Reduced speaking-induced suppression of auditory evoked-potentials in early Parkinson's disease suggests deficits in monitoring self-produced speech

Henry Railo[1,2], Niklas Nokelainen[3], Saara Savolainen[3], & Valtteri Kaasinen[2,4,5]

[1] Department of Clinical Neurophysiology, University of Turku, Turku, Finland
[2] Turku Brain and Mind Centre, University of Turku, Turku, Finland
[3] Department of Psychology, University of Turku, Turku, Finland
[4] Department of Neurology, University of Turku, Turku, Finland
[5] Division of Clinical Neurosciences, Turku University Hospital, Turku, Finland

Corresponding author:
Henry Railo,
Department of Clinical Neurophysiology,
University of Turku and Turku University Hospital,
Finland,
henry.railo@utu.fi

Abstract

Speech deficits are a common symptom of Parkinson's disease (PD), but the neural mechanisms that cause them are not well understood. Paradoxically, the patients may not subjectively notice their own speech deficits such as hypophonic speech. We hypothesized that this reflects an impairment in monitoring self-produced speech: updating the internal speech model may malfunction in PD, which could explain their speech deficit. We measured evoked EEG responses to self-produced and passively heard phonemes in PD patients (N=17) with mild speech impairment and age-matched controls (N=18). When the speech sounds are self-produced, the amplitude of the evoked response is lower when compared to the condition where the same sounds are passively heard. This effect is known as speaking-induced suppression, and it reflects the modulation of sensory cortical processing through top-down corollary discharge predictions. We observed that the speaking-induced suppression was significantly reduced in PD patients, suggesting of abnormally high prediction errors (when compared to healthy controls). In addition, source reconstruction showed that activity in auditory cortex was stronger in PD patients than control participants already -200 ms before phonation as well as during a later time period (400-500 ms after phonation). We conclude that altered internal monitoring of speech may underlie speech deficits in PD.

Key words: Parkinson's disease, speech impairment, speech deficit, hypokinetic dysarthria, corollary discharge, efference copy, speech-induced suppression, prediction, EEG, auditory cortex

Introduction

The majority of patients with Parkinson's disease (PD) show speech impairments typically classified as hypokinetic dysarthria (Ho et al., 1998; Logemann et al., 1978; Sapir et al., 2001). The major speech symptoms are inaccurate articulation, variable speech rate, monotonic and hypophonic loudness of speech, and harsh voice quality. Speech deficits are difficult to treat. Often they do not respond to dopamine therapy (Louis, Winfield, Fahn, & Ford, 2001; Skodda, Visser, & Schlegel, 2010), and dysarthria is a well-known side-effect of deep-brain stimulation (Follett et al., 2010). Treatment of speech impairments is difficult in part because its neural basis is not understood.

Speech, just like any other actions, requires the brain to continuously monitor how well the produced actions match the desired outcome (e.g. Hickok, 2012). We hypothesized that speech deficits in PD reflect problems in monitoring self-produced speech. PD patients are often not fully aware of their speech deficits (Kwan & Whitehill, 2011), whereas they are typically well aware of their primary motor symptoms (Leritz, Loftis, Crucian, Friedman, & Bowers, 2004). PD patients overestimate the volume of their own speech, and report experiencing speaking in loud voice, although their speech sound pressure level is actually reduced (Fox & Ramig, 1997; Ho, Bradshaw, & Iansek, 2008). PD patients also respond abnormally to perturbation in their speech, suggesting deficits in monitoring mechanisms (Huang et al., 2016; Kiran & Larson, 2001; Liu, Wang, Metman, & Larson, 2012).

According to the State Feedback Control model (SFC) of speech production, the brain generates a continuously updated internal model of the desired action (Houde & Nagarajan, 2011). Motor commands are produced by first estimating the current state of the system, and then generating a command based on these estimates. The observer tracks and updates actions by relying on copies of the motor command called corollary discharge (CD), or efference copy signals. During speech one of the major functions of these predictions modulate the activity of the auditory cortex. These predictions allow the brain to internally monitor whether or not the produced speech sounds match the predictions (i.e. motor commands). This enables rapid and efficient modulation of speech production.

The functioning of the auditory speech monitoring system is examined by comparing the auditory evoked activity produced by self-produced speech to the same sounds when they are not self-produced. A large body of studies shows that during speech, the amplitude of auditory evoked cortical activity is suppressed when compared to a condition where the same sounds passively listened (Chang, Niziolek, Knight, Nagarajan, & Houde, 2013; Curio, Neuloh, Numminen, Jousmäki, & Hari, 2000; Ford & Mathalon, 2005; Ford et al., 2013; Houde, Nagarajan, Sekihara, & Merzenich, 2002). This phenomenon is called the speaking-induced suppression (SIS; Houde et al., 2002), and it is assumed to reflect the influence of the CD signals on auditory cortex. On scalp recorded electroencephalography (EEG), SIS is strongest over central sites around 100 ms after speech/sound onset (i.e. N100 wave of the event-related potential, ERP). We hypothesized that PD patients' internal model incorrectly signals that strong speech was produced, when in fact speech output was hypophonic. This could be due to abnormal CD signals, or deficits in comparing the CD prediction to the incoming auditory activation (see Friston et al., 2012). These deficits may, in turn, reflect pathological changes in e.g. dopamine system in the PD (Friston, Bastos, Pinotsis, & Litvak, 2015). Demonstration of an altered SIS in PD would be an important discovery because it

would suggest that the neural mechanism underlying speech impairment relates to abnormal monitoring of speech production. By "monitoring" we here refer to all processes using which the individual keeps track of how the produced sounds match the desired outcome.

Methods

Participants

We tested 20 PD patients and 20 age-matched control participants without neurological disease. PD patients were diagnosed using the UK Brain Bank criteria. One control participant was left-handed, others were right-handed. Comparison of the demographical information, and clinical and neuropsychological measurements are presented in Table 1. The two groups are similar with respect to other measures than motor and speech deficits, which were assessed using MDS-UPDRS scale (Movement Disease Society, Unified Parkinson's Disease Rating Scale, part III, motor symptoms). None of the patients or control participants reported experiencing any problems with speech.

Mini Mental-State (MMSE) questionnaire was used to assess cognitive impairment (Folstein, Folstein, & McHugh, 1975). One PD patient was excluded from the data analysis because his MMSE-score suggested possible mild cognitive impairment (MMSE score was 23). Beck's Depression Inventory was used to asses depression (Beck, 1976). One PD patient was excluded due to relatively high score (27 total score) in BDI. Finally, one patient had to be excluded because of a technical failure during EEG recording, and two control participants were excluded because of massive artifacts in EEG. Therefore, the final sample was 17 patients and 18 control participants. We also run the analyses by including the two PD patients who were excluded (due to low MMSE and high BDI), and the results did not change.

Table 1. Comparison of demographic, clinical and neuropsychological data between PD and control groups

	Control	PD	test statistic*	p
N	18	17	n/a	n/a
gender (male/female)	(6/12)	(7/10)	5.1	0.08
Age (mean +/- SD)	67.0 (6.1)	69.5 (7.7)	0.9	0.4
medication break (count)	n/a	12	n/a	n/a
Self-reported speech deficits	0	0	n/a	n/a
Disease duration, years (mean +/- SD)	n/a	6 (5)	n/a	n/a
Mini-mental state score (mean +/- SD)	28.1 (1.3)	28.1 (1.6)	0.6	0.5
Beck's depression inventory	4.7 (3.0)	6.3 (2.5)	1.4	0.2
MDS-UPDRS, motor (mean +/- SD)	4.9 (3.5)	23.0 (10.2)	6.7	< 0.001
MDS-UPDRS, speech (mean +/- SD)	0.05 (0.2)	0.5 (0.7)	2.4	0.02
* t statistic (continuous variables) or Chi-squared statistic (categorical data)				

Eleven patients were on levodopa medication, and sixteen were treated with MAO-inhibitors and/or dopamine agonists. One patient was on anti-cholinergic medication. The mean levodopa equivalent dose across the patients was 568 mg (SD = 492 mg). The patients were asked to go through a voluntary medication break before the test session to counteract possible intervening effects of medication. Twelve patients underwent the medication break, and did not take their morning medication during the day of the study (at least 12 h break in medication). These patients can therefore be considered being in OFF phase. However, because the present study did not

include a systematic manipulation/control of medication, possible intervening effects of medication cannot be fully excluded.

The experiment was approved by the Ethics committee of the hospital district of South-West Finland. A written informed consent was obtained prior the study. The study was conducted according to the principles of the Declaration of Helsinki. Participants received a small fee for participating.

Task and procedure

The present study included two parts, a psychophysical procedure designed to measure hypophonia, and a protocol where EEG was measured during speech (phoneme production) and passive listening conditions.

In the *psychophysical task* the participants listened to /a/ phonemes from a loudspeaker (Creative T40 Series II), and they were asked to repeat the phoneme with a matching volume. There were five different volume levels, ranging from weak to loud speaking volumes (on average 67 dB), but otherwise the phonemes were identical. This procedure was repeated 65 times in total (13 times for each volume level). The participant's response was measured using a microphone (Vivitar TVM-1) positioned halfway between the loudspeaker and the participant. The volume of the spoken response was measured from the microphone signal using a Matlab script. Response amplitude was determined by calculating the mean absolute amplitude of samples that were higher than 50% of the maximum absolute amplitude. Before each test session the position of the microphone was calibrated to ensure that volume was measured accurately. This calibration consisted of playing an identical phoneme as a response from a loudspeaker positioned where the participants head should be. This procedure yielded near perfect correspondence between stimulus and response, verifying that the volume measurement was accurate. No EEG was measured during this task. Our hypothesis was that if the patients show hypophonic speech they should systematically produce lower volume responses than control participants.

The participant was first asked to utter the phoneme /a/ approximately every 5 seconds for 10 minutes (this is called the "Speak condition"). They were asked to do this with their normal speaking volume while minimizing facial movements to reduce motor EEG artifacts. The participants wore headphones with a microphone. White noise and the participant's own speech was played back through the headphones. The volume of the participant's own speech heard through the microphone was adjusted to be similar to what is heard without microphones. The reason for this procedure was that it enabled us to precisely and straightforwardly measure the onset of the phoneme from the signal of the headphones to use it as a marker in EEG analysis (see below for details). During the task the participants were asked to fix their gaze on a dot presented in front of them on a computer screen. This was done to minimize eye movements during EEG recording. In the second condition (below referred to as the "Listen condition"), the participants sat quietly, fixating the dot in front of them, and the same phonemes they had just spoken were played back to them through the headphones (again with white noise in the background). In sum, the EEG recording consists of two conditions which are identical with respect to auditory stimulation, but in one of them the sounds are spoken by the participant, and in the other they passively listen to the sounds.

To accurately measure the onset time of the phonemes, the signal from the loudspeakers of the headphones was recorded by the EEG amplifier. The auditory signal is thus measured by the same system as brain activity, which means that the two signals are perfectly synchronized. Markers were inserted to the EEG data to the time-points that corresponded to the onsets of phonemes following a semi-automated procedure. First, the signal that contained the signal from the headphones was high-pass filtered at 100 Hz; because only high-frequency information is present, the transient onset of sound could be easily detected. Second, if the signal remained above a predefined threshold for minimum of 10 samples, a marker was added to the timepoint where the threshold was first crossed. All markers and trials were inspected by the researcher, and any trials where the onset of the sound was unclear, were excluded from the analysis. The threshold was adjusted for individual participants. As seen in Figure 2A, which shows the EEG-recorded audio signal averaged over the two groups, the marker was accurately added to the onset of the sound.

Electroencephalography

EEG was measured with 64 active electrodes with NeurOne Tesla amplifier. Sampling rate was 500 Hz. In addition, EOG was measured from electrodes placed beside and below the right eye.

EEG was processed using EEGLAB v14.1.1 software (Delorme & Makeig, 2004) on Matlab 2014b. Channels were first automatically inspected (pop_rejchan function examining probability, kurtosis, and spectrum; SD = 4), and bad channels interpolated. EEG was fist high-pass (0.1 Hz cutoff, filter-order = 16500) and then low-pass (40 Hz cutoff, filter order = 166) filtered (pop_eegfiltnew). Speech related and other EEG artefacts were removed using artifact subspace reconstruction (ASR), using the following parameters: channel criterion = 0.8, burst criterion = 5, and window criterion = 0.05. ASR resembles principle component analysis (PCA) in that data is reconstructed after first excluding artifactual components. ASR rejects components automatically by determining thresholds based on clean segments of EEG data. Here, ASR was utilized instead of PCA or independent component analysis (ICA) because large speech-related artefacts remained in the data after PCA or ICA-based artifact rejection. For technical details of ASR, the reader is referred to (Mullen et al., 2015). After ASR, data was epoched, removed electrodes were interpolated, and the data referenced to linked mastoids. Epochs containing abnormal data were rejected using pop_jointprob function (local and global cutoff = 2 and 4 SD, respectively). This left on average 105 (SD = 53) trials per participant for the EEG analyses.

Source reconstruction analysis was performed in Brainstorm software (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011), following the guidelines provided by Stropahl et al. (Stropahl, Bauer, Debener, & Bleichner, 2018). The source reconstruction procedure estimates the current in large number of sources distributed over the cortex. Dynamic statistical parametric mapping, which estimates the location of deep sources more accurately than minimum-norm solutions (Lin et al., 2006), was applied to the data (Dale et al., 2000). Because we did not have the individual participants' anatomical images, an MNI anatomical template (Holmes et al., 1998) was used. This means that the localization accuracy of the current approach cannot be compared to brain imaging. However, here our motivation to use source reconstruction was not spatially precise localization, but extract rough time course of activity from auditory cortices. During source reconstruction, prestimulus interval between -600 and -400 ms was used to calculate the noise covariance matrices. This time-window was chosen because we assumed that preparatory activation could be present right before speech onset. The boundary element method (BEM), as implemented in OpenMEEG, was used as a head model (default parameters) (Gramfort et al.,

2013; Stenroos, Hunold, & Haueisen, 2014). The orientation of the current dipole was constrained to be perpendicular with respect to cortical surface. The source localization was performed on each participant's averaged evoked response.

The time course of auditory cortical activity was extracted from the source reconstructed data. We extracted the maximum amplitude of dipoles in the posterior ramus of the lateral sulcus, based on the Destrieux atlas (region 41) (Destrieux, Fischl, Dale, & Halgren, 2010). This region was selected because it showed the strongest auditory evoked activation in the listen condition (across all participants; see Fig. 3). The absolute mean value of dipoles in this region was used as a measure of activation strength.

Statistical analysis

The results of the psychophysical task were analyzed using linear mixed-effects model. The model included the intensity of the test stimulus, group, and their interaction as predictors. Control group was coded as 0 (i.e. baseline), and the PD group as 1.

Evoked EEG analysis was performed in two steps. First, we characterize scalp recorded ERPs and SIS. Second, source reconstruction was run to estimate the time-course of activations in auditory cortices. Statistical analysis of scalp recorded EEG was performed on electrode Cz as this site typically (and also in our study) shows strongest SIS. We used the non-parametric permutation approach to test for statistical significance (Maris & Oostenveld, 2007). First, we calculated the SIS by subtracting Listen condition ERPs from the Speak condition ERPs. We then compared if the SIS was different in PD when compared to control group using a t test performed on each time point between -200 and 600 ms relative to stimulus/speech onset. Statistical significance of the difference was determined by comparing the observed test statistics to a null distribution obtained by random permutation of the data. To obtain the null distribution we randomly shuffled the group labels and calculated the t tests 1000 times. A cluster of consecutive statistically significant ($p < .05$) effects whose summed absolute t-value was highest was saved after each permutation; these values make up the null distribution. A difference between PD and control group was counted as statistically significant if the probability of observing such cluster mass of t values was smaller than 2.5% when compared to the null distribution. This type of clustering method takes into account the fact that true differences cluster together in time. The source reconstructed activity (auditory cortical ROI) was analyzed in the same manner.

The datasets are available for download at https://osf.io/eys97/.

Results

Speech deficits, and the psychophysical task

None of the patients or control participants subjectively reported experiencing any speech deficits. While most participants and PD patients were assessed to not have speech deficits according to the item representing speech deficit in MDS-UPDRS, PD patients were observed to have speech deficits more frequently than control participants (Table 1). Six of the 17 patients were rated with some speech deficits in the UPDRS (i.e. score 1 or higher).

Our hypothesis was that even though the patient's report not experiencing speech deficits, hypophonia would still be detectable using psychophysical assessment. Both patients and control participants produced higher volume speech to louder test stimuli, as shown in Figure 1. In

general, for low volume phonemes participants/patients tended to produce a higher volume sound than the test stimulus, whereas for higher sound volumes they tended to reproduce lower volume sounds. When analyzed using Test Stimulus Amplitude x Group linear mixed-effects regression (df = 1696) where individual participants' intercepts are allowed to vary randomly, the main effect of Test Stimulus Amplitude strongly predicted their response (β = 0.49, CI = [0.45–0.52], t = 28.05, p < .0001), and the interaction between Test Stimulus Amplitude and Group also reached statistical significance (β = -0.08, CI = [-0.13–-0.035], t = -3.36, p = 0.0008). This interaction indicates that, on average, the influence of Test Stimulus Amplitude (i.e. slope) was smaller in PD patients (this result is visualized in Figure 1). However, it should be noted that this effect is influenced by few individual patients, and when the between-subject variation in the slope is taken into account as a random-effect factor (in addition to between-subject variation in the intercept), the interaction is not statistically significant (β = -0.068, CI = [-0.20–0.06], t = -0.99, p = 0.32). Altogether these results indicate that while PD patients as a group were less successful in matching their speech volume to the test stimulus when compared to the control participants, there was substantial variation between individual participants/patients.

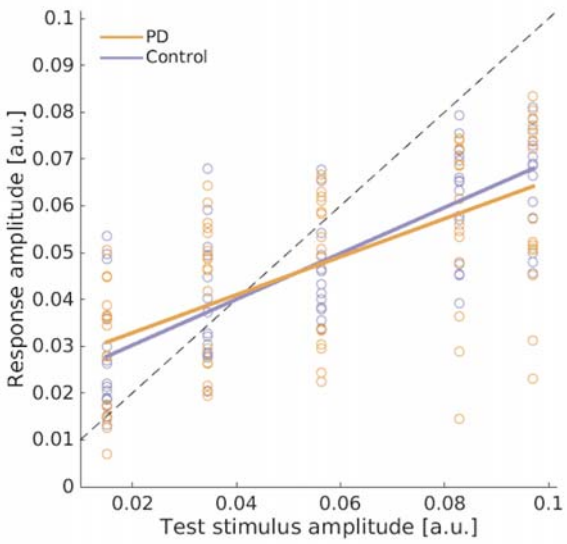

Figure 1. Results of the psychophysical task. PD patients and control participants are depicted as different colors (PD = orange). Dots represent individual participants/patients. The solid lines show the results of the linear model.

EEG

As shown in Figure 2, ERPs time-locked to the onset of the spoken (Speak condition; Fig. 2B) or heard (Listen condition, Fig. 2C) phoneme were calculated. In both conditions, N1 (negative peak at 80 ms), and P2 (positive peak at 190 ms) peaks are observed. The topography of the ERPs is shown in Figure 2D (averaged across patients and control participants). Previous studies employing this approach have shown that at central electrode sites the amplitude of the N1 ERP produced by the spoken phoneme is decreased when compared to the same stimuli when passively heard. As can be seen in Fig. 2D at time point 80 ms (N1), this same pattern is observed in the present study. This difference between the Speak and Listen conditions is assumed to reflect the CD signal: the speech production systems of the brain rely top-down predictions (CD)

about the upcoming speech, and hence the auditory cortex "knows what to expect". The P2 time window reveals a similar, but weaker effect.

The SIS – that is, the difference between the Speak and Listen conditions – is shown in Figure 2E (electrode Cz). The red markers on the x axis show the statistically significant differences between the groups: PD patients show a reduced SIS around 100 ms (i.e. N1 time window). Figure 2F shows the scalp distribution of the SIS at 80 ms after speech/sound onset in control participants and patients.

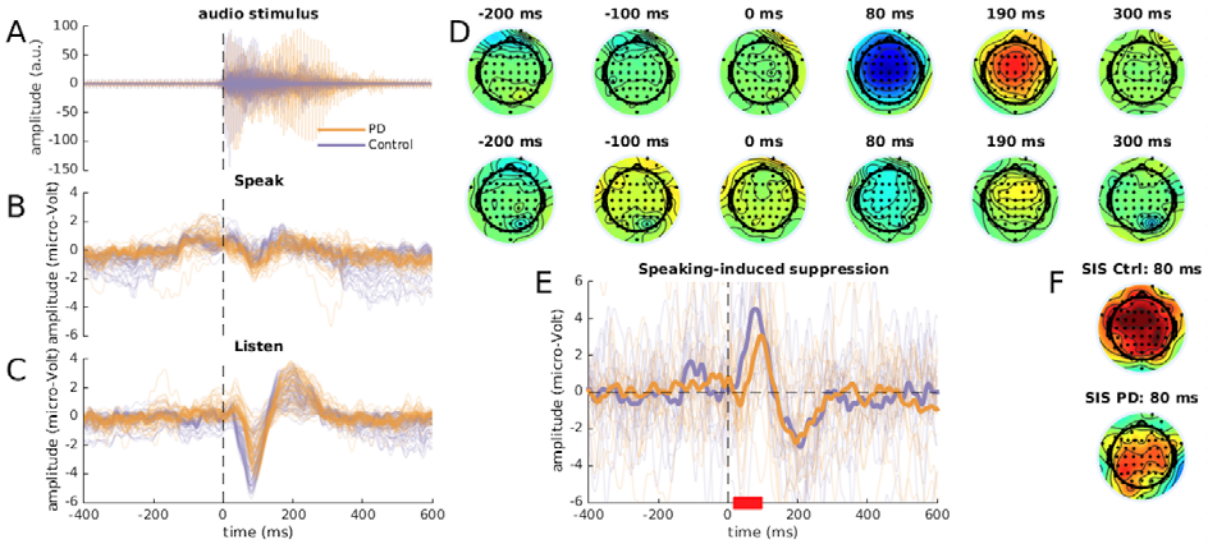

Figure 2. Auditory ERPs. In each panel the orange line is PD patients, and blue line is the control group. A) The auditory signal recorded by EEG (average over all participants/patients). B) A butterfly plot of ERPs in the Speak condition. C) A butterfly plot of ERPs in the Listen condition. Each line depicts one channel. D) Scalp topographies at different timepoints in Listen (above) and Speak (below) conditions, averaged across PD patients and control participants. Colormap limits are -5 to 5 microVolt. E) SIS in central parietal electrode (Cz). Faint lines in the background are individual participants' data, and the bold lines are group averages. The red box on the x axis shows the time window where the two groups differ according to the cluster-mass permutation test. F) Topography of SIS in control participants and patients at 80 ms after stimulus onset. The color here is the t statistic (limits -6 to 6) of a one sample t test against zero (not corrected for multiple comparisons).

During speech, the CD signal is assumed to modulate activity in auditory cortex. Source-reconstruction was run to estimate the time-course of activation of the auditory cortex. As shown in Fig. 3A, when sounds were passively listened, the auditory cortex activation revealed two activation peaks whose latencies closely resembled the latencies of the N1 and P2 ERP waves. During speech the activation is smaller, and begins before the onset of the spoken sound, especially in PD patients. The estimated auditory cortical SIS is shown in Figure 3B: negative values here indicate suppression of activity, and positive values stronger activity in the Speak condition relative to Listen condition. As can be seen the two groups differ in three time-windows: about 200 ms before phonation, and after speech onset at an early (100-200 ms) and late (400-500 ms) time window.

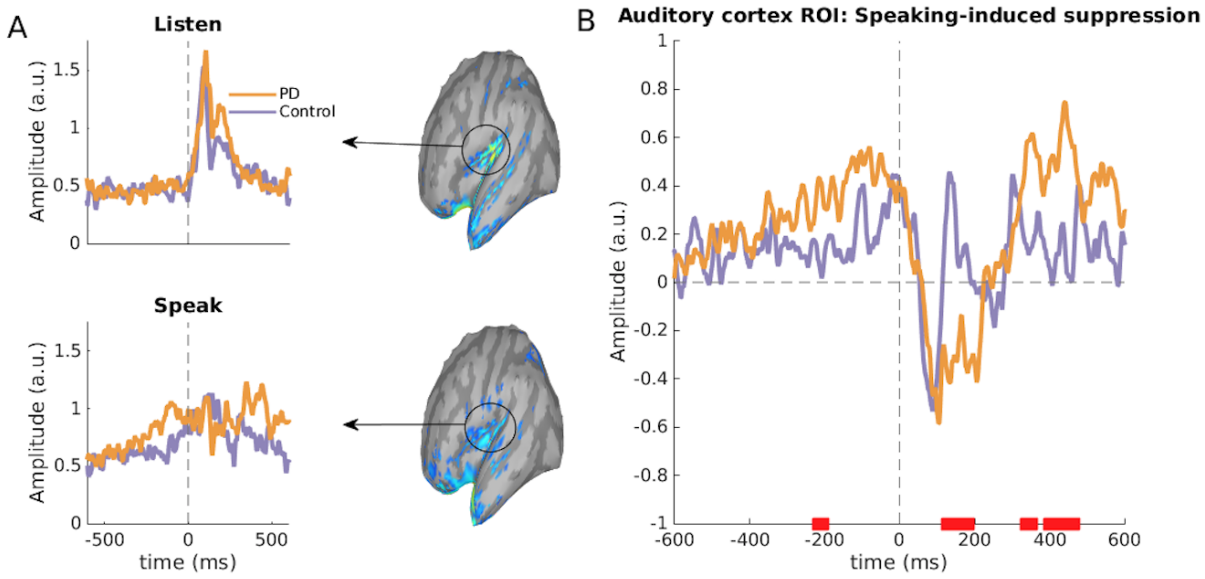

Figure 3. Source reconstruction results visualized. A) Cerebral activations in the Listen (top) and Speak (below) conditions, presented on an inflated cerebral hemisphere. The left panels show the time course of the averaged signal from the left and right auditory cortical ROIs. The orange line is PD, and the blue line the control group. B) SIS in PD patients and control participants in the ROI. The red markers on the x axis show statistically significant effects corrected for multiple comparisons.

Discussion

We compared whether speaking-induced suppression (SIS) of evoked auditory activity differs between PD patients and age-matched healthy controls. SIS reflects top-down modulation of auditory cortex activity during speaking. In a sample of PD patients with very mild speech deficits, we observed reduced SIS. This finding suggests that the monitoring of self-produced speech is impaired in PD. While we expected increased SIS (indicating that predictions strongly suppress the activity related to self-produced speech), the results nevertheless indicate that speech monitoring malfunctions in PD when compared to healthy controls. To what extent the observed effect explains speech deficits of PD remains to be determined. Subjectively, the patients reported not experiencing any speech impairment, although based on perceptual assessment (MDS-UPDRS speech item) the patients did show mild speech impairment.

We did not observe clear differences between speech volume with our psychophysical task. This could be due to various reasons. First, it is possible that the patients did not display speech hypophonia. Second, it is possible that the presence of an external cue stimulus (to which the participants were asked match their own speech volume) could have improved performance. In hindsight, an objective measurement of different speech characteristics using e.g. sustained phonation task would have been a better approach. However, the fact that PD patients with very mild speech impairment show relatively strong modulation in SIS is also a fascinating finding, as it suggests that changes in speech-production occur early in PD. To examine whether abnormal SIS is associated with specific aspects of speech deficit (e.g. hypophonia or altered articulation), a larger sample of patients with clear speech deficits needs to be collected.

Previously, Huang et al. (2016) reported that perturbed speech differently modulated auditory evoked activity in P2 time-window (200–400 ms after sound onset) in PD patients when compared to healthy controls. Our results show that PD patients display abnormal speech-related activity even when speech is not artificially perturbed, and that this effect is observed early (around 100 ms after phonation) in scalp recorded EEG. That the difference occurs in this early (N1) time-window is important because CD signals are assumed to modulate auditory-evoked activity at this specific time-window. In other words, models of speech production (e.g. Houde & Nagarajan, 2011) assume that self-produced sounds are compared (via CD signals) to predicted speech output during initial cortical processing stages, enabling reflexive modulation of speech (Niziolek, Nagarajan, & Houde, 2013). While the present investigation focused on auditory monitoring mechanisms, also somatosensory predictions are known to mediate speech (Niemi, Laaksonen, Ojala, Aaltonen, & Happonen, 2006; Tremblay, Shiller, & Ostry, 2003), and also they could be altered in PD.

The N1 time-window in which SIS is strongest in scalp recorded EEG is assumed to generated by a network of areas including primary and association auditory cortices in the superior temporal cortex and planum temporale. In addition it may be modulated by frontal cortical areas (Godey, Schwartz, De Graaf, Chauvel, & Liégeois-Chauvel, 2001; Näätänen & Picton, 1987; Zouridakis, Simos, & Papanicolaou, 1998). Because N1 is known to be modulated by various cortical sources, and especially frontal cortical areas are active during active phonation, we estimated auditory cortical activity in the Heschl's gyrus using source reconstruction. These results suggest that posterior superior temporal cortex is modulated by speech already before the onset of phonation (Fig. 3B). Interestingly, PD patients showed stronger activity than healthy controls during this pre-phonation time-window. This suggests that top-down modulation of auditory cortex activity is incorrectly timed in PD when compared to healthy controls. In addition, auditory cortex was activated more strongly in PD patients than controls during a later time-window (400-500 ms after phonation). Because these activations took place partly before phonation, it shows that the activity was not caused by bottom-up sensory information, but reflects top-down modulation from other brain areas such as frontal cortex.

Brain imaging studies have reported increased activation in frontal areas during speech in PD (Liotti, Ramig, Vogel, New, & Cook, 2003; Pinto et al., 2004). These activations have been associated with “compensatory mechanisms”. That is, they have not assumed to be related to the pathophysiology of speech deficits per se, but be a consequence of them: the patients’ attempts to compensate for their speech difficulties (e.g. attempts to speak louder) cause increased activations in frontal areas. In contrast to this prevailing view, we suggest that the increased activations may be part of the cause of speech deficits. Because the increased auditory cortical activation in PD was partly observed before phonation, the activity is not likely to reflect compensation of speech deficits. Specifically, we propose that the increased activation may reflect deficits in timing the communication of top-down predictions to auditory cortex. This deficit may lead to abnormal SIS after phonation.

According to common interpretation SIS is a signature of the mismatch between predicted and produced speech (i.e. prediction error). If this is correct, our results suggest that PD patients’ brain detect a higher than normal mismatch between desired and produced speech. This could be due to the fact that produced phonation does not match the intended phonation. If true, an

interesting question is why are PD patients often subjectively unaware of their speech deficit? Possibly the mismatch signal is not successfully communicated to speech production processes; in the terms of state control feedback model, possibly the current state of the system is not properly estimated, and this is why speech impairments emerge in PD. Possibly because the prediction error is not successfully taken into account by speech production mechanisms the patients may not notice their speech impairment. However, this is speculative and subsequent research should aim to disentangle if speech prediction errors are indeed not successfully assimilated by frontal speech production areas.

Because PD is associated with changes to multiple neurotransmitter systems (Jellinger, 1991), multiple different mechanisms could underlie the observed differences between PD patients and control participants. That said, dopamine deficiency has been proposed to lead to deficits in sensory attenuation (Friston et al., 2015, 2012). Previously, we observed that in PD patients performance in the double-saccade task – which is used to probe CD signals in the oculomotor system – correlated with changes in dopamine transported binding in striatum (Railo et al., 2018). PD is known to lead to altered activation in basal ganglia–cortex loops (Rivlin-Etzion et al., 2006; Rodriguez-Oroz et al., 2009). This could lead to abnormal activation in frontal speech production mechanisms, which in turn communicate CD signals to sensory cortical areas and receive speech prediction errors from sensory areas. Dopamine deficit has been suggested to alter activity in the beta band, which is assumed to mediate top-down predictions (Friston et al., 2015). There is lots of evidence for altered beta band activity in PD (Rivlin-Etzion et al., 2006; Rodriguez-Oroz et al., 2009), and possibly it could also contribute to speech deficits. Results by Söros et al. (2017) suggest that beta-band activity is increased during speech preparation in PD when compared to healthy controls.

Previously, patients with schizophrenia have been reported to reveal reduced SIS (Ford & Mathalon, 2005; Ford et al., 2013), but to our knowledge this is the first report of impaired SIS in PD. We have argued that altered SIS in PD may be due to aberrant timing between frontal cortical speech production areas and sensory cortical areas, and that this mechanism may cause speech deficits in PD.

Acknowledgements

Henry Railo was funded by Academy of Finland (grant #308533).

References

Beck, A. T. (1976). Cognitive therapy and the emotional disorders. *INT.UNIV.PRESS,NEW YORK*. https://doi.org/10.1176/appi.psychotherapy.1977.31.4.633

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 of the United States of America*. https://doi.org/10.1073/pnas.1216827110

Curio, G., Neuloh, G., Numminen, J., Jousmäki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*. https://doi.org/10.1002/(SICI)1097-0193(200004)9:4<183::AID-HBM1>3.0.CO;2-Z

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E.

(2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*. https://doi.org/10.1016/S0896-6273(00)81138-1

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*. https://doi.org/10.1016/j.jneumeth.2003.10.009

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2010.06.010

Follett, K. A., Weaver, F. M., Stern, M., Hur, K., Harris, C. L., Luo, P., … Reda, D. J. (2010). Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa0907083

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. https://doi.org/10.1016/0022-3956(75)90026-6

Ford, J. M., & Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? In *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2005.01.014

Ford, J. M., Mathalon, D. H., Roach, B. J., Keedy, S. K., Reilly, J. L., Gershon, E. S., & Sweeney, J. A. (2013). Neurophysiological evidence of corollary discharge function during vocalization in psychotic patients and their nonpsychotic first-degree relatives. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbs129

Fox, C. M., & Ramig, L. O. (1997). Vocal Sound Pressure Level and Self-Perception of Speech and Voice in Men and Women with Idiopathic Parkinson Disease. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/1058-0360.0602.85

Friston, K. J., Bastos, A. M., Pinotsis, D., & Litvak, V. (2015). LFP and oscillations-what do they tell us? *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2014.05.004

Friston, K. J., Shiner, T., FitzGerald, T., Galea, J. M., Adams, R., Brown, H., … Bestmann, S. (2012). Dopamine, affordance and active inference. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1002327

Godey, B., Schwartz, D., De Graaf, J. B., Chauvel, P., & Liégeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clinical Neurophysiology*. https://doi.org/10.1016/S1388-2457(01)00636-8

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., … Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2013.00267

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn3158

Ho, A. K., Bradshaw, J. L., & Iansek, R. (2008). For better or worse: The effect of Levodopa on speech in Parkinson's disease. *Movement Disorders*. https://doi.org/10.1002/mds.21899

Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S. (1998). Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology*.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography*. https://doi.org/10.1097/00004728-199803000-00032

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2011.00082

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/089892902760807140

Huang, X., Chen, X., Yan, N., Jones, J. A., Wang, E. Q., Chen, L., … Liu, H. (2016). The impact of parkinson's disease on the cortical mechanisms that support auditory–motor integration for voice control. *Human Brain Mapping*. https://doi.org/10.1002/hbm.23306

Jellinger, K. A. (1991). Pathology of Parkinson's disease - Changes other than the nigrostriatal pathway. *Molecular and Chemical Neuropathology*. https://doi.org/10.1007/BF03159935

Kiran, S., & Larson, C. R. (2001). Effect of duration of pitch-shifted feedback on vocal responses in patients with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2001/076)

Kwan, L. C., & Whitehill, T. L. (2011). Perception of speech by individuals with Parkinson's disease: A review. *Parkinson's Disease*. https://doi.org/10.4061/2011/389767

Leritz, E., Loftis, C., Crucian, G., Friedman, W., & Bowers, D. (2004). Self-awareness of deficits in Parkinson disease. *Clinical Neuropsychologist*. https://doi.org/10.1080/1385404049052412

Lin, F. H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2005.11.054

Liotti, M., Ramig, L. O., Vogel, D., New, P., & Cook, C. I. (2003). Hypophonia in Parkinson's disease Neural correlates of voice treatment revealed by PET. *Neurology*.

Liu, H., Wang, E. Q., Metman, L. V., & Larson, C. R. (2012). Vocal responses to perturbations in voice auditory feedback in individuals with parkinson's disease. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0033629

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders*. https://doi.org/10.1044/jshd.4301.47

Louis, E. D., Winfield, L., Fahn, S., & Ford, B. (2001). Speech dysfluency exacerbated by levodopa in Parkinson's disease. *Movement Disorders*. https://doi.org/10.1002/mds.1081

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mullen, T. R., Kothe, C. A. E., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., … Cauwenberghs, G. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomedical Engineering*. https://doi.org/10.1109/TBME.2015.2481482

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x

Niemi, M., Laaksonen, J. P., Ojala, S., Aaltonen, O., & Happonen, R. P. (2006). Effects of transitory lingual nerve impairment on speech: an acoustic study of sibilant sound /s. *Int J Oral Maxillofac Surg*.

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? evidence from speech production. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2137-13.2013

Pinto, S., Thobois, S., Costes, N., Le Bars, D., Benabid, A. L., Broussolle, E., … Gentil, M. (2004). Subthalamic nucleus stimulation and dysarthria in Parkinson's disease: A PET study. *Brain*. https://doi.org/10.1093/brain/awh074

Railo, H., Olkoniemi, H., Eeronheimo, E., Pääkkönen, O., Joutsa, J., & Kaasinen, V. (2018). Dopamine and eye movement control in Parkinson's disease: Deficits in corollary discharge

signals? *PeerJ*. https://doi.org/10.7717/peerj.6038

Rivlin-Etzion, M., Marmor, O., Heimer, G., Raz, A., Nini, A., & Bergman, H. (2006). Basal ganglia oscillations and pathophysiology of movement disorders. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2006.10.002

Rodriguez-Oroz, M. C., Jahanshahi, M., Krack, P., Litvan, I., Macias, R., Bezard, E., & Obeso, J. A. (2009). Initial clinical manifestations of Parkinson's disease: features and pathophysiological mechanisms. *The Lancet Neurology*. https://doi.org/10.1016/S1474-4422(09)70293-5

Sapir, S., Pawlas, A. A., Ramig, L. O., Countryman, S., O'Brien, C., Hoehn, M. M., & Thompson, L. A. (2001). Voice and Speech Abnormalities in Parkinson Disease: Relation to Severity of Motor Impairment, Duration of Disease, Medication, Depression, Gender, and Age. *Journal of Medical Speech-Language Pathology*.

Skodda, S., Visser, W., & Schlegel, U. (2010). Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease. *Journal of Neural Transmission*. https://doi.org/10.1007/s00702-009-0351-5

Stenroos, M., Hunold, A., & Haueisen, J. (2014). Comparison of three-shell and simplified volume conductor models in magnetoencephalography. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2014.01.006

Stropahl, M., Bauer, A. K. R., Debener, S., & Bleichner, M. G. (2018). Source-Modeling auditory processes of EEG data using EEGLAB and brainstorm. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2018.00309

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2011/879716

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

Zouridakis, G., Simos, P. G., & Papanicolaou, A. C. (1998). Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. *Brain Topography*. https://doi.org/10.1023/A:1022246825461