which participants reported the detection of feedback perturbations. The results showed that compensation to feedback perturbations occurred in response to perturbations as small as 10 cents, and that vocal response magnitudes increased linearly as the feedback perturbation magnitude increased from 0 to 40 cents. This ‘motor threshold’ of 10 cents is of similar magnitude to the threshold reported by Hafke (2008), who found that participants compensated to pitch shifts as small as 9 cents. On the other hand, the hit rate results indicated that the magnitude at which participants reliably reported detecting feedback perturbations was around 15 cents. In this study, 5 and 10 cent perturbations were reported less than 50% of the time (at or below chance), while feedback perturbations 15 cents and larger were detected a minimum of 69% of the time (well above chance). The threshold at which participants reliably reported feedback perturbations in this study is slightly lower than the 21 cent threshold reported by Hafke (2008). However, as this current study utilized multiple feedback perturbations per utterance, and Hafke (2008) utilized a single mid-utterance shift, it is plausible that the multiple perturbations per utterance facilitated detection of the altered feedback. Despite the difference in the magnitude of feedback perturbation required for participants to report a pitch change in these studies, both studies suggest that the magnitude of feedback perturbation required to elicit a compensatory response, and for an individual to report a pitch change, is different.

In addition to vocal and behavioural responses, ERPs elicited by the altered feedback were also measured. In accordance with previous studies, P1-N1-P2 amplitudes were all modulated by the feedback perturbations. All feedback perturbation magnitudes elicited larger P1 amplitudes than the unaltered feedback, while the 40 cent
feedback perturbation also elicited larger P1 amplitudes than the 15 cent perturbation. These results are consistent with the idea that the P1 component responds non-specifically to changes in an acoustic stimulus (Chait et al., 2004; Korzyukov et al., 2012; Scheerer et al., 2013). On the other hand, N1 amplitudes showed increased specificity, as N1 amplitudes were found to increase (become more negative) in response to large feedback perturbations (20, 30, and 40-cents) relative to small (5-cent) perturbations, and unaltered feedback. Scheerer and colleagues (2013) suggest that increases in N1 amplitude occur when sensory feedback is in violation of the prediction created by an efference copy. The current pattern of results suggests that small feedback perturbations (i.e., 5-cents) may be too small to exceed the normal variability in an individual’s voice, preventing a violation of the efference copy and resulting in similar N1 amplitudes as the unaltered condition. On the other hand, larger feedback perturbations (20, 30, and 40-cents) exceed the normal variability of an individual’s voice, creating a violation of the prediction set out by the efference copy, and resulting in increased N1 amplitudes. Similarly, P2 amplitudes were also found to increase in response to larger feedback perturbations. P2 amplitudes increased linearly with increasing perturbation magnitudes, which is in accordance with previous studies that have reported that P2 amplitudes increase in a graded manner with increasing feedback perturbation magnitudes (Behroozmand et al., 2009; Scheerer et al., 2013). In addition to modulation as a function of feedback perturbation magnitude, P2 amplitudes were also found to correlate significantly with vocal response magnitudes. This result has been reported previously by Scheerer and colleagues (2013) and suggests that P2 amplitudes may reflect the computation of the mismatch between perceived and expected sensory feedback.

The results of this study suggest that the P1 and P2 components, in addition to their modulation by perturbation magnitude, also varied as a function of electrode site. Both the P1 and P2 components displayed larger amplitudes at medial sites, relative to the left and right electrode sites. These results are consistent with previous studies that have reported larger P1 and P2 amplitudes at fronto-medial electrode sites (Behroozmand et al., 2009; Scheerer et al., 2013).

In addition to investigating ERP amplitudes, changes in P1-N1-P2 latencies were also investigated across the different perturbation magnitudes. For feedback perturbations between 5 and 40 cents, the latencies did not vary as a function of perturbation magnitude. Although previous studies have shown ERP latency modulation as a function of perturbation magnitude, these studies used larger, more discrete shifts (Liu, Meshman, Behroozmand, & Larson, 2011; Scheerer et al., 2013).

In attempt to uncover potential relationships between the threshold at which feedback perturbations are reported and vocal and ERP responses, correlational analyses were conducted. The proportion of detected feedback perturbations, or ‘hit rate’ was found to correlate with vocal response magnitudes, P1, and P2 amplitudes. Despite the fact that differences were found in the magnitude of perturbations required to elicit a compensatory response, and for individuals to report a feedback error, the significant correlation between hit rate and vocal response magnitude suggests that the reliability with which individuals can report errors in their auditory feedback is related to the size of their compensatory response to the error. In addition, the correlation between hit rate and P1 and P2 amplitudes provides further evidence that these ERP components index neural activity related to the detection and correction of unintended changes in vocal output.

Previous studies investigating the relationship between compensation to altered auditory feedback and speakers’ awareness of the alteration, have suggested that differences in the magnitude of feedback error required to elicit a motor and perceptual response support an action-perception dissociation. However, due to the inherent variability in individuals’ voices, reported differences in these motor and perceptual thresholds may in fact reflect differences in the criteria for triggering a compensatory response and a conscious report of the perceived feedback deviation. The results of the current study suggest that individuals begin compensating for feedback perturbations in the range of 5-10 cents. On the other hand, individuals are able to reliably report feedback perturbations between 10-15 cents. Although the feedback perturbation magnitude required to elicit a motor response and a conscious detection of an error differed slightly in this study, the negligibility of the difference provides little support for an action-perception dissociation in speech motor control. Rather, a more parsimonious explanation for the observed results would be that due to the inherent variability in the speech signal, the criteria for triggering a compensation response, or a conscious report of a vocal perturbation are different. Finally, we should point out that it is currently unknown whether participants are reporting their awareness of the altered auditory feedback based exclusively on the auditory stimuli. It is possible that, at least for very small perturbations, participants’ reports are based on their awareness of their motor responses and not what they hear in the auditory stimuli. It is our hope that further ERP investigations can elucidate these types of issues.
ACKNOWLEDGMENTS

This work was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES


