



## Self-monitoring in the cerebral cortex: Neural responses to small pitch shifts in auditory feedback during speech production



Matthias K. Franken<sup>a,b,1</sup>, Frank Eisner<sup>a</sup>, Daniel J. Acheson<sup>a,b</sup>, James M. McQueen<sup>a</sup>, Peter Hagoort<sup>a,b</sup>, Jan-Mathijs Schoffelen<sup>a,\*</sup>

<sup>a</sup> Donders Institute for Brain, Cognition and Behaviour, Radboud University, P.O. Box. 9101, Nijmegen, 6500 HB, The Netherlands

<sup>b</sup> Max Planck Institute for Psycholinguistics, P.O. Box 310, Nijmegen, 6500 AH, The Netherlands

### ARTICLE INFO

#### Keywords:

Altered auditory feedback  
Event-related fields  
Forward model  
MEG  
Oscillations

### ABSTRACT

Speaking is a complex motor skill which requires near instantaneous integration of sensory and motor-related information. Current theory hypothesizes a complex interplay between motor and auditory processes during speech production, involving the online comparison of the speech output with an internally generated forward model. To examine the neural correlates of this intricate interplay between sensory and motor processes, the current study uses altered auditory feedback (AAF) in combination with magnetoencephalography (MEG). Participants vocalized the vowel /e/ and heard auditory feedback that was temporarily pitch-shifted by only 25 cents, while neural activity was recorded with MEG. As a control condition, participants also heard the recordings of the same auditory feedback that they heard in the first half of the experiment, now without vocalizing. The participants were not aware of any perturbation of the auditory feedback. We found auditory cortical areas responded more strongly to the pitch shifts during vocalization. In addition, auditory feedback perturbation resulted in spectral power increases in the  $\theta$  and lower  $\beta$  bands, predominantly in sensorimotor areas. These results are in line with current models of speech production, suggesting auditory cortical areas are involved in an active comparison between a forward model's prediction and the actual sensory input. Subsequently, these areas interact with motor areas to generate a motor response. Furthermore, the results suggest that  $\theta$  and  $\beta$  power increases support auditory-motor interaction, motor error detection and/or sensory prediction processing.

### Introduction

Speaking is a remarkably complex motor skill. We speak at a rate of often more than 10 speech sounds per second, each of which require accurate coordination of more than 100 different muscles. We make use of this skill day in day out, throughout our lives, usually without conscious awareness of the complexity of the task. If attention is paid to phonological aspects of speech production, it is mostly focused on wording, while articulation follows effortlessly. In order to perform this motor task almost without errors, a good quality control system is needed. Recent developments in speech motor control have shown that integration of sensorimotor information, including auditory feedback (i.e. the sound of our own voice), is key in this respect. The current study investigates the neural underpinnings of sensorimotor integration during speech production.

The role of auditory feedback in speech production has been

investigated by providing speakers with online manipulated feedback (Houde and Jordan, 1998; Burnett et al., 1998; Jones and Munhall, 2000). For example, speakers could be hearing their own speech in real time at a higher pitch or with a lower first formant. It turns out that speakers usually compensate for these manipulations by changing their speech in the opposite direction (that is, by lowering the pitch, or by increasing the frequency in the first formant, which results in a change in vowel quality). This compensatory response occurs even when participants are told to ignore the altered feedback (Keough et al., 2013). This suggests that speakers automatically monitor their auditory feedback during speech production. Cognitive modeling work in this context has drawn from principles in motor control more generally, in order to explain such a fast feedback monitoring mechanism (Wolpert et al., 1995; Wolpert and Ghahramani, 2000). These models hypothesize the use of internally generated forward models (Houde and Nagarajan, 2011; Tourville and Guenther, 2011). Specifically, all articulatory motor

\* Corresponding author.

E-mail addresses: [matthias.franken@ugent.be](mailto:matthias.franken@ugent.be) (M.K. Franken), [jan.schoffelen@donders.ru.nl](mailto:jan.schoffelen@donders.ru.nl) (J.-M. Schoffelen).

<sup>1</sup> Present address: Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium.

programs which are generated in (and will be executed by) the motor system are sent to the auditory system. Each of these efference copies can be used to create a forward model, which models the sensory (auditory) consequences of the articulation. This sensory prediction can then be compared with the observed sensory consequences, and if necessary generate a prediction error that could signal the need for behavioral adaptation.

Using the altered auditory feedback paradigm, several functional magnetic resonance imaging studies have shown that feedback processing is supported by an extended bilateral functional neural network including auditory and motor-related areas (Behroozmand et al., 2015b; Zarate et al., 2010; Zarate and Zatorre, 2005; Zheng et al., 2010; Zheng et al., 2013). Electrophysiological studies using electroencephalography (EEG) to investigate the temporal dynamics of feedback processing have shown that altered feedback leads to a brain response as early as 100 ms after perturbation onset (Behroozmand et al., 2009; Behroozmand et al., 2011; Behroozmand and Larson, 2011; Hawco et al., 2009). The early latency of these findings suggests that auditory processing and motor control already interact at an early processing stage. In addition, in a MEG study, Kort et al., (2014) show responses of a broad bilateral cortical network to an unexpected 100-cent pitch shift in auditory feedback. These authors found enhanced neural activity in response to pitch perturbations in sensorimotor, auditory and premotor cortices.

The current study investigates the neural correlates of pitch perturbation processing and of the subsequent automatic responses to these perturbations. Importantly, we used a small perturbation magnitude (25 cents), to make sure that the participants did not consciously detect the perturbation. This was done to substantiate the claim that speakers' responses to altered auditory feedback are not subject to conscious awareness (Behroozmand et al., 2015a). In most studies, the perturbations used are large enough to trigger conscious processing, and therefore possibly recruit attentional resources. Since it has been established that attention can indeed modulate speakers' responses to unexpected auditory feedback (Hu et al., 2015; Korzyukov et al., 2012; Liu et al., 2015), it is crucial to avoid attentional effects by keeping the perturbation small.

In addition, in this study we performed a detailed analysis of neural oscillatory activity in relation to the feedback perturbations. So far, only a small number of studies on feedback perturbations have looked beyond evoked responses. This may be surprising, as recent dynamic approaches to cognition have linked cortical oscillations to predictive processing (Engel et al., 2001) and sensorimotor integration more generally (Caplan et al., 2003), as well as to speech production specifically (Cruikshank et al., 2012; Gehrig et al., 2012; Jenson et al., 2014). Two recent studies suggested that spectral power increases in the  $\delta$  (1–4 Hz),  $\theta$  (4–8 Hz) and  $\gamma$  (65–150 Hz) bands over motor and sensory areas reflect sensorimotor speech processing (Behroozmand et al., 2015a; Kort et al., 2016). The current study looks at responses in the lower frequency range to a much smaller pitch shift (only 25 cents instead of 100 cents).

We also investigated the neural correlates of the different types of response (opposing versus following) to the perturbation. Although the typical response to a feedback pitch perturbation (for instance: an increase) is a compensatory change in the opposing direction (for instance: a decrease), occasionally participants respond by actually following the direction of the perturbation (Behroozmand et al., 2012; Franken et al., 2018; Larson et al., 2007).

## Materials and methods

### Subjects

Thirty-nine healthy volunteers (age:  $M = 22$ , range = 18–34; 27 females) participated after providing written informed consent in accordance with the Declaration of Helsinki and the local ethics board committee (CMO region Arnhem/Nijmegen). All participants had normal hearing, were native speakers of Dutch and had no history of speech and/or language pathology.

### Paradigm

An experimental session consisted of two tasks, a speaking and a listening task, always performed in the same order (speaking, then listening), while brain activity was measured using MEG.

In the speaking task, participants performed a tone-matching task (Liu and Larson, 2007; Hawco et al., 2009). This task was chosen to keep participants attentive. A trial started with the presentation of a short tone (duration 700 ms). 200 ms after the tone offset, a visual cue (“EE”, in Dutch pronounced as/e/) instructed the participants to start vocalizing/e/, while trying to match the pitch of the tone they just heard. The visual cue disappeared after 3s, cueing the participant to stop vocalizing. During speech production/vocalization, the participant's voice was recorded using a microphone, positioned about 1.5 m from the participant to avoid any artifacts in the MEG signal. The recorded signal was used to provide the participants with online auditory feedback. In half of the trials, participants received normal auditory feedback throughout the trial, i.e. participants' speech was recorded and played back to them unaltered (henceforth control trials). In the other half of the trials (perturbation trials), auditory feedback was normal at first, but, starting between 500 and 1500 ms after speech onset (randomly jittered), the feedback's pitch was increased by 25 cents for a duration of 500 ms, before returning back to normal feedback for the remainder of the trial. The only difference in auditory feedback between control and perturbation trials was this 500 ms pitch shift. The duration of the pitch shift is rather long compared to previous studies (Burnett et al., 1998; Hain et al., 2000), in order to have a broad time window during the shift for time-frequency analyses. The shift duration is not much longer compared to the 400 ms shifts in Kort et al., (2014, 2016). Overall, participants received 99 perturbation trials and 99 control trials, randomly mixed in two blocks of 99 trials each. After the speaking task, participants did the passive listening task, in which the participants were shown the same visual cues as in the production task, but were instructed not to speak. Instead, they listened to recordings of the very same feedback they were given in the speaking task.

Finally, after the experiment, participants filled out a short debriefing questionnaire, which asked whether they noticed any feedback manipulations and if so, what kind of manipulations.

### Materials

The tone stimuli were 700 ms pure tones at one of three pitch frequencies. The pitch of the tones was individually tailored to the participants at 4, 8 and 11 semitones above their conversational pitch. This was done by having participants produce the vowel/e/five times (they were not yet aware the experiment would involve pitch), and the average pitch was considered their conversational pitch.

The auditory feedback shifts were implemented using Audapter software (Cai et al., 2008; Tourville et al., 2013). In brief, the software performs a near-real-time autocorrelation analysis to track the pitch. In order to shift the pitch, the short-time Fourier spectra were stretched and interpolated along the frequency axis. The pitch-shifted sounds were played back to the speaker through audio air tubes with a latency of 10–20 ms.

All voice recordings were made on one channel using a Sennheiser ME64 cardioid microphone, which was set up in the MEG magnetically shielded room and connected through an in-house-built audio mixer to a dedicated soundcard Motu MicroBook II outside the room, which was connected to a Windows computer. Auditory feedback was delivered through the same soundcard which was connected to CTF (VSM/CTF systems, Port Coquitlam, Canada) audio air tubes. Stimulus presentation and sound recording times were controlled by the same Windows computer running Audapter and MathWorks Matlab (MathWorks, Version 8 Release 5, Natick, MA).

Download English Version:

<https://daneshyari.com/en/article/8686706>

Download Persian Version:

<https://daneshyari.com/article/8686706>

[Daneshyari.com](https://daneshyari.com)