Conscious and unconscious contributions to auditory feedback control of speech — Neural correlates and behavioral consequences

Daniel Suchý[1,2], Roozbeh Behroozmand[3], Henry Railo[1,2]

[1] Department of Psychology and Speech-Language Pathology, University of Turku, Finland
[2] Turku Brain and Mind Centre, University of Turku, Finland
[3] Speech Neuroscience Lab, Department of Speech Language and Hearing, Callier Center for Communication Disorders, School of Behavioral and Brain Sciences, The University of Texas at Dallas, 2811 N. Floyd Rd, Richardson, TX 75080, United States

Abstract

During speaking, errors in the auditory feedback are processed rapidly and automatically, suggesting that feedback control of speech operates unconsciously. Yet, humans are also sensitive to consciously detect errors in their auditory feedback during speaking, but research has not probed whether conscious processing of feedback errors modulates speech, or what neural mechanisms underlie conscious detection of errors during speaking. To address these questions, we concurrently measured behavioral vocal responses and electroencephalography (EEG) in 30 participants while their online auditory feedback was altered using individually calibrated threshold-level pitch changes (i.e., errors). Following each vocalization trial, participants reported if they consciously detected pitch errors in their auditory feedback. Our results revealed that irrespective of conscious error detection, participants generated vocal compensation to the altered auditory feedback; however, conscious error detection was associated with increased magnitude of vocal responses roughly 500-700 ms after the altered feedback compared to the unconscious trials. In addition, we found that conscious detection of the altered feedback correlated with early and late neural activity responses as indexed by the modulation of event-related potentials (ERPs). Source localization of ERP responses suggested that conscious error detection was attributed to increased neural activity within the temporal, frontal and parietal cortical networks known to be involved in speech motor control. These findings emphasize the importance of investigating the role of higher-level conscious processing mechanisms in regulating speech feedback control, and their effect on the underlying neural and behavioral functions.

Keywords: Speech motor control; auditory feedback; sensorimotor integration; conscious perception; neural correlates of consciousness

Introduction

Speech is a complex task that requires coordination of cognitive, motor, and sensory programs. Even small changes in acoustical parameters of the sound can lead to meaningful changes in communication. Therefore, the process of speech production needs to be subjected to feedback monitoring that at the level of cortex involves frontal, temporal and

parietal areas (Tourville & Guenther, 2011). While speech feedback control is often assumed to operate unconsciously, in some situations such as during speech therapy or phonemic learning, speech feedback may need to be consciously accessed. However, the functional role and neural underpinnings of conscious speech feedback control remain open.

Current models of speech production (Houde & Nagarajan, 2011; Tourville & Guenther, 2011) do not distinguish between conscious and unconscious speech feedback control. Rather, the models are built around the idea that speech feedback control relies on an automatic comparison of how well the acoustic features of produced speech match the intended targets. Speech-related motor commands arise in the left-lateralized ventral premotor areas of the cortex, and are executed by the motor areas. In addition to sending the feedforward motor commands, axons from the premotor cortex are assumed to relay information about the expected sensory outcome to areas such the auditory cortex in the posterior superior temporal gyrus. This copy of the motor command is known as the efference copy (or corollary discharge). The consequence of the efference copy is that the sensory areas have an inhibited response to vocalizations that match the target of the motor program, but are excited by motor commands that result in vocalization errors (Behroozmand et al., 2022; Behroozmand & Larson, 2011; Chang et al., 2013; Eliades & Wang, 2008). In case of errors, the motor areas are recruited to initiate corrections to the motor commands (Burnett et al., 1997; Hain et al., 2000). We call the change in vocalization that follow errors in sensory feedback the *vocal response*.

Auditory feedback control of speech is measured behaviorally by “perturbing” auditory feedback in real-time with artificial vocalization errors. In a widely used version of the paradigm, participants unexpectedly hear a pitch-shifted version of their own voice during vocalization (Burnett et al., 1997; Hafke, 2008; Hain et al., 2000; Hawco et al., 2009; Villacorta et al., 2007). If the participants are not given instructions on how to respond to the perturbation, their vocal responses are on average compensatory. That is, if the perturbation increases the pitch of auditory feedback, participants typically react by lowering their voice. Given that the onset of the vocal response is rapid (just over 100 ms), it is assumed to be mediated by unconscious processes. However, as discussed next, previous studies that have

tested the unconscious nature of vocal responding have significant methodological limitations.

If speech feedback control is mediated by unconscious processes, participants should reveal vocal responses even when they fail to consciously notice the perturbation. Supporting evidence for this comes from a study suggesting that trained singers adjust their voice in response to the perturbation unconsciously, but their analysis also included trials where the participants reported hearing the perturbation (Hafke, 2008). Another study replicated the finding of Hafke (2008) on a sample of participants without professional vocal experience, but only assessed conscious perception of the perturbation by having the participants passively listen to recorded samples of perturbed or non-perturbed feedback (Xu et al., 2020). This is problematic because participants are actually more sensitive to detect perturbations during "active" vocalization than during "passive" listening (Scheerer & Jones, 2018). In fact, Scheerer and Jones (2018) found that the differences between the thresholds for vocal responses and for conscious perception are at best minimal (5 cents), and suggested that the supposedly unconscious vocal responses to very small perturbations are instead biased by other factors such as vocal variability (i.e., participants may attribute the artificial error to the natural variability of their own voice). In this context, the fact that participants can consciously notice significantly smaller perturbations during vocalization than passive listening (Scheerer & Jones, 2018) begs the question about the functional role of consciousness in speech feedback control.

In addition to the rapid vocal response component (called VR1), the vocal response often also contains a second, later component (VR2, typical onset after 300 ms). Asking participants to voluntarily control how they respond to a perturbation predominantly modifies the VR2 (Hain et al., 2000). For this reason, the VR2 is often assumed to correspond to a conscious vocal response mode. In the present study we aimed to test the behavioral function of conscious perception on vocal responses. To isolate the effect of conscious perception, one needs to compare trials that differ only with respect to consciousness while keeping physical stimulation (i.e., the perturbation) constant for each participant. This type of contrastive approach is also used in the search for the neural correlates of conscious perception (Crick & Koch, 1990).

Conscious perception of external sensory stimuli correlates with activity in sensory cortical areas, and frontal and parietal areas(Dembski et al., 2021; Koch et al., 2016; Mashour et al., 2020), although the role of extrasensory areas remains debated (Boly et al., 2017; Odegaard et al., 2017). Event-related potential (ERP) studies show that conscious perception of external auditory stimuli correlates with activity in two time-windows. An earlier correlate known as the Auditory Awareness Negativity (AAN) has a negative polarity, and takes place around 100-250 ms after stimulus onset (Dembski et al., 2021; Eklund et al., 2021; Eklund & Wiens, 2019; Filimonov et al., 2022). This neural correlate is assumed to reflect the subjective experience of the sensory stimulus. A later correlate known as the Late Positivity (LP)—which coincides with the P3b event-related potential (ERP) wave—is often argued to reflect higher-order conscious perception (Derda et al., 2019; Jimenez et al., 2021) or post-sensory processing (e.g., decision making; (Dembski et al., 2021; Schlossmacher et al., 2020)). Other sensory modalities show similar correlates (Dembski et al., 2021).

Research on the neural correlates of consciousness has focused on investigating how individuals consciously perceive external, environmental stimuli, but how individuals become conscious of the sensory consequences of their own behavior remains little studied. This is a significant gap in research given that motor-related activity strongly modulates the processing self-initiated sensory stimulation (Schneider & Mooney, 2018) and the perception of self-produced sounds differs from the perception of external sounds (Lane et al., 1961). Are the neural processes that mediate conscious perception of one's own voice different from the processes that enable conscious perception of external sounds? What role does conscious perception of one's own voice play in speech motor control? Answers to these questions are important from the perspective of basic science, but could also turn out significant for our understanding of clinical speech deficits and their rehabilitation.

To address this gap in research, we examine the neural and behavioral correlates of consciously detecting errors in one's own speech (Fig. 1). Using a novel approach, we calibrate pitch perturbations for individual participants using a psychophysical staircase procedure. To test three preregistered hypotheses in the subsequent EEG experiment, we contrast trials where participants reported conscious perception of the perturbation with trials where they

failed to detect the perturbation. First, based on earlier research (Hafke, 2008; Xu et al., 2020), we hypothesize that vocal responses are initiated even when participants report not noticing the perturbation. In addition to this unconscious mode of feedback control, we hypothesize that speech feedback control can also operate at a conscious level, enabling additional corrections to speech. Second, based on studies examining the neural correlates of conscious perception of external sounds (Dembski et al., 2021; Eklund et al., 2021; Eklund & Wiens, 2019; Filimonov et al., 2022), we expected that conscious perception of the perturbation correlates with an early negative amplitude difference in central electrode locations (i.e., AAN). Finally, assuming that the ERPs in part reflect the initiation of vocal responses (e.g., Behroozmand et al., 2022), we expected to observe a correlation between ERP and vocal responses amplitudes.

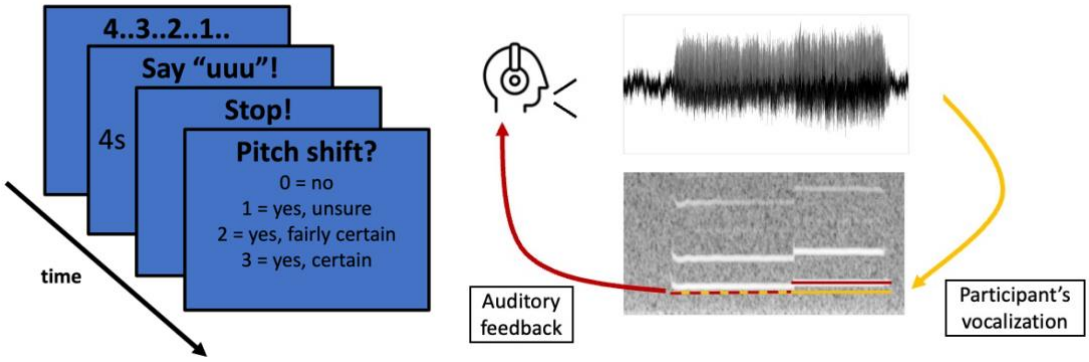

Figure 1: Graphical illustration of the experiment task. The participants heard their online auditory feedback through headphones while it was sometimes perturbed 1.5 seconds after the start of vocalization (upward pitch shift) and remained on until the end of the 4-second-long vocalization phase. In the panel on right, the yellow line denotes the true pitch of vocalization (i.e., unmodified auditory feedback), and the red line shows the perturbed (i.e., pitch-shifted) auditory feedback. After each vocalization, the participant reported if they detected the perturbation (on right).

Results

Behavioral results

Results showed that the participants greatly differed in their ability to consciously detect the perturbation. The mean perturbation magnitude determined by the staircase procedure was 23 cents (min 1 cent, max 71 cents). The mean proportion of critical trials (i.e., trials with a perturbation) reported as aware during the experiment was 70%, but performance varied considerably between participants (min 16%, max 95%). The overall result is in line with that

of Hafke (2008), who showed that perturbation magnitudes of 19 cents are, on average, below participants' 75% threshold of conscious perception.

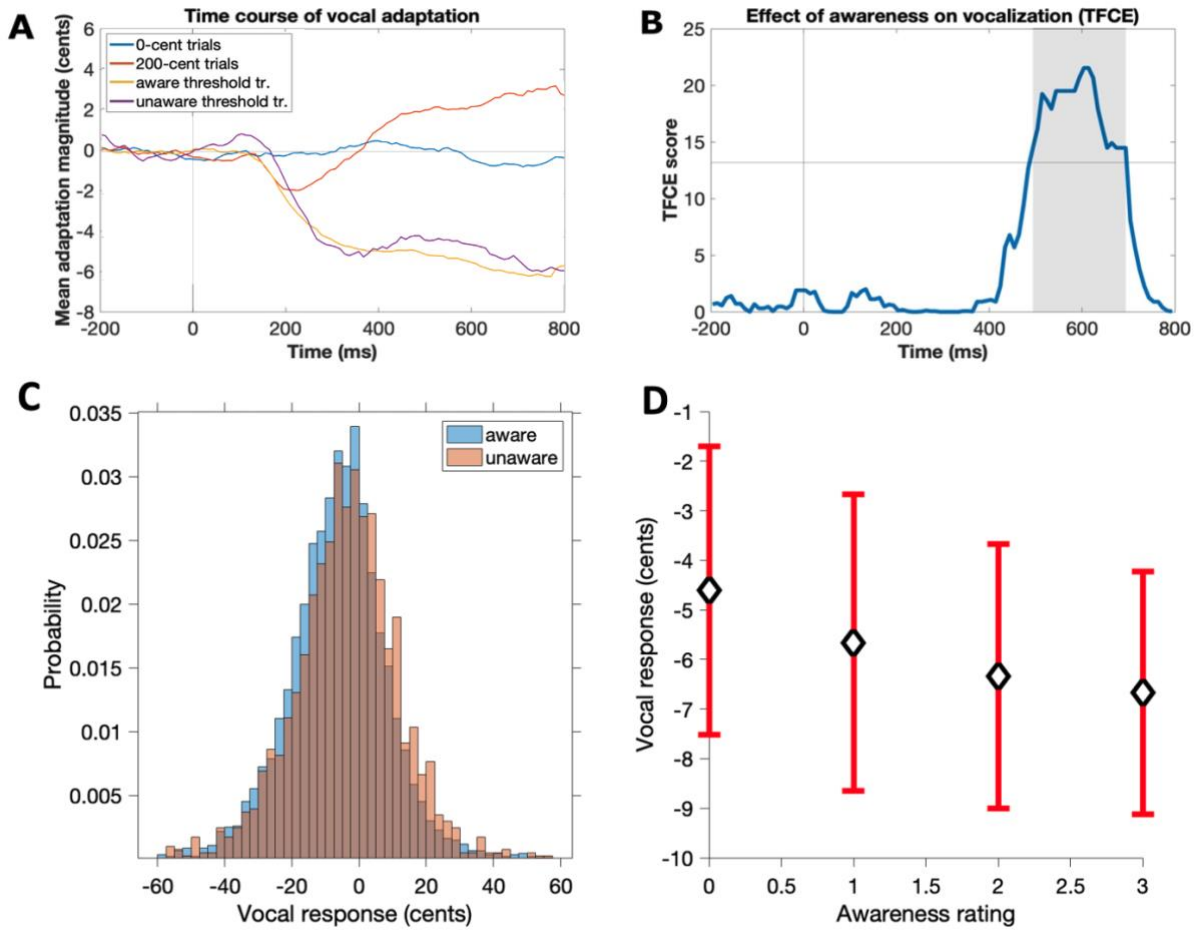

Figure 2: Vocal response results. (A) Mean vocal response. On the x axis, zero represents perturbation onset. On the Y axis, zero represents 200 ms pre-perturbation mean pitch. Consciousness ratings are here treated as binary for the ease of viewing (unaware = rating 0). (B) Results of the mass univariate analysis. Highlighted area represents the time interval with a statistically significant effect of consciousness, meaning that significant t-values cluster around this point. Horizontal line represents the statistical significance threshold (p < .05, corrected for multiple comparisons). (C) Histogram of the vocal responses 494 to 694 ms after perturbation onset separately for unaware (rating 0), and aware (ratings 1-3) trials. (D) Mean vocal response magnitude 494 to 694 ms after perturbation onset as a function of consciousness rating (error bars represent the SEM).

Vocal responses

Mean time course of vocalization during each type of trial used in the experiment is visualized in Figure 2A. Figure 2A shows that participants also performed vocal responses to threshold-level stimulus, both when they reported being conscious (ratings 1-3; yellow line) or

unconscious (rating = 0; purple line) of it. Compared to the biphasic vocal response in high perturbation trials (red line; 200 cent perturbation), threshold level perturbations produced a longer duration vocal response. We quantified the statistical significance ($p < .05$, corrected for multiple comparisons) of the effect of conscious perception using mass univariate analysis whose result is visualized in Figure 2B. Permutation testing indicated that conscious perception of the perturbation increased vocal response magnitude in the time interval from 494 ms until 694 ms after perturbation onset (gray shaded area in Fig. 2B). As shown in Figure 2C, conscious perception equally modulated both compensatory (i.e., negative) and "following" (i.e., positive) vocal responses. Vocal response magnitude increased as a function of conscious perception rating in the 494-694 ms time-window (Fig. 2D).

ERPs

Graphical summary of the ERP results is shown in Figure 3. Trials rated as unconscious produced an oscillatory response without clear ERP peaks, while conscious trials show early negative (100–200 ms) and late positive (300–500 ms) ERP waves (Fig. 3A). We performed linear mixed effects regression tests in order to measure the effect of conscious perception on ERPs produced by the perturbation (**Error! Reference source not found.**). These analyses focused on two time-windows (AAN and LP) hypothesized to reflect conscious perception. The effect of conscious perception was significant for both the early component and the late time-window, reported in Table 1 (the intercept is the amplitude in unconscious trials, and the Consciousness Estimate indicates how ERP amplitude changes as rating increases one unit). As shown in Fig. 3B and Fig. 3C, the amplitude of the early and late responses increases as a function of consciousness rating, suggesting that these correspond to AAN and LP. The scalp distributions (Fig. 3D and Fig. 3E) of the early and late time window correlates are consistent with what has been previously reported for AAN and LP, although both time windows suggest somewhat a right-lateralized activation.

Vocal–ERP correlation

We tested for correlations between ERP and vocal response magnitudes. We selected the early and late ERP correlates as presented above, and tested if they correlate with vocal response amplitudes from (a) the same time interval, or (b) the time interval with a significant effect of consciousness (i.e., 464-694 ms), obtained from the mass univariate analysis of vocal

responses (Figure 2). We used linear mixed effects regression to compare these variables on a single-trial basis (pitch ~ erp_amplitude + (1|participant)). None of these 4 tests showed a near significant effect of ERP magnitude on vocal response magnitude.

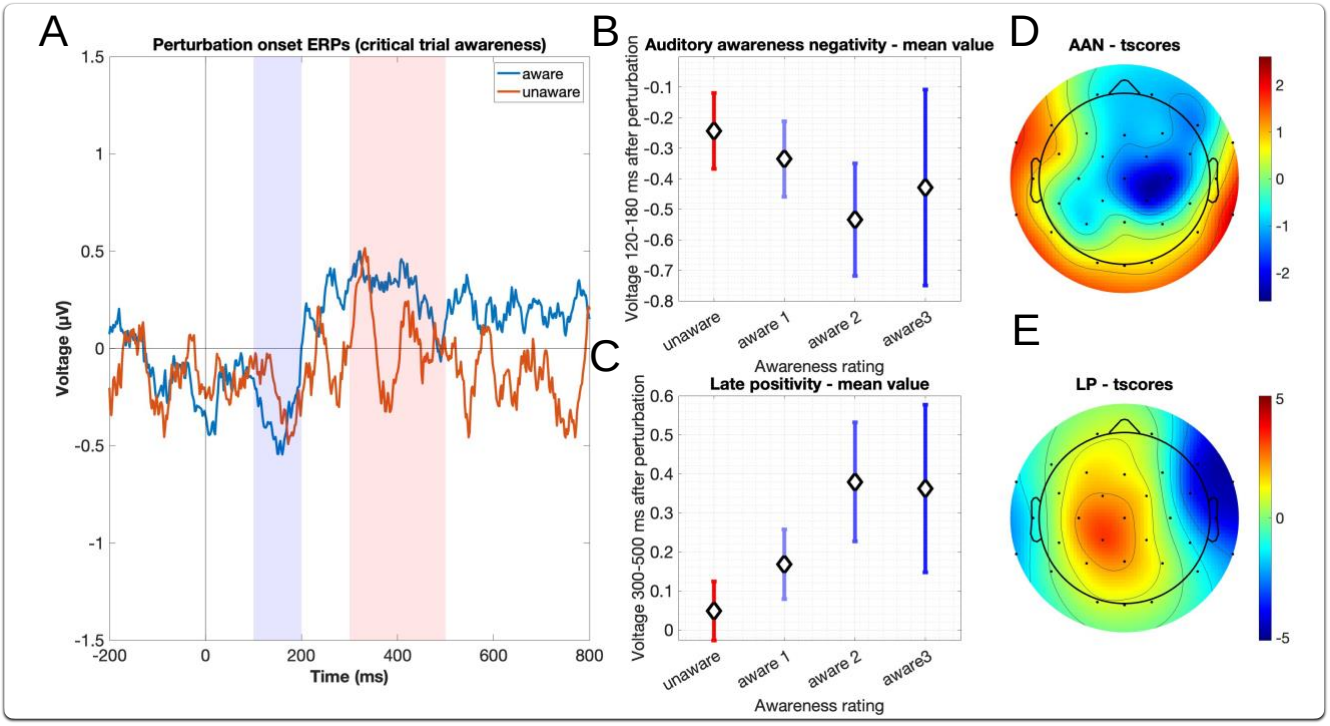

Figure 3: ERP results. (A) Mean ERP responses for trials rated as conscious and unconscious. Perturbation onset is at 0 ms. The consciousness rating is binarized for easier presentation (unaware = rating 0). (B) AAN amplitude as a function of consciousness. Error bars represent standard error of mean. (C) LP amplitude as a function of consciousness. Error bars represent standard error of mean. (D) Scalp distribution of AAN (map of t-values). (E) Scalp distribution of LP (map of t-values).

Table 1: Results of linear mixed effects regression on the effects of consciousness of early and late ERPs.					
Auditory Awareness Negativity, AAN (100-200 ms)					
Parameter	*Estimate*	*Lower 95% CI*	*Upper 95% CI*	*t*	*p*
Intercept	-0.19	-0.30	-0.08	-3.36	0.0008
Consciousness	-0.10	-0.17	-0.03	-2.98	0.0029
Late Positivity, LP (300-500 ms)					
Intercept	0.04	-0.16	0.25	0.39	0.69
Consciousness	0.18	0.11	0.25	4.96	< 0.0001

Source estimation

Figure 3 visualizes sources that were statistically significantly modulated by conscious perception of the perturbation during AAN (top panels), and LP (bottom panel) time-windows. Table 2 and Table 3 lists the MNI coordinates of the sources maximally modulated by conscious processing of vocal feedback pitch changes in the AAN, and LP time-windows, respectively. If multiple significant sources exist within a single structure, only the source with the highest effect of consciousness is presented in Tables 2 and 3.

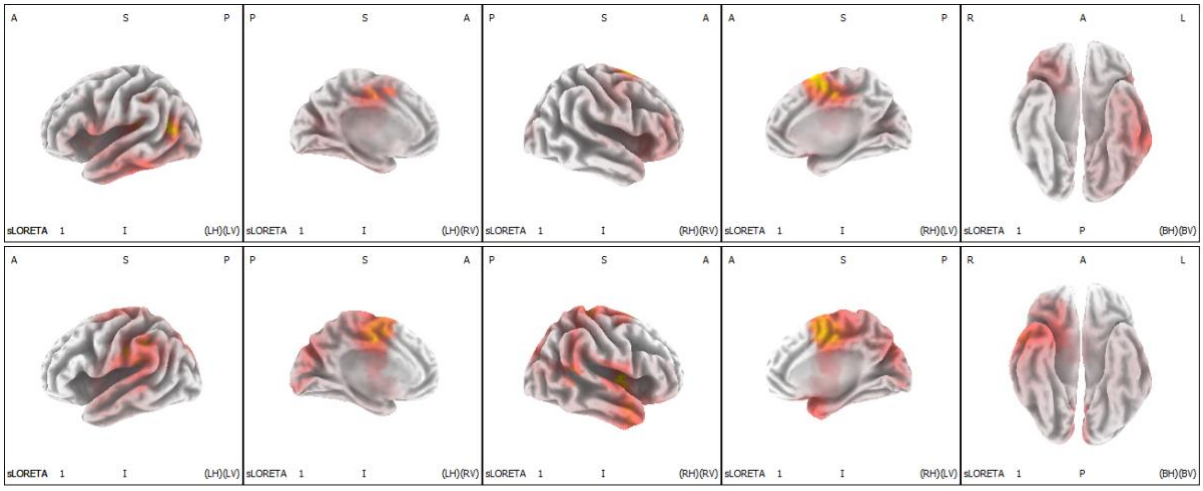

Figure 3: The estimated current source density maps of statistically significant ($p < .05$) modulation by consciousness *(t test results, corrected for multiple comparisons). First row: AAN (100-200 ms post perturbation onset), second row: LP (300-500 ms post perturbation onset).*

During the AAN time window, activity was modulated in distributed areas involving temporal, frontal, parietal, and even occipital cortex. In the temporal cortex, the effects of consciousness were strongest in the posterior part of left superior temporal gyrus, and left inferior temporal cortex. Inferior frontal cortex and insula were activated bilaterally. Activations were also present in areas corresponding to motor and somatosensory areas, and strong activity was present especially in bilateral paracentral lobule and middle part of the cingulate cortex.

In the LP time window, bilateral paracentral lobule and middle cingulate cortex again showed strong modulation by conscious perception. Activity in the superior temporal gyrus

was also modulated by conscious perception bilaterally. Conscious perception related activity is also present near the boundary of temporal and parietal cortex (this effect is somewhat left lateralized), in addition to more posterior parietal/occipital effects. Unlike in the AAN time window, the LP time-window showed considerably weaker effects in the frontal lobe.

Table 2. MNI coordinates of the maximum source activations in the AAN time window					
X	Y	Z	t-value	BA	Structure
-40	-55	20	-3,92	22	Superior Temporal Gyrus
10	10	70	-3,91	6	Superior Frontal Gyrus
10	0	60	-3,87	6	Medial Frontal Gyrus
15	10	65	-3,79	6	Middle Frontal Gyrus
-55	-65	15	-3,76	19	Middle Temporal Gyrus
-35	-40	20	-3,66	13	Insula
10	-10	50	-3,61	24	Cingulate Gyrus
5	-15	50	-3,55	31	Paracentral Lobule
20	0	60	-3,54	6	Sub-Gyral
-40	-30	30	-3,46	2	Postcentral Gyrus
-45	-30	30	-3,39	40	Inferior Parietal Lobule
-60	-35	-25	-3,23	20	Inferior Temporal Gyrus
-40	-45	35	-3,18	40	Supramarginal Gyrus
-55	10	10	-3,16	44	Precentral Gyrus
-55	-35	-25	-3,13	20	Fusiform Gyrus
-55	-70	5	-3,12	19	Middle Occipital Gyrus
-35	-35	10	-3,08	41	Transverse Temporal Gyrus
25	30	-5	-3,07	47	Inferior Frontal Gyrus
35	20	0	-2,93	47	Extra-Nuclear
-20	-45	30	-2,88	31	Precuneus
-40	-30	-15	-2,82	36	Parahippocampal Gyrus
-20	-65	15	-2,77	31	Posterior Cingulate
20	15	-15	-2,71	47	Subcallosal Gyrus
-15	-70	15	-2,67	18	Cuneus
20	30	-25	-2,65	47	Orbital Gyrus
-25	0	-45	-2,64	20	Uncus
-50	-70	30	-2,61	39	Angular Gyrus

Table 3. MNI coordinates of the maximum source activations in the LP time window					
X	Y	Z	t-value	BA	Structure
40	0	10	3,39	13	Insula

5	-5	55	3,25	6	Medial Frontal Gyrus
5	-5	50	3,24	24	Cingulate Gyrus
-45	-30	30	3,23	40	Inferior Parietal Lobule
5	5	60	3,20	6	Superior Frontal Gyrus
50	-5	10	3,19	6	Precentral Gyrus
-50	-25	30	3,17	2	Postcentral Gyrus
5	-15	50	3,08	31	Paracentral Lobule
65	-45	15	3,03	22	Superior Temporal Gyrus

Discussion

Using a novel experimental approach that allowed us to contrast instances where participants consciously perceive errors in their auditory feedback with trials where the same error goes unnoticed, we characterized the neural and behavioral correlates of conscious perception of errors in one's own voice. Our results showed that short latency vocal responses to pitch perturbations were elicited even when participants failed to consciously notice the error, but conscious perception of the error in auditory feedback correlated with stronger vocal responses at late time windows (around 500-700 ms after the error). That consciousness modulated vocal responses suggests that models of speech feedback control need to explain how unconscious and conscious modes of speech feedback control differ in terms of their function, and neural correlates. Our results indicate that conscious perception of an error in one's own vocalization correlated with ERPs in early (AAN) and late (LP) time windows. Source reconstruction results suggest that during both these time windows, activity in distributed cortical areas correlated with conscious perception.

Effect of consciousness on vocal responses

The present results confirm the hypothesis that vocal responses to pitch perturbation are mediated by unconscious processes (Hafke, 2008; Xu et al., 2020). Because consciousness correlated with vocal responses only in relatively late time windows, our result is consistent with the proposal that the early vocal response component (VR1) is determined by unconscious processes, whereas the later component (VR2) can be modulated by consciousness (Hain et al., 2000). In the present study, consciousness was associated with relatively minor chances in the vocal response magnitude. This is understandable, given that

we did not instruct participants to modulate their vocalization. In our results, clearly consciously perceived perturbations (i.e., rating 3) were associated with almost 2 cents larger vocal responses than unconscious perturbation (i.e., rating 0; Fig. 2D), corresponding to over 10% increase in vocal response magnitude relative to (average) perturbation magnitude. The effect of conscious perception on vocal responses was uniform, shifting the whole distribution (Fig. 2C), indicating that "opposing" and "following" responses (see, (Franken et al., 2018)) were similarly affected.

That consciousness only modulated late vocal responses aligns with current models of conscious perception. The early, unconscious part of vocal response is assumed to be mediated by feedforward circuitry: mismatch between intended and produced vocalization activate error units in auditory cortex, which communicate the error to frontal cortex to rapidly initiate corrective movements. Theories of consciousness assume that feedforward neural activity is not associated with conscious perception, although such activity can nevertheless trigger behavioral responses (Hurme et al., 2017; Lamme & Roelfsema, 2000). Consequently, consciousness likely only correlated with late vocal responses because conscious perception required the activation of feedback connections to trigger sustained, synchronized brain activity.

Neural correlates of consciously perceiving errors in auditory feedback

The ERP observed in the conscious condition in the present study resembles previously reported ERPs to pitch perturbations (Behroozmand & Larson, 2011; Scheerer & Jones, 2018). The ERP correlates of conscious perception in the present study reveal a similar pattern of correlates as conscious perception of purely "external" auditory stimuli (i.e., stimuli not-related to the individual's own behavior), although, unlike in previous studies with "external" stimuli, the scalp distributions of AAN and LP are somewhat right lateralized (Dembski et al., 2021; Eklund et al., 2021; Eklund & Wiens, 2019; Filimonov et al., 2022). The primary sources of AAN have been in previous studies (that use external stimuli) localized to posterior superior temporal cortex (Dykstra et al., 2016; Eklund et al., 2019; Gärtner & Gutschalk, 2021; Wiegand & Gutschalk, 2012), although some results also report activity in frontal cortex (Dykstra et al., 2016; Eklund et al., 2019). Intriguingly, the present source

localization results suggest that activity in distributed cortical areas contributed to AAN, in addition to sources in the left auditory cortex. Frontal activations were largely lacking in LP time window in the present study. This contrasts the typical interpretation that LP underlies "accessing" conscious contents for task-relevant processing through mechanisms that involve the frontal cortex (Pitts et al., 2012; Schlossmacheret al., 2021).

While the present source localization is limited in resolution (due to the small number of electrodes whose locations were not digitized, and the use of template brain anatomy; (Song et al., 2015)), the observed sources are encouragingly consistent with the neural basis of speech production (Behroozmand et al., 2022; Tourville et al., 2008; Toyomura et al., 2007). Next, we first discuss the source activations in light of neural correlates of consciousness. After this, in the next chapter, we discuss how the observed sources activations may relate to neural mechanisms of speech motor control.

The widespread nature of the sources correlating with conscious perception of the perturbation in the AAN time window raises the possibility that conscious perception of errors in one's own vocalization is mediated by activity in frontal and parietal areas. Theoretically, this is a plausible suggestion. Conscious perception of errors in one's own voice tells individuals that their motor behavior may need to be corrected. Given that the sources in AAN time-window strongly overlapped with areas known to be involved in speech feedback control (Guenther, 2016), the present results suggest that that the same circuit that underlies speech feedback control may also mediates conscious perception of the errors in vocalization. Put in terms frequently used in consciousness research, the present results suggest that participants may be able to consciously "access" information processed by the speech feedback control circuitry.

Based on theories of consciousness, conscious perception of error in vocalization could require that information about the detected error is integrated into a singular network with its own cause-effect power (often correlating with parietal activity; Tononi et al., 2016), that sensory activation "ignites" strong, sustained activation is distributed areas (Mashour et al., 2020), that higher-order frontal mechanisms take the auditory error information as content (Brown et al., 2019), or that auditory cortex activates feedback connections to form

recurrently communicating neural coalitions (Lamme, 2010; Lamme & Roelfsema, 2000). The present results are largely consistent with these theories, and future research is required to contrast different theories.

In addition to reflecting processes that truly underlie conscious perception of the vocalization error, the observed correlates of consciousness could also reflect processes that do not directly produce conscious perception (Aru et al., 2012). For example, the correlates observed in the present study likely include neural correlates of attention to one's vocalization as the participants closely attended to their voice to detect the perturbation. Attention can modulate ERP amplitudes in same time windows and electrodes as conscious perception, and the two processes interact, but can also be dissociated from each other (Koivisto et al., 2005, 2006; Pitts et al., 2018).

In the present study, detecting the perturbations required participants also to separate the artificial pitch shift from normal variation in the pitch of their vocalization. This type of conscious monitoring of the source of the error is known to modulate vocal response amplitudes (Franken et al., 2023; Subramaniam et al., 2018). The neural correlates of source monitoring of vocalization remain unknown, but based on studies in other modalities, it may be mediated by efference copy circuits similar to speech feedback control (David et al., 2008), and activate frontal and parietal regions such as the insula and temporoparietal junction (Kühn et al., 2013; Zito et al., 2020).

Also, methodological limitations complicate the interpretation of the results. When participants hear pitch shifted feedback through headphones, they may concurrently detect so called "beat tones" (vibration caused by the two simultaneous tones that differ slightly in frequency) if they simultaneously hear the non-perturbed voice in the background. We used sound-isolating in-ear earphones and embedded auditory feedback in background noise to eliminate the beat-tones. However, it is difficult to verify that we were able to eliminate beat tones completely.

The role of consciousness in speech feedback control

As mentioned, the sources that correlated with conscious perception overlap with the circuitry of speech feedback control. Vocalization errors are assumed to be detected in higher-order auditory cortex in posterior superior temporal gyrus (pSTG; which receive efferent copies from the motor cortex) (Hickok, 2022). Consistent with this, in the present study, left pSTG was activated during AAN and LP time windows. In the DIVA model, auditory errors are assumed to be translated to motor coordinates in the right inferior frontal gyrus (rIFG) (Tourville et al., 2008; Toyomura et al., 2007). Activity in this area correlated with consciousness in the AAN time-window. Activity in the left IFG also correlated with consciousness during AAN. This area is an integral part of all models of speech production (Guenther, 2016), and it has been suggested to contribute to speech feedback control by regulating activity in sensorimotor areas (Behroozmand et al., 2022). Activity in the left IFG has also previously been reported to correlate with AAN in (Dykstra et al., 2016). Also consistent with present results, brain imaging results suggest that right insular cortex activity correlates with errors in speech feedback (Toyomura et al., 2007). Bilateral paracentral activity (i.e., supplementary motor area), and the adjacent Cingular cortex correlated strongly with consciousness in both AAN and LP time windows. Both of these areas are activated during speaking (Guenther, 2016; Picard & Strick, 1996), and their stimulation or lesions results in speech-arrest or mutism (Filevich et al., 2012; Guenther, 2016). In the LP time-window, the area near the boundary between parietal and temporal areas correlated with consciousness. This corresponds to area referred to as spt in the dual-stream model, assumed to underlie sensory-to-motor transformations (Hickok, 2022). Activation in nearby areas during LP also includes the angular gyri, which have been reported to correlate with processing altered auditory feedback (Behroozmand et al., 2022).

What is the relationship between the observed correlates of conscious perception of the perturbation and vocal control? There are at least two not mutually exclusive explanations. First, the observed correlates of consciousness may have causally modulated vocal responses. Given that both AAN and LP preceded the time-window when vocalization was modulated by consciousness, this conclusion is in line with the observed data. According to the standard interpretation, conscious control of behavior is mediated by processes taking place in the LP time window, although some evidence suggests that already early correlates of consciousness

could mediate behavioral changes (Railo et al., 2015). Given that source reconstruction suggested that activity in frontal and parietal areas correlated with conscious perception in AAN time window, our findings are consistent with the interpretation that these processes are part of the neural mechanisms that initiated corrections to vocalization when the error was consciously noticed. However, neither AAN, or LP amplitude predicted vocal response magnitudes, suggesting that the two ERP correlates of consciousness may not causally drive vocal responses.

The alternative explanation is that conscious perception of the perturbation (and therefore its neural correlates) did not causally influence the vocal responses. Based on this interpretation, a "third factor" explains the observed association between conscious perception and vocal responses. It could be, for example, argued that participants consciously detected the perturbations only on trials where they succeeded to selectively attend to the perturbation—and the same attentional selection possibly mediated amplified vocal responses. This interpretation is consistent with the fact that attention can modulate vocal responses (Liu et al., 2015).

Conclusion

Out results show that speech feedback control is modulated by conscious perception. This suggests that feedback control of speech may function in two different modes: A rapid unconscious mode, and a slower conscious mode that builds on the unconsciously determined vocal response. Activity in distributed cortical areas, including temporal, frontal and parietal cortices correlated with conscious perception of errors in vocalization, overlapping with circuits that underlie speech feedback motor control. Lack of understanding about how humans become conscious of their own voice, and how this enables them to voluntarily modulate their speech represents a significant gap in our understanding of the mechanisms of speech. Our results represent the first steps towards understanding this process, but significant questions remain open: What are the functional contributions of different cortical areas to conscious perception of errors in one's own vocalization? Through what mechanism does conscious perception modulate vocalization? Could biases in how one consciously perceives one's own voice lead to clinically meaningful speech deficits?

Methods

Pre-registration and data availability

The present study was preregistered prior to collecting data and running analyses (https://osf.io/am3rq/). All procedures described in this paper conform to the pre-registration, unless otherwise specified. All computer code used to run the experiment and analyze the results are publicly available at osf.io.

Participants

35 participants (M_{age} = 26, range: 20-36; 27 female) participated the study. All participants gave written informed consent. Sample size was determined based on previous studies, as described in the preregistration document. The participants were recruited using the university's participant recruitment website, by advertising among psychology, speech pathology, and biomedical students, and by word-of-mouth. The participants completed the experiment in exchange for course credit, or small financial reward. The exclusion criteria for participating in the study included left-handedness, and presence of neurological, hearing, or voice disorders that would interfere with completing the experiment. We excluded five participants from the analyses. Three of them did not complete the tasks as intended (they reported detecting the perturbation on every trial, systematically varied the pitch of their own voice regardless of trial type, or systematically reported not detecting large control trial perturbations), and two participants were excluded due to insufficient EEG quality. The research was approved by the ethics board of the Wellbeing services of the Southwest Finland.

Procedure

The experimental procedure was completed in a single session, lasting up to two hours and 30 minutes. The perturbations were created using the Audapter package (Cai et al., 2008) and the experiment was run using the Psychtoolbox package (Brainard, 1997), both running in Matlab (version 2016B). Participants' vocalization was recorded using a Motu MicroBook IIc audio interface, and a microphone (Audio-technica AT2035) positioned at a 1-2 centimeters distance from the mouth. The participants heard their own voice through sound-isolating in-

ear earphones. During each trial of all stages of the experiment, the auditory feedback was mixed with pink noise in order to reduce the participant's ability to hear their own voice due to insufficient headphone isolation, and bone conduction. We kept the sound level the same for all participants, except for a single participant, who considered the feedback to be too loud.

At the beginning of the experiment, we first presented the participant with an interactive warm-up task, with the goal of exposing them to both unperturbed and highly perturbed feedback (+ 200 cents, present until the end of vocalization). We instructed the participant to produce the vowel /u/ of sufficient length and loudness. The participant was not instructed about the pitch of the vocalization, and we did not instruct the participant on how to respond to the perturbation, other than to keep vocalizing until the preset time limit. During the warmup, the participants also practiced how to respond to the consciousness question that followed each trial. The warm-up continued until the desired quality of vocalization had been achieved and the participant was able to consistently identify the perturbed feedback.

After the warm-up, we presented the participant with a staircase calibration procedure, with a goal of estimating their perturbation perception threshold. The perturbation perception threshold represents a magnitude of pitch change (e.g., +20 cents) that the participant can consciously detect in a predefined proportion of trials (e.g., in 65 percent of the total trial count). We used the QUEST staircase algorithm (Watson & Pelli, 1983). The participant produced the /u/ phoneme for approximately 4 seconds. During the first 1.5 seconds of every trial, the participant was exposed to the unperturbed headphone feedback. Auditory feedback was perturbed 1.5 seconds after the beginning of vocalization, and the perturbation continued until the end of the trial. After each trial, the participant was asked to report on a 4-step scale whether they heard the change in the pitch (see below for details). The QUEST algorithm then placed the next perturbation at the value most likely to be near the specified threshold. A positive response (ratings 1-3, indicating a detection perturbation) lead to a decreased perturbation magnitude being used in the next trial of the experiment, while a negative response (lowest rating) lead to an increase of perturbation magnitude. Finally, we calculated the mean of the probability density function (as recommended by (King-Smith et

al., 1994)), and used this value as the individual perturbation perception threshold in the main phase of the experiment.

The calibration procedure had a fixed length of 50 trials, but it was re-run in case the results suggested an unreliable threshold estimate (e.g., unexpectedly high or low threshold value). All perturbations presented during the calibration as well as during the other phases of the experiment were upward, and the perturbation lasted until the end of vocalization. We initially estimated 75% perturbation detection thresholds, as it was considered an adequate trade-off between the amount of acquired unaware trials while maximizing perturbation size (to obtain clear ERPs). However, the 75% threshold resulted in a higher-than-expected amount of consciously perceived trials in the main experiment for the first participants, and therefore we lowered the threshold for later participants. In the participants included in the statistical analysis, the estimated thresholds were as follows: 75% threshold in 2 participants, 50% threshold in one participant, and 60% threshold in the remaining 27 participants. While this change is at odds with the pre-registration, it enabled us to maximize the number of aware and unaware trials per participant.

The main experiment, summarized in **Error! Reference source not found.**, took place after the calibration session. We recorded the EEG signal only during this phase. This experiment contained 300 trials per participant, divided into blocks of 50 trials. Each block was followed by a short break. The participants were presented with perturbations for which the intensity was adjusted to the individual conscious pitch error detection threshold. Control trials without perturbations, and control trials with large perturbations (+200 cents) were presented as well. The control trials amounted for ~33% of all trials, and were equally divided into trials with no perturbations and trials with strong perturbations. These trials were randomly distributed across the whole session (i.e., different blocks did not necessarily contain the same distribution of different trial types). The participant was instructed that they will be able to hear their unmodified voice at the beginning of each trial, and that this voice may then be modified as they continue vocalizing. After each trial, the participant was asked if they heard that the pitch of their vocalization was modified. The participant provided a response on a scale of 0 to 3, with higher numbers indicating clearer consciousness of the perturbation. The participants were told that the question concerns the auditory signal they hear from the

headphones (e.g., not how they thought they vocalized), and that the numbers on the scale corresponded to these statements: 0) No, I did not hear any change in pitch; 1) I might have heard a change in pitch, but I am unsure: 2) I am fairly certain that I heard a change; and 3) I am very confident that I heard a change.

The difference between the options (0) and (1), in particular, was stressed to the participants. In the analysis the lowest alternative corresponds to the "unconscious" condition. This question measuring conscious perception of a stimulus is based on a widely-used scale known as the Perceptual Awareness Scale (Ramsøy & Overgaard, 2004). Use of graded rating scale is important because previous studies in the visual modality suggest that the use of binary ratings scales may contaminate the "unconscious" category with stimuli the participants faintly consciously perceived, leading to confounded results regarding unconscious processing (Koivisto et al., 2021; Overgaard et al., 2008; Phillips, 2021).

Voice data processing and analysis

Participants' vocalization was processed in Matlab version R2022b with the perturbation onset data obtained from the voice-modulation software Audapter (Cai et al., 2008). During the preprocessing of vocalization data, we first lowpass filtered the data at 500 Hz and extracted the fundamental frequency (F0) from each vocalization using Matlab's *pitch* function. We removed the last 300 ms from the end of each trial as these contained unstable F0s. We removed trials which did not contain any vocalization, or contained vocalization that was too short, or too quiet to activate Audapter's perturbation algorithm. Next, we determined the start time of the perturbation relative to the start of each trial, and excluded trials which did not contain at least 800 ms of perturbed phonation. We used the 200 ms before the perturbation onset as a baseline period. For every sample from 200 ms before perturbation onset until 800 ms after perturbation onset, we calculated the relative deviation of this sample from the baseline period, using the standard equation used in auditory perturbation research (Hafke, 2008; Hain et al., 2000):

$$cents = 100 * \left(\frac{10 * Log\left(\frac{F_S}{F_B}\right)}{log(2)} \right)$$

where F_S is the fundamental frequency of an individual audio sample, and F_B is the mean fundamental frequency during the 200 ms pre-perturbation period. This calculation takes absolute units of frequency (Hertz) and translates them to relative units (cents, 100 cents = 1 semitone), which enables comparing participants and trials with varying pitch. Lastly, we excluded trials which contained vocal responses larger than 200 cents at any time point during the trial.

EEG data acquisition and analysis

The EEG signal was be recorded from 32 scalp electrodes with NeurOne Tesla amplifier (Bittium, 2022), software version 1.5.0. Two additional electrodes were used during data collection. One was placed on the forehead (in-between the FP1 and FP2 electrodes) and served as the ground. The second one was placed on the nose and served as the reference during data collection. Sampling rate was 500 Hz. Markers that denoted participants' keyboard responses (conscious perception ratings), and trial type were added to the EEG signal during recording. The onset of the pitch perturbation was marked in the EEG signal during data preprocessing based on Audapter data.

The EEG data was processed using EEGLAB, version 2022.1 (Delorme & Makeig, 2004) running in Matlab version R2022b. During preprocessing, channels containing a large proportion of noise and artifacts were rejected using the function pop_rejchan, using kurtosis and probability measures. The signal was high-pass filtered at 1 Hz using EEGLAB's function *pop_eegfiltnew*. Line noise around 50 Hz was reduced using the *zapline* function (de Cheveigné, 2020). The data was re-referenced to average reference of all scalp electrodes. The data was epoched relative to the perturbation onset, from 500 ms before the perturbation onset up to 1000 ms after the perturbation onset was used. Each trial was baseline corrected to the 500 ms time window before the perturbation onset. Next, we ran the Independent Component analysis (ICA), using the *pop_runica* function in 'extended' mode. After ICA,

removed or missing channels were interpolated using function pop_interp with default parameters. Independent components were automatically labelled using ICLabel, and only those that were classified as being brain based with more than 70 percent probability were kept. Finally, we removed artifactual trials using the *pop_jointprob* function (local and global thresholds = 3).

Statistical analysis

Vocal responses

We analyzed the vocal responses using a mass univariate analysis, that is, the effect of consciousness on the vocal response is examined at every time point before and after the onset of the perturbation. While our preregistration did not specify a mass univariate analysis, its use here is justified because it allows comparisons at high temporal resolution while controlling family-wise statistical error rate. Second, the method is suitable to situations where prior information about the timing of effects is limited, as is the case here. To run the mass univariate analysis of vocal responses, we have adopted a methodology typically used for mass univariate analysis of ERPs (Groppe et al., 2011; Pernet et al., 2015). First, we selected all critical trials, and ran a linear mixed effects regression model for each time point of the vocalization from 200 ms before the perturbation onset up to 800 ms after the perturbation onset. In the model, we treated the magnitude of the vocal response as the dependent variable, consciousness as a fixed-effect predictor, and the by-participant intercept as a random effect (pitch ~ consciousness + (1|participant)). This analysis resulted in a large number of statistical tests. To account for multiple comparisons, we next ran the same series of tests with 1000 random permutations of the original trials. To do this, we randomly assigned trials to either in the aware or the unaware conditions, and repeatedly ran the previously described statistical model. To identify clusters with a significant effect of consciousness, we ran threshold-free cluster enhancement (TFCE; Smith & Nichols, 2009) for both real regression results, and the results based on the random permutations using the *limo_tfce* function from the LIMO EEG toolbox (Pernet et al., 2011). We selected the maximum TFCE-value (out of regression models for each time sample between -200-800 ms) for each of the random permutations, and saved these to form a null-distribution to which the TFCE values of the real (i.e., not randomized) data was compared to. TFCE-values of real data higher than 97.5 of

randomly permuted TFCE-values were considered significant effect of consciousness (this is equal to two-tailed alpha level of 0.05).

ERPs

For the analysis of ERPs, we also used linear mixed-effects models on single-trial data. Unlike in the vocal response analysis, in the ERPs we expected to see two previously reported (Eklund et al., 2019, 2021; Eklund & Wiens, 2019; Filimonov et al., 2022; Schlossmacher et al., 2021) correlates: AAN (100-200 ms after stimulus onset), and LP (300-500 ms after stimulus onset). We analyzed these using two separate models, although only the analysis in the AAN time window was preregistered. In these models, we treated mean voltage during the time period (measured on a single-trial basis) as the dependent variable, consciousness as the fixed effect predictor, and the by-participant intercept as the random effect predictor (voltage ~ consciousness + (1|participant)). As specified in the preregistration, we only included the electrodes from the central electrode cluster (Cz, FC1, FC2, CP1, CP2) in this model (both AAN, and also LP are typically strongest in these electrodes (Eklund et al., 2019, 2021; Eklund & Wiens, 2019; Filimonov et al., 2022; Schlossmacher et al., 2021). To visualize the scalp distribution of AAN and LP in terms of t values, we examined how consciousness influenced ERPs by running the model separately for each channel.

To model which brain sources contribute to AAN and LP, we performed source estimation using sLoreta (standardized low resolution brain electromagnetic tomography), which solves the inverse problem of EEG source localization by generating a reference-free and polarity-independent single linear solution based on the linearly weighted sum of all scalp electrode potentials (Pascual-Marqui, 2002). Previous intracranial and brain imaging studies have validated the accuracy of the sLORETA (Mulert et al., 2004; Pizzagalli et al., 2004; Zumsteg et al., 2006). For each subject, sLORETA images of activation on cortical gray matter were generated using Montreal Neurological Institute's standardized template (MNI152). The resolution of the present reconstruction is limited by the small number of EEG electrodes, and the lack of individual MRIs (Song et al., 2015). Statistical analysis of source activations was performed separately for early (AAN; 100-200 ms) and late (LP; 300-500 ms) time windows by comparing the source density distributions in the conscious and unconscious conditions voxel-by-voxel using t tests. The results were corrected for multiple comparisons using permutation

testing. Unlike the analysis on ERPs (which was performed on single-trial data), sLoreta analysis was based on average ERPs. This analysis was not preregistered.

Author contributions

H.R. and D.S. designed the study. D.S. collected the data, and performed the statistical analysis, and visualizations. H.R. supervised data preprocessing and analysis, and R.B. supervised the source reconstruction and analysis. D.S. and H.R. wrote the first draft of the paper. All authors provided feedback and modifications to the manuscript, and approved the final version of the manuscript.

Acknowledgements

This research was funded by the Research Council of Finland (grant number 351109).

References

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience & Biobehavioral Reviews*, *36*(2), 737–746. https://doi.org/10.1016/j.neubiorev.2011.12.003

Behroozmand, R., Bonilha, L., Rorden, C., Hickok, G., & Fridriksson, J. (2022). Neural correlates of impaired vocal feedback control in post-stroke aphasia. *NeuroImage*, *250*, 118938. https://doi.org/10.1016/j.neuroimage.2022.118938

Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, *12*(1), 54. https://doi.org/10.1186/1471-2202-12-54

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *The Journal of Neuroscience*, *37*(40), 9603–9613. https://doi.org/10.1523/JNEUROSCI.3218-16.2017

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brown, R., Lau, H., & LeDoux, J. E. (2019). Understanding the Higher-Order Approach to Consciousness. *Trends in Cognitive Sciences*, *23*(9), 754–768. https://doi.org/10.1016/j.tics.2019.06.009

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

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

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences*, *110*(7), 2653–2658. https://doi.org/10.1073/pnas.1216827110

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*.

David, N., Newen, A., & Vogeley, K. (2008). The "sense of agency" and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, *17*(2), 523–534. https://doi.org/10.1016/j.concog.2008.03.004

de Cheveigné, A. (2020). ZapLine: A simple and effective method to remove power line artifacts. *NeuroImage*, *207*, 116356. https://doi.org/10.1016/j.neuroimage.2019.116356

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dembski, C., Koch, C., & Pitts, M. (2021). Perceptual awareness negativity: a physiological correlate of sensory consciousness. *Trends in Cognitive Sciences*, *25*(8), 660–670. https://doi.org/10.1016/j.tics.2021.05.009

Derda, M., Koculak, M., Windey, B., Gociewicz, K., Wierzchoń, M., Cleeremans, A., & Binder, M. (2019). The role of levels of processing in disentangling the ERP signatures of conscious visual processing. *Consciousness and Cognition*, *73*, 102767. https://doi.org/10.1016/j.concog.2019.102767

Dykstra, A. R., Halgren, E., Gutschalk, A., Eskandar, E. N., & Cash, S. S. (2016). Neural Correlates of Auditory Perceptual Awareness and Release from Informational Masking Recorded Directly from Human Cortex: A Case Study. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00472

Eklund, R., Gerdfeldter, B., & Wiens, S. (2019). Effects of a Manual Response Requirement on Early and Late Correlates of Auditory Awareness. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02083

Eklund, R., Gerdfeldter, B., & Wiens, S. (2021). The early but not the late neural correlate of auditory awareness reflects lateralized experiences. *Neuropsychologia*, *158*, 107910. https://doi.org/10.1016/j.neuropsychologia.2021.107910

Eklund, R., & Wiens, S. (2019). Auditory awareness negativity is an electrophysiological correlate of awareness in an auditory threshold task. *Consciousness and Cognition*, *71*, 70–78. https://doi.org/10.1016/j.concog.2019.03.008

Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*(7198), 1102–1106. https://doi.org/10.1038/nature06910

Filevich, E., Kühn, S., & Haggard, P. (2012). Negative motor phenomena in cortical stimulation: implications for inhibitory control of human action. *Cortex*, *48*(10), 1251–1261. https://doi.org/10.1016/j.cortex.2012.04.014

Filimonov, D., Railo, H., Revonsuo, A., & Koivisto, M. (2022). Modality-specific and modality-general electrophysiological correlates of visual and auditory awareness: Evidence from a bimodal ERP experiment. *Neuropsychologia*, *166*, 108154. https://doi.org/10.1016/j.neuropsychologia.2022.108154

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., Hartsuiker, R. J., Johansson, P., Hall, L., & Lind, A. (2023). Don't blame yourself: Conscious source monitoring modulates feedback control during speech production. *Quarterly Journal of Experimental Psychology*, *76*(1), 15–27. https://doi.org/10.1177/17470218221075632

Gärtner, K., & Gutschalk, A. (2021). Auditory cortex activity related to perceptual awareness versus masking of tone sequences. *NeuroImage*, *228*, 117681. https://doi.org/10.1016/j.neuroimage.2020.117681

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Guenther, F. H. (2016). *Neural Control of Speech*. The MIT Press. https://doi.org/10.7551/mitpress/10471.001.0001

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

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

Hickok, G. (2022). *The dual stream model of speech and language processing* (pp. 57–69). https://doi.org/10.1016/B978-0-12-823384-9.00003-7

Houde, J. F., & Nagarajan, S. S. (2011). Speech Production as State Feedback Control. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00082

Hurme, M., Koivisto, M., Revonsuo, A., & Railo, H. (2017). Early processing in primary visual cortex is necessary for conscious and unconscious vision while late processing is necessary only for conscious vision in neurologically healthy humans. *NeuroImage*, *150*, 230–238. https://doi.org/10.1016/j.neuroimage.2017.02.060

Jimenez, M., Poch, C., Villalba-García, C., Sabater, L., Hinojosa, J. A., Montoro, P. R., & Koivisto, M. (2021). The Level of Processing Modulates Visual Awareness: Evidence from Behavioral and Electrophysiological Measures. *Journal of Cognitive Neuroscience*, *33*(7), 1295–1310. https://doi.org/10.1162/jocn_a_01712

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, *34*(7), 885–912. https://doi.org/10.1016/0042-6989(94)90039-6

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22

Koivisto, M., Leino, K., Pekkarinen, A., Karttunen, J., Railo, H., & Hurme, M. (2021). Transcranial magnetic stimulation (TMS)-induced Blindsight of Orientation is Degraded Conscious Vision. *Neuroscience*, *475*, 206–219. https://doi.org/10.1016/j.neuroscience.2021.08.025

Koivisto, M., Revonsuo, A., & Lehtonen, M. (2006). Independence of Visual Awareness from the Scope of Attention: an Electrophysiological Study. *Cerebral Cortex*, *16*(3), 415–424. https://doi.org/10.1093/cercor/bhi121

Koivisto, M., Revonsuo, A., & Salminen, N. (2005). Independence of visual awareness from attention at early processing stages. *NeuroReport*, *16*(8), 817–821. https://doi.org/10.1097/00001756-200505310-00008

Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex*, *49*(7), 1935–1942. https://doi.org/10.1016/j.cortex.2012.09.002

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*(3), 204–220. https://doi.org/10.1080/17588921003731586

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X

Lane, H. L., Catania, A. C., & Stevens, S. S. (1961). Voice Level: Autophonic Scale, Perceived Loudness, and Effects of Sidetone. *The Journal of the Acoustical Society of America*, *33*(2), 160–167. https://doi.org/10.1121/1.1908608

Liu, Y., Hu, H., Jones, J. A., Guo, Z., Li, W., Chen, X., Liu, P., & Liu, H. (2015). Selective and divided attention modulates auditory-vocal integration in the processing of pitch feedback errors. *European Journal of Neuroscience*, *42*(3), 1895–1904. https://doi.org/10.1111/ejn.12949

Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

Mulert, C., Jäger, L., Schmitt, R., Bussfeld, P., Pogarell, O., Möller, H.-J., Juckel, G., & Hegerl, U. (2004). Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *NeuroImage*, *22*(1), 83–94. https://doi.org/10.1016/j.neuroimage.2003.10.051

Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception? *The Journal of Neuroscience*, *37*(40), 9593–9602. https://doi.org/10.1523/JNEUROSCI.3217-16.2017

Overgaard, M., Fehl, K., Mouridsen, K., Bergholt, B., & Cleeremans, A. (2008). Seeing without Seeing? Degraded Conscious Vision in a Blindsight Patient. *PLoS ONE*, *3*(8), e3028. https://doi.org/10.1371/journal.pone.0003028

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, *24 Suppl D*, 5–12.

Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: A Toolbox for Hierarchical LInear MOdeling of ElectroEncephaloGraphic Data. *Computational Intelligence and Neuroscience*, *2011*, 1–11. https://doi.org/10.1155/2011/831409

Pernet, C. R., Latinus, M., Nichols, T. E., & Rousselet, G. A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods*, *250*, 85–93. https://doi.org/10.1016/j.jneumeth.2014.08.003

Phillips, I. (2021). Blindsight is qualitatively degraded conscious vision. *Psychological Review*, *128*(3), 558–584. https://doi.org/10.1037/rev0000254

Picard, N., & Strick, P. L. (1996). Motor Areas of the Medial Wall: A Review of Their Location and Functional Activation. *Cerebral Cortex*, *6*(3), 342–353. https://doi.org/10.1093/cercor/6.3.342

Pitts, M. A., Lutsyshyna, L. A., & Hillyard, S. A. (2018). The relationship between attention and consciousness: an expanded taxonomy and implications for 'no-report' paradigms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1755), 20170348. https://doi.org/10.1098/rstb.2017.0348

Pitts, M. A., Martínez, A., & Hillyard, S. A. (2012). Visual Processing of Contour Patterns under Conditions of Inattentional Blindness. *Journal of Cognitive Neuroscience*, *24*(2), 287–303. https://doi.org/10.1162/jocn_a_00111

Pizzagalli, D. A., Oakes, T. R., Fox, A. S., Chung, M. K., Larson, C. L., Abercrombie, H. C., Schaefer, S. M., Benca, R. M., & Davidson, R. J. (2004). Functional but not structural subgenual prefrontal cortex abnormalities in melancholia. *Molecular Psychiatry*, *9*(4), 325–325. https://doi.org/10.1038/sj.mp.4001501

Railo, H., Revonsuo, A., & Koivisto, M. (2015). Behavioral and electrophysiological evidence for fast emergence of visual consciousness. *Neuroscience of Consciousness*, *2015*(1). https://doi.org/10.1093/nc/niv004

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, *3*(1), 1–23. https://doi.org/10.1023/B:PHEN.0000041900.30172.e8

Scheerer, N. E., & Jones, J. A. (2018). Detecting our own vocal errors: An event-related study of the thresholds for perceiving and compensating for vocal pitch errors. *Neuropsychologia*, *114*, 158–167. https://doi.org/10.1016/j.neuropsychologia.2017.12.007

Schlossmacher, I., Dellert, T., Bruchmann, M., & Straube, T. (2021). Dissociating neural correlates of consciousness and task relevance during auditory processing. *NeuroImage*, *228*, 117712. https://doi.org/10.1016/j.neuroimage.2020.117712

Schlossmacher, I., Dellert, T., Pitts, M., Bruchmann, M., & Straube, T. (2020). Differential Effects of Awareness and Task Relevance on Early and Late ERPs in a No-Report Visual Oddball Paradigm. *The Journal of Neuroscience*, *40*(14), 2906–2913. https://doi.org/10.1523/JNEUROSCI.2077-19.2020

Schneider, D. M., & Mooney, R. (2018). How Movement Modulates Hearing. *Annual Review of Neuroscience*, *41*(1), 553–572. https://doi.org/10.1146/annurev-neuro-072116-031215

SMITH, S., & NICHOLS, T. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Song, J., Davey, C., Poulsen, C., Luu, P., Turovets, S., Anderson, E., Li, K., & Tucker, D. (2015). EEG source localization: Sensor density and head surface coverage. *Journal of Neuroscience Methods*, *256*, 9–21. https://doi.org/10.1016/j.jneumeth.2015.08.015

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*. https://doi.org/10.3389/fnhum.2018.00082

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, *17*(7), 450–461. https://doi.org/10.1038/nrn.2016.44

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981. https://doi.org/10.1080/01690960903498424

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, *39*(3), 1429–1443. https://doi.org/10.1016/j.neuroimage.2007.09.054

Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., & Kuriki, S. (2007). Neural correlates of auditory feedback control in human. *Neuroscience*, *146*(2), 499–503. https://doi.org/10.1016/j.neuroscience.2007.02.023

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306–2319. https://doi.org/10.1121/1.2773966

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120. https://doi.org/10.3758/BF03202828

Wiegand, K., & Gutschalk, A. (2012). Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. *NeuroImage*, *61*(1), 62–69. https://doi.org/10.1016/j.neuroimage.2012.02.067

Xu, M., Tachibana, R. O., Okanoya, K., Hagiwara, H., Hashimoto, R., & Homae, F. (2020). Unconscious and Distinctive Control of Vocal Pitch and Timbre During Altered Auditory Feedback. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01224

Zito, G. A., Wiest, R., & Aybek, S. (2020). Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLOS ONE*, *15*(6), e0234321. https://doi.org/10.1371/journal.pone.0234321

Zumsteg, D., Lozano, A. M., & Wennberg, R. A. (2006). Mesial Temporal Inhibition in a Patient with Deep Brain Stimulation of the Anterior Thalamus for Epilepsy. *Epilepsia*, *47*(11), 1958–1962. https://doi.org/10.1111/j.1528-1167.2006.00824.x