Recognition of emotional prosody is altered after subthalamic nucleus deep brain stimulation in Parkinson’s disease

Julie Péron a, b, c, Didier Grandjean c, Florence Le Jeune b, d, Paul Sauleau a, b, Claire Haegelen b, e, Dominique Drapier b, f, Tiphaine Rouaud a, b, Sophie Drapier a, b, Marc Vérin a, b

4 Clinique Neurologique, Hôpital Pontchaillou, CHU de Rennes, rue Henri Le Guillox, 35033 Rennes, France
b URE-EM 425 Comportement et Nouveaux Crus Centraux, Université Rennes 1, Hôpital Pontchaillou, CHU de Rennes, rue Henri Le Guillox, 35033 Rennes, France
c Neuroscience of Emotion and Affective Dynamics, Department of Psychology, University of Geneva, 40 bd du Pont d’Arve, 1205 Geneva - CH, Switzerland
d Service de Médecine Nucléaire, Centre Eugène Marquis, rue de la Bataille Flandres Dunkerque, 35042 Rennes, France
e Service de Neurochirurgie, Hôpital Pontchaillou, CHU de Rennes, rue Henri Le Guillox, 35033 Rennes, France
f URU-EM 425 Comportement et Noyaux Gris Centraux, Université Rennes 1, Hôpital Pontchaillou, CHU de Rennes, rue Henri Le Guilloux, 35033 Rennes, France

A R T I C L E   I N F O

Article history:
Received 3 February 2009
Received in revised form 20 November 2009
Accepted 3 December 2009
Available online 11 December 2009

Keywords:
Emotion
Prosody
Deep brain stimulation
Parkinson’s disease

A B S T R A C T

The recognition of facial emotions is impaired following subthalamic nucleus (STN) deep brain stimulation (DBS) in Parkinson’s disease (PD). These changes have been linked to a disturbance in the STN’s limbic territory, which is thought to be involved in emotional processing. This was confirmed by a recent PET study where these emotional modifications were correlated with changes in glucose metabolism in different brain regions, including the amygdala and the orbitofrontal regions that are well known for their involvement in emotional processing. Nevertheless, the question as to whether these emotional changes induced by STN DBS in PD are modality-specific has yet to be answered. The objective of this study was therefore to examine the effects of STN DBS in PD on the recognition of emotional prosody.

An original emotional prosody paradigm was administered to twenty-one post-operative PD patients, twenty-one pre-operative PD patients and twenty-one matched controls. Results showed that both the pre- and post-operative groups differed from the healthy controls. There was also a significant difference between the pre and post groups. More specifically, an analysis of their continuous judgments revealed that the performance of the post-operative group compared with that of the other two groups was characterized by a systematic emotional bias whereby they perceived emotions more strongly. These results suggest that the impaired recognition of emotions may not be specific to the visual modality but may also be present when emotions are expressed through the human voice, implying the involvement of the STN in the brain network underlying the recognition of emotional prosody.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

High-frequency deep brain stimulation (DBS) of the subthalamic nucleus (STN) constitutes a therapeutic advance for severely disabled patients with Parkinson’s disease (PD), in whom long-term pharmacological treatment has failed. Although there is growing evidence of the beneficial effects of chronic DBS on PD motor symptoms, several clinical studies have reported cognitive (For a review, see Parsons, Rogers, Braaten, Woods, & Troster, 2006), behavioural and emotional impairments (For a review, see Temel et al., 2006) associated with STN DBS in PD.

Clinical studies have revealed that the recognition of facial expressions of emotions is impaired following STN DBS in PD (Biseul et al., 2005; Drapier et al., 2008; Dujardin et al., 2004; Le Jeune et al., 2008; Péron et al., in press; Schroeder et al., 2004). Although the materials and procedure differed across studies, it has been argued that one consistent finding is that the impairment seems to selectively concern negative emotions. Anger recognition was found to be selectively impaired in Schroeder et al.’s study (2004), recognition of three negative emotions (anger, disgust, sadness) was impaired in Dujardin et al.’s study (Dujardin et al., 2004), fear recognition was selectively impaired in Biseul et al.’s study (Biseul et al., 2005), and fear and sadness were impaired in Drapier et al. and Péron et al.’s studies (Drapier et al., 2008; Péron et al., in press). Nevertheless, the selective nature of the disruption of
negative emotions following STN DBS in PD remains under debate. For instance, the design of the above-mentioned studies appears to have been biased towards negative emotions. While they always had several negative stimuli, they only had one positive stimulus (happiness), so there was a greater likelihood of missing a negative scale. Another important finding is that the impairment cannot be attributed to secondary variables, such as anxiety or depression, a visuospatial deficit or general cognitive decline. These results suggest that the STN may be part of a broadly distributed neural network involved in the recognition of facial expressions, either via computation within the STN itself or by virtue of its impact on other brain regions involved in emotional processing, such as the orbitofrontal cortices and the amygdala (for a review, see e.g., Adolphs, Damasio, Tranel, & Damasio, 1996).

STN involvement in emotional processing has also been suggested by intraoperative studies, such as that conducted by Kuhn et al. (2005), which showed a differential and dynamic modulation of STN potentials in response to the presentation of affective pictures. However, Kuhn et al.’s result should be not be specifically applied to some parameters, as such as the visual stimulus features, were not properly controlled in the study and could therefore account for this result (Delplanque, N’Diaye, Scherer, & Grandjean, 2007).

PET studies have also been conducted in order to test the hypothesis of the involvement of the STN in the recognition of facial expressions (Geday, Ostergaard, & Gjedde, 2006; Le Jeune et al., 2008). In one of them (Le Jeune et al., 2008), our group recently obtained a positive correlation between a reduction in the recognition of fearful facial expressions following STN DBS and changes in glucose metabolism, especially in the right orbitofrontal cortex, using 18F-FDG-PET. These results suggest that the STN is a key basal ganglia structure in brain circuits involved in emotional processing.

Nevertheless, this research has left several questions unanswered and it has yet to be determined whether these emotional changes induced by STN DBS in PD are modality-specific or supramodal. Whereas previous studies have focused on the influence of STN stimulation on the recognition of emotional facial expressions, the present study examined the effects of STN stimulation on the recognition of emotional expressions conveyed by the human voice i.e., the recognition of emotional prosody.

Emotional prosody is defined as modifications in segmental and supra-segmental speech parameters during an emotional episode (Grandjean, Banziger, & Scherer, 2006). Physiological modifications in the vocal tract and vocal folds during an emotional episode, such as modifications in the tension of the vocal folds related to vagus nerve activity or in the tension of the soft palate, induce modulations in different acoustic components of the human voice (Ghazanfar & Rendall, 2008; Grandjean et al., 2006). These voice modulations can be studied and described in terms of modifications in acoustic parameters such as fundamental frequency, amplitude and distribution of energy across different frequency bands. These objective acoustic measures correspond, at the level of perception, to the modulations in intonation (i.e. pitch), loudness and voice quality perceived by the listener. Investigations of the perception of emotional prosody through fMRI and patient studies have allowed researchers to delineate a distributed neural network involved in the identification and recognition of emotional prosody. In addition to primary and secondary auditory regions, modulation of neuronal activity within the superior temporal sulcus and gyrus has also been reported in response to exposure to emotional prosodic stimuli (Ethofer, Anders, Wiethoff, et al., 2006; Grandjean et al., 2005; Sander et al., 2005). An increase in activity within the amygdala has also been observed, not only in response to emotional prosody (Grandjean et al., 2005; Sander et al., 2005) but also in response to emotional animal vocalizations (Fectue, Belin, Joanette, & Armony, 2007). Modulations in activity within anterior regions, such as the orbitofrontal cortex and inferior frontal areas, have been described, especially when participants pay attention to an auditory stimulus or are asked to identify emotional prosody (Ethofer, Anders, Erb, et al., 2006; Sander et al., 2005; Wildgruber et al., 2004). Recently, a study of neglect patients confirmed that these regions – the superior temporal sulcus and gyrus and the orbitofrontal cortex – are important for detecting emotional voices in the environment (Grandjean, Sander, Lucas, Scherer, & Vuilleumier, 2008). In addition to these regions, the involvement of the basal ganglia in the processing of emotional prosody, particularly the caudate nucleus and putamen, has also been revealed by fMRI and patient studies (Bach et al., 2008; Grandjean et al., 2005; Kotz et al., 2003; Morris, Scott, & Dolan, 1999).

These findings highlight the overlap between the neuroanatomical substrates of the recognition of emotional prosody and the neuroanatomical circuit that STN DBS is thought to modify, particularly the amygdala and the orbitofrontal cortex regions (Le Jeune et al., 2008). Accordingly, in the present study, we hypothesized that the impaired recognition of emotions in PD patients following STN DBS is not specific to the visual sensory modality but is also present when emotions are expressed through the human voice.

To test this hypothesis, we explored the recognition of emotions in the vocal modality, using an original emotional prosody paradigm, in twenty-one PD patients in a pre-operative condition, twenty-one PD patients in a post-operative condition, and twenty-one matched healthy controls (HC).

2. Participants and methods

2.1. Participants (Table 1)

Two groups of patients with PD (pre-operative and post-operative groups) and a HC group took part in the study.

All patients met the clinical criteria of the United Kingdom Parkinson’s Disease Society Brain Bank for Idiopathic PD (Hughes, Daniel, Kilford, & Lees, 1992). The first patient group included 21 patients (10 men, 11 women) with advanced PD who were candidates for STN DBS (pre-operative group), while the second included 21 patients (10 men, 11 women) who had already undergone bilateral STN DBS (35 ± 29.7 months post-operation; range 3–72 months) at Rennes University Hospital, France. Standard selection criteria for surgery were applied to all patients (Welter et al., 2002).

The two patient groups were comparable for disease duration and cognitive functions, as well as for dopamine replacement therapy, calculated on the basis of correspondences adapted from Looza et al. (1995). The characteristics of the two patient groups and the HC group are presented in Table 1.

The HC group consisted of 21 healthy individuals (10 men and 11 women) who had no history of neurological disease or alcohol abuse and no signs of dementia, as attested by their scores on the MMSE (Dérouesné, 2001).

All three groups were matched for gender, age and education level (see Table 1). The study was approved by the Ethics Committee of Rennes University Hospital. After a complete description of the study, written informed consent was obtained from each participant, and the study was conducted in accordance with the Declaration of Helsinki.

2.2. Methods

All the PD patients (pre-operative and post-operative) were assessed using motor, neuropsychological and emotional evaluations. All the PD patients were receiving their normal dopamine replacement therapy (i.e. were “on-dopa”) when they underwent the neuropsychological and emotional assessments. The post-operative patients were on-dopa and on-stimulation in the post-operative condition.

2.2.1. Neurosurgery (post-operative group)

2.2.1.1. Methodology. Quadrupolar deep brain stimulation electrodes (3389 Medtronic, Minneapolis, MN, USA) were implanted bilaterally in the STN. The overall methodology was similar to that previously described by Benabid et al. (2000).

2.2.1.2. Electrode location. The location of the unipolar chronic electrode contacts at M3 was determined using the technique we have previously reported (Sauleau et al., 2005). The focus of each stimulation contact was located in relation to the middle of the bicommissural line (AC–PC), by superimposing the electrode positioning picture on the corresponding ventriculogram. Distances were measured on a squared transparent sheet, and then readjusted using a computerised spreadsheet. In all
patients, chronic stimulation was monopolar, using a single contact of the quadripolar electrode. The stimulation characteristics were as follows: mean (±S.D.) electrical variables 2.6 (±0.6) for voltage (volts), 65.7 (±12.0) for pulse width (microseconds) and 140.7 (±17.4) for frequency (hertz) on the right side, and 2.5 (±0.4) for voltage, 64.2 (±10.7) for pulse width and 141.4 (±17.2) for frequency on the left side.

2.2.2. Motor assessment

All the PD patients (pre-operative and post-operative) were evaluated according to the Core Assessment Program for Intracerebral Transplantation (Langston et al., 1992) and were scored on the Unified Parkinson’s Disease Rating Scale motor part III (UPDRS III) (Fahn & Elton, 1987), the Hoehn and Yahr (Hoehn & Yahr, 1967) and Schwab and England (Schwab & England, 1969) scales. All the PD patients were assessed on- and off-dopa. For the post-operative group, stimulation remained on after surgery.

2.2.3. Neuropsychological and psychiatric assessments

A short neuropsychological battery was administered to all participants prior to the vocal emotion recognition sessions. This battery included the Mattis scale (Mattis, 1988) and a series of tests assessing frontal executive functions: Nelson’s modified version of the Wisconsin Card Sorting Test (MCST) (Nelson, 1976), the Trail Making Test (Reitan, 1958), the Categorical and Literal Fluency Test (Cardébat, Doyon, Puel, Goulet, & Joanette, 1990), the Action Verbs fluency task (Woods et al., 2005), and the Stroop Test (Stroop, 1935). Depression was assessed using the Montgomery–Asberg Depression Rating Scale (MADRS, Montgomery & Asberg, 1979). The MADRS was chosen because of the predominance of psychic items over happiness and sadness, together with a neutral condition, were used in the present study. To ensure that participants had normal hearing, they were assessed by means of a standard audiometric screening procedure (AT-1-B audiometric test) to measure tonal and vocal sensitivity. This was carried out in the ENT Department of Reennes University Hospital. None of the patients included in the study were hearing aids or had a history of tinnitus or a hearing impairment.

2.2.4. Vocal emotion recognition

A set of vocal stimuli consisting of meaningless speech (pseudowords) was played to all participants.

2.2.4.1. Pseudoword stimuli. The vocal stimuli were extracted from the database developed by Banse and Scherer (1996) and validated in their study. They consisted of short segments of meaningless speech (pseudowords), obtained by concatenating different syllables found in Indo-European languages so that they would be perceived as natural utterances, with emotional intonation (across different cultures) but no semantic content. Four different categories of emotional prosody (anger, fear, happiness and sadness), together with a neutral condition, were used in the present study. The mean duration of the stimuli was 2044 ms (range: 1205–5236 ms). An ANOVA test failed to reveal any significant difference in duration between the different prosodic categories (neutral, anger, fear, happiness and sadness, F(4,156) = 1.43, p > .10); and there was no significant difference between the mean acoustic energy expended, F(4,156) = 1.86, p > .10 (none of the systematic pairwise comparisons between the neutral condition and the emotional prosodies was significant). Likewise, there was no significant difference between categories for the standard deviation of the mean energy of the sounds, F(4,156) = 1.9, p > .10.

We selected utterances produced by 12 different actors (6 women and 6 men), each expressed by five different prosodies (anger, fear, happiness, neutral and sadness). The set of vocal stimuli (pseudowords) comprised 60 stimuli: 12 actors × 5 emotion conditions.

2.2.4.2. Vocal emotion recognition procedure. All stimuli were presented bilaterally through stereo headphones using an Authorware programme designed specially for this study. Participants sat comfortably in a quiet room, in front of the computer, and looked at a fixation cross while listening to the stimuli. They were told that they would hear meaningless speech uttered by male and female actors and that these actors would express emotions through their utterances. Participants were required to listen to each stimulus, after which they were asked to judge the emotional content of each stimulus using a set of visual analogue scales displayed simultaneously on the computer screen. More specifically, participants were instructed to judge the extent to which the different emotions were expressed on a visual analogue scale ranging from “not at all” to “very much”. Participants rated six scales: one scale for each emotion presented (anger, fear, happiness and sadness) and one for the neutral utterance, plus a scale to rate the “surprise” emotion, in order to see whether the fear emotion expressed by the human voice was confused with surprise, as is the case with facial expressions (Ekman, 2003; Scherer & Ellgring, 2007). Participants were told that they could listen again to each stimulus as many as six times, by clicking on a button on the computer interface (click count). Participants were presented with two examples in order to familiarize themselves with the task. An example of the computer interface used for the recognition of emotional prosody task is provided in Appendix A.

In order to avoid a list effect between the pre-operative and post-operative groups, a second equivalent version of the recognition of emotions from voices test was constructed and the two versions were counterbalanced. In the pre-operative group, half the PD patients were assessed with version “A” of the recognition of emotions from voices test and the other half with version “B”. The same procedure was applied to the post-operative group. Consequently, half the HC were assessed with version “A” and the other half with version “B”.

The entire protocol was completed within a single session lasting approximately 90 min.

2.2.4.3. Audiometric screening procedure. To ensure that participants had normal hearing, they were assessed by means of a standard audiometric screening procedure (AT-1-B audiometric test) to measure tonal and vocal sensitivity. This was carried out in the ENT Department of Reennes University Hospital. None of the patients included in the study wore hearing aids or had a history of tinnitus or a hearing impairment.

2.2.5. Statistical analysis

For the sociodemographic, neuropsychological and psychiatric data, comparisons between the three groups (HC, pre-operative PD patient group and post-operative PD patient group) were performed using a single-factor analysis of variance (ANOVA). Whenever the ANOVA test yielded a significant difference, pairwise t-tests for two independent groups were carried out to determine which groups differed from one another. Moreover, comparisons of the two PD patient groups on clinical variables were conducted using t-tests for two independent groups. Within-group comparisons were conducted in the post-operative group, in order to compare patients’ ‘off-dopa–off-stim’ score vs. their ‘off-dopa–on-stim’ score on the UPDRS III using t-tests for dependent groups.

For the vocal emotion recognition data, we performed two levels of analysis. First, we compared the three groups’ performances on categorical judgements according to the percentages of correct responses using the chi-square ($\chi^2$) test. Second, we compared the three groups’ performances on the continuous rating scales for each prosodic category (plus neutral) and for each scale. To do so, we conducted a repeated-measures ANOVA with two within-participants factors – emotion (5 levels) and scale (6 levels) – and one between-participants factor – group (pre, post, HC; 3 levels). We then performed a repeated-measures ANOVA with one within-participants factor – scale (6 levels) – and one between-participants factor – group (pre, post, HC) for each type of prosody. If the results of the latter were significant, and in order to investigate the effects in greater detail, contrasts were performed between the three groups (‘pre vs. post’, ‘pre vs. HC’ and ‘post vs. HC’) for each prosodic category and each rating scale.
Table 2

**Summary of results of statistical analyses.**

1. Sociodemographic variables: Single-factor analysis of variance (ANOVA) for the three groups, i.e. pre, post, and HC; if significant, pairwise t-tests for neuropsychological and psychiatric data.

2. Vocal emotion recognition data
   - 2.1 Categorical judgments
     - 2.1.1 Chi-square ($\chi^2$) for overall performances for the three groups (pre, post, HC); if significant, pairwise comparisons ('pre vs. post', 'pre vs. HC', and 'post vs. HC').
     - 2.1.2 Chi-square ($\chi^2$) for each emotion separately for the three groups (pre, post, HC); if significant, pairwise comparisons ('pre vs. post', 'pre vs. HC', and 'post vs. HC').
   - 2.2 Continuous judgments (Table 5)
     - 2.2.1 Overall performances: Repeated-measures ANOVA with two within-participants factors – emotion (5 levels) and scale (6 levels) – and one between-participants factor: group (pre, post, and HC; 3 levels).
     - 2.2.2 Each emotion separately: Repeated-measures ANOVA with one within-participants factor – emotion (5 levels); if significant, pairwise comparisons for each separate prosodic category.
   - 2.3 Click count: Repeated-measures ANOVA with one within-participants factor – emotion (5 levels) – and one between-participants factor – group (3 levels).
   - 2.4 Version A vs. Version B: Chi-square ($\chi^2$).

3. Correlations between neuropsychological, sociodemographic and clinical variables and recognition of emotions from voices variables.

Additionally, we compared the three groups' performances on the number of clicks for all prosodies (plus neutral) using a repeated-measures ANOVA with one within-participants factor – emotion (5 levels) – and one between-participants factor – group (3 levels). The 'click count' variable refers to the number of times that the participants listened again to each stimulus by clicking on a button on the computer interface. This variable can be regarded as an indicator of the judgment difficulty of each item.

We performed an analysis in order to compare the two parallel versions (A and B) of the emotional prosody recognition task using the chi-square ($\chi^2$) test. Correlations were assessed using the Bravais–Pearson correlation coefficient between neuropsychological, sociodemographic and clinical variables and recognition of emotions from voices variables.

A statistical analysis was performed using Statistica 8.0. The significance threshold was set at $p < .05$, except for Bravais–Pearson correlations where the $p$-value was adjusted for multiple comparisons.

Table 2 summarizes the statistical methodology.

3. Results

3.1 Motor, neuropsychological and psychiatric assessments (Tables 1 and 3)

The analysis failed to reveal any significant difference between the two PD patient groups for any of the variables of the motor assessment, except for the Schwab & England score in the "On" condition, and there was no significant difference between the three groups for any of the variables of the neuropsychological background assessment. Interestingly, there was a significant difference between the UPDRS III 'off-dopa–off-stim' and 'off-dopa–on-stim' scores in the post-operative PD patient group, attesting to the motor improvement induced by the surgery ($t = 8.86 (20), p = .0001$).

There was a significant difference between the three groups on the MADRS score. Pairwise comparisons revealed a significant difference between the pre-operative group and HC ($t = 2.16(40), p = .04$) and between the post-operative group and HC ($t = 3.38(40), p = .002$), but none between the pre and post groups ($t = 0.16(40), p = .99$). As shown in Table 2, the pre- and post-operative PD groups scored significantly higher on MADRS than the HC group, meaning that PD patients were more depressed than HC (Table 3).

3.2 Vocal emotion recognition

3.2.1 Categorical judgments (Table 4)

A response was deemed to be correct when the subject rated the “Target” scale (e.g. the "Anger" scale when the stimulus was “Anger”) higher than all the other scales.

There was a significant difference between the three groups on the distribution of the overall score, $\chi^2(2) = 6.34, p = .04$. Pairwise comparisons of overall scores revealed that the pre-operative PD patient group performed more poorly than the HC group, $\chi^2(1) = 6.21, p = .01$. There was no difference between the pre and post groups, $\chi^2(1) = 0.85, p = .1$, or between the post and HC groups, $\chi^2(1) = 2.47, p = .12$.

Additional analyses were performed in order to compare the three groups performances for each separate prosodic category. This analysis did not reveal any significant difference for any of the comparisons (Anger: $\chi^2(2) = 3.95, p = .14$; Fear: $\chi^2(2) = 2.62, p = .27$; Happiness: $\chi^2(2) = 2.05, p = .36$; Neutral: $\chi^2(2) = .48, p = .79$; Sadness: $\chi^2(2) = 5.60, p = .06$).

3.2.2 Continuous judgments (Table 5)

Overall, analysis revealed a main effect of Emotion, $F(4,240) = 38.93, p < .00001$, an effect of Group, $F(2,60) = 7.25, p = .001$, and, more interestingly, an interaction between the Group $\times$ Emotion $\times$ Scale factors, $F(4,1200) = 1.71, p = .003$ showing that the pre, post, and HC displayed different patterns of responses on the different scales and different emotions.
### Table 3
Neuropsychological background data for the two groups of PD patients and the HC group.

<table>
<thead>
<tr>
<th></th>
<th>Pre-op. (n=21)</th>
<th>Post-op. (n=21)</th>
<th>HC (n=21)</th>
<th>Stat. val. (F)</th>
<th>df</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MMSE</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mattis (/144)</strong></td>
<td>141.1</td>
<td>2.3</td>
<td>139.9</td>
<td>2.8</td>
<td>140.9</td>
<td>2.0</td>
</tr>
<tr>
<td><strong>Stroop</strong></td>
<td>3.8</td>
<td>10.2</td>
<td>2.1</td>
<td>8.6</td>
<td>6.8</td>
<td>9.2</td>
</tr>
<tr>
<td><strong>TMT</strong></td>
<td>42.5</td>
<td>13.9</td>
<td>50.0</td>
<td>20.6</td>
<td>42.7</td>
<td>15.1</td>
</tr>
<tr>
<td><strong>Verbal fluency</strong></td>
<td>34.8</td>
<td>9.4</td>
<td>29.0</td>
<td>12.3</td>
<td>32.0</td>
<td>9.0</td>
</tr>
<tr>
<td><strong>Categorical</strong></td>
<td>24.7</td>
<td>7.2</td>
<td>21.0</td>
<td>7.4</td>
<td>20.8</td>
<td>6.1</td>
</tr>
<tr>
<td><strong>Phonemic</strong></td>
<td>16.7</td>
<td>5.8</td>
<td>14.9</td>
<td>5.3</td>
<td>17.8</td>
<td>6.5</td>
</tr>
<tr>
<td><strong>Action verbs</strong></td>
<td>34.8</td>
<td>9.4</td>
<td>29.0</td>
<td>12.3</td>
<td>32.0</td>
<td>9.0</td>
</tr>
<tr>
<td><strong>MCST</strong></td>
<td>45.21</td>
<td>26.85</td>
<td>50.7</td>
<td>27.3</td>
<td>56.3</td>
<td>19.17</td>
</tr>
<tr>
<td><strong>Errors</strong></td>
<td>3.95</td>
<td>2.62</td>
<td>3.96</td>
<td>2.62</td>
<td>3.96</td>
<td>2.62</td>
</tr>
<tr>
<td><strong>Perseverations</strong></td>
<td>1.0</td>
<td>1.6</td>
<td>1.6</td>
<td>3.0</td>
<td>0.4</td>
<td>2.60 .64</td>
</tr>
<tr>
<td><strong>MADRS</strong></td>
<td>5.7</td>
<td>8.1</td>
<td>5.6</td>
<td>8.0</td>
<td>4.9</td>
<td>2.06</td>
</tr>
</tbody>
</table>

*Statistical values (stat. val.), the degree of freedom (df) and p-values between pre, post and HC groups are reported (single-factor ANOVA).

MMSE = Mini Mental State Examination; TMT = Trail Making Test; MCST = modified version of the Wisconsin Card Sorting Test; MADRS = Montgomery–Asberg Depression Rating Scale.

* Significant if p-value less than 0.05.

### Table 4
Percentage of correct responses (and standard deviation (S.D.)) on categorical judgments in the emotional prosody task for the two groups of PD patients and the HC group.

<table>
<thead>
<tr>
<th></th>
<th>Pre-op. (n=21)</th>
<th>Post-op. (n=21)</th>
<th>HC (n=21)</th>
<th>Stat. val. (χ²)</th>
<th>df</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anger</strong></td>
<td>69.05</td>
<td>12.25</td>
<td>70.63</td>
<td>14.82</td>
<td>76.59</td>
<td>12.25</td>
</tr>
<tr>
<td><strong>Fear</strong></td>
<td>49.21</td>
<td>8.72</td>
<td>53.57</td>
<td>11.95</td>
<td>56.35</td>
<td>19.17</td>
</tr>
<tr>
<td><strong>Happiness</strong></td>
<td>45.24</td>
<td>16.99</td>
<td>42.06</td>
<td>17.77</td>
<td>48.41</td>
<td>13.85</td>
</tr>
<tr>
<td><strong>Neutral</strong></td>
<td>63.09</td>
<td>15.93</td>
<td>60.32</td>
<td>25.13</td>
<td>60.71</td>
<td>17.71</td>
</tr>
<tr>
<td><strong>Sadness</strong></td>
<td>43.65</td>
<td>16.65</td>
<td>52.78</td>
<td>20.80</td>
<td>52.78</td>
<td>15.44</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>54.05*</td>
<td>7.63</td>
<td>55.87</td>
<td>9.70</td>
<td>58.97</td>
<td>8.99</td>
</tr>
</tbody>
</table>

*Statistical values (stat. val.), the degree of freedom (df) and p-values between pre, post and HC groups are reported (χ²).

HC = healthy controls.

* p = 0.01 in comparison with HC.

### Table 5
Means (and standard deviations (S.D.)) of continuous judgment in the emotional prosody task for the two groups of PD patients (pre-operative and post-operative) and the HC group.

<table>
<thead>
<tr>
<th></th>
<th>Pre-operative (n=21)</th>
<th>Post-operative (n=21)</th>
<th>HC (n=21)</th>
<th>Stat. val. (F)</th>
<th>df</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Happiness scale</strong></td>
<td>0.90</td>
<td>12.14</td>
<td>6.64</td>
<td>42.50*</td>
<td>33.29</td>
<td>5.73</td>
</tr>
<tr>
<td><strong>Fear scale</strong></td>
<td>1.84</td>
<td>30.76**</td>
<td>17.71</td>
<td>5.70</td>
<td>16.45</td>
<td>6.17</td>
</tr>
<tr>
<td><strong>Sadness scale</strong></td>
<td>26.88</td>
<td>21.31</td>
<td>16.46</td>
<td>5.03</td>
<td>14.67</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Anger scale</strong></td>
<td>4.80</td>
<td>21.14</td>
<td>14.64</td>
<td>5.03</td>
<td>14.67</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Neutral scale</strong></td>
<td>0.53</td>
<td>21.31</td>
<td>16.46</td>
<td>5.03</td>
<td>14.67</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Surprise scale</strong></td>
<td>5.40*</td>
<td>21.31</td>
<td>16.46</td>
<td>5.03</td>
<td>14.67</td>
<td>4.00</td>
</tr>
</tbody>
</table>

*Statistical values (stat. val.), the degree of freedom (df) and p-values between pre, post and HC groups are reported (F).

HC = healthy controls.

* Significant if p-value less than 0.05 and **significant if p-value less than 0.01 in comparison to healthy controls (HC).

† Significant if p-value less than 0.05 and †† significant if p-value less than 0.01 in comparison to the pre-operative group.

† Significant if p-value less than 0.05 and †† significant if p-value less than 0.01 in comparison to the post-operative group.
In order to investigate these effects in greater detail, Group × Scale interaction analyses were performed for each separate emotion. If the results were significant, the contrasts were performed between the three groups for each prosodic category and (a) the target scale, i.e. the value on the target scale (e.g. Anger scale) corresponding to the participant’s rating of the relevant stimulus (e.g. Anger stimulus), and (b) the non-target scales, i.e. the values on the scales that did not correspond to the stimulus emotion (e.g. Fear scale for the Anger stimulus). Contrasts were performed for the “non-target” scales in order to investigate the patterns of confusion between the different emotions.

3.2.2.1. Anger. For Anger prosody, analyses revealed an interaction between the Group × Scale factors, $F(10, 300) = 2.47, p = .007$.

For the Anger scale when the stimulus was Anger, contrasts showed a significant difference between the pre and HC groups, $F(1,60) = 4.51, p = .03$, and the post and pre groups, $F(1,60) = 7.05, p = .01$. No difference was found between the post and HC groups, $F<1$.

When the Stimulus was Anger and the scale Fear (Fig. 1A), contrasts showed significant differences between the post and HC groups, $F(1,60) = 7.46, p = .0083$, and the post and pre groups, $F(1,60) = 7.96, p = .006$, but not between the pre and HC groups, $F<1$. There were no differences when the scales were Happiness, Neutral or Sadness ($F<1$ for all contrasts). When the scale was Surprise, contrasts revealed a difference between the post and HC groups, $F(1,60) = 4.05, p = .049$, but no difference between the pre and HC groups, $F<1$.

3.2.2.2. Fear. For Fear prosody, analyses revealed an interaction between the Group × Scale factors, $F(10, 300) = 2.02, p = .03$.

For the Fear scale when the stimulus was Fear, contrasts revealed a significant difference between the post and pre groups, $F(1,60) = 7.46, p = .008$, and the pre and HC groups, $F(1,60) = 4.05, p = .049$. There was no difference between the post and HC groups, $F<1$.

When the Stimulus was Fear and the scale Happiness (Fig. 1B), contrasts showed a difference between the post and HC groups, $F(1,60) = 10.32, p = .002$, and the post and pre groups, $F(1,60) = 5.81, p = .019$, but no difference between the pre and HC groups, $F<1$. There was no difference when the scales were Sadness, Anger or Neutral ($F<1$ for all contrasts). When the scale was Surprise, contrasts revealed a difference between the post and HC groups, $F(1,60) = 5.38, p = .024$, and the post and pre groups, $F(1,60) = 9.00, p = .0039$. There was no difference between the pre and HC groups, $F<1$ (Fig. 2).

3.2.2.3. Happiness. For Happiness prosody, there was no interaction effect between the Group × Scale factors, $F<1$.

3.2.2.4. Sadness. For Sadness prosody, analyses revealed an interaction between the Group × Scale factors, $F(10, 300) = 1.94, p = .04$.

For the Sadness scale when the stimulus was Sadness, contrasts revealed a significant difference between the post and pre groups, $F(1,60) = 9.95, p = .002$, and a marginal difference between the pre and HC groups, $F(1,60) = 3.53, p = .065$. There was no difference between the post and HC groups, $F(1,60) = 1.62, p = .21$.

When the Stimulus was Sadness and the scale Happiness, contrasts revealed a difference between the post and HC groups, $F(1,60) = 4.26, p = .043$, and the post and pre groups, $F(1,60) = 5.27, p = .025$, but no effect between the pre and HC groups, $F<1$. When the scale was Fear, contrasts revealed a significant difference between the post and HC groups, $F(1,60) = 4.34, p = .041$, but no difference between the post and pre groups, $F(1,60) = 1.49, p = .23$, or the pre and HC groups, $F<1$. There was no difference when the scales were Neutral or Anger ($F<1$ for all contrasts). When the scale was Surprise, there were no significant differences (post vs. HC, $F(1,60) = 1.17, p = .28$, post vs. pre, $F(1,60) = 1.94, p = .17$, pre vs. HC, $F<1$).

3.2.2.5. Neutral. For Neutral prosody, there was no interaction effect between the Group × Scale factors, $F<1$.

3.2.3. Click count analysis

This analysis revealed a main effect of Emotion, $F(4,180) = 3.74, p = .006$, but no main effect of Group and no interaction (click count for Anger: HC = 180, pre = 105, post = 135; for Fear: HC = 204, pre = 104, post = 165; for Happiness: HC = 221, pre = 118, post = 137;
for Neutral: HC = 160, pre = 110, post = 115; for Sadness: HC = 149, pre = 94, post = 98; Total: HC = 914, pre = 531, post = 650).

3.2.4. Version A vs. Version B comparisons
No significant difference was found between the percentages of correct responses for version “A” vs. version “B” ($\chi^2(5) = 5.63, p = .3$).

3.3. Correlations
No significant correlation was observed between performances on recognition of emotions from voices and either the tests included in the general neuropsychological battery or age, level of education or disease duration, for either the post-operative PD group or the pre-operative PD group. It should be noted that there was no significant correlation between recognition of emotions from voices performances and MADRS scores.

4. Discussion
The aim of the present study was to explore the recognition of emotions in the vocal modality (i.e. emotional prosody) in PD following STN DBS, in order to measure the extent to which the disruption in emotion recognition observed after this type of neurosurgery is specific to the visual modality or are also present in the auditory modality. We explored performances on the recognition of emotions from voices by comparing the ratings given by a pre-operative, a post-operative and an HC group ($n=21$ in each group) using an original emotional prosody recognition paradigm.

The participants’ responses were investigated using two complementary methods adapted to the experimental paradigm. First, we compared the three groups’ performances on categorical judgments in terms of percentages of correct responses. Second, we compared the three groups’ performances on continuous judgments for each type of prosody (plus a neutral condition), on the basis of (a) target scales and (b) non-target scales.

At the first level of analysis, an investigation of categorical judgments (i.e. percentages of correct responses) showed a significant difference between the three groups on the distribution of the overall score. This effect was mainly driven by the performances of the pre-operative group, who performed more poorly than the HC group (Table 4). There was no difference between either the pre and post groups or the post and HC groups. The result obtained for the pre-operative group replicated previous findings reported in the emotional prosody literature. To date, studies have consistently found that adults with mild to moderate PD display impaired recognition of the emotional meaning of prosodic cues in speech when compared with matched HC (Ariatti, Benuzzi, & Nichelli, 2008; Benke, Bosch, & Andree, 1998; Blonder, Gur, & Gur, 1989; Borod et al., 1990; Breitenstein, Van Lancker, Daum, & Waters, 2001; Caekebeke, Jennekens-Schinkel, van der Linden, Buruma, & Roos, 1991; Clark, Nergardar, & Cronin-Golomb, 2008; Dara, Monetta, & Pell, 2008; Kan, Kawamura, Hasegawa, Mochizuki, & Nakamura, 2002; Lloyd, 1999; Mitchell & Boucas, 2008; Pell & Leonard, 2003; Schroder et al., 2006; Scott, Caird, & Williams, 1984; Velez Feijo, Rieder, & Chaves, 2008; Yip, Lee, Ho, Tsang, & Li, 2003). The performances of the post-operative group on the categorical judgments can be interpreted in several ways. One could say that STN DBS enhances emotional abilities, as has been claimed in previous studies (Funkiewiez et al., 2006; Schneider et al., 2003). Nevertheless, while the post-operative group did not differ from the HC, nor did it differ from the pre-operative group. A more relevant hypothesis, from a methodological point of view, would therefore be that the categorical judgment analysis is less sensitive than the continuous one for detecting the emotional effects of STN DBS, notably because it induces “categorization biases” (Scherer & Ekman, 2008).

The second level of analysis, which consisted of an investigation of continuous judgments, allowed us to probe our data in greater depth, and results seemed to confirm our second hypothesis on the poor sensitivity of the categorical judgment methodology (Figs. 1 and 2, and Table 5). Overall, the analysis revealed an interaction effect showing that the pre, post, and HC groups displayed different patterns of responses on the different scales and different emotions. In order to be more specific and to characterize this overall difference more accurately, systematic contrasts were performed between all three groups.

These contrasts revealed that, compared with the other two groups, the post-operative group displayed disturbances on the fear and anger target scales; they also displayed disturbances on sadness scale, compared with the post-operative group. These contrasts also revealed a significant difference between the post and HC groups on the non-target scales, whereas there were no such differences between the pre-operative and HC groups. Compared with the pre-operative and HC groups, the post-operative group rated the fear scale significantly more intensely when they listened to anger stimuli. Similarly, this same group rated the happiness scale more intensely when they listened to fear stimuli. Furthermore, contrasts revealed that, compared with the other two groups, the post-operative patients were biased in their rating of the surprise scale, rating this scale significantly more intensely when they listened to angry or fearful utterances. These results would appear to support our initial hypothesis that STN DBS modifies the ability of PD patients to recognize emotions, not only when these are conveyed by faces but also when they are expressed through the human voice.

It is noteworthy that all three groups were matched for age, education and gender, and that the two PD patient groups were matched for disease duration, severity of motor symptoms and dopamine replacement therapy, thus indicating that neither demographic variables nor clinical characteristics accounted for the observed differences. It is now well documented that dopamine is involved in emotional processes (Salgado-Pineda, Delaveau, Blin, & Nieoullon, 2005); the absence of any significant difference between the two PD patient groups for dopamine replacement therapy suggests that modifications in the recognition of emotions from voices performances in the post-operative group cannot be explained by this variable. Similarly, the absence of any significant difference between the three groups regarding neuropsychological performances suggests that the group effects observed for the recognition of emotions from voices cannot be attributed to differences in general cognitive efficiency or executive function.

Moreover, all participants were subjected to an audiometric screening procedure to check that they had normal hearing; none of the patients included in the study wore hearing aids or had a history of tinnitus or hearing impairment. Lastly, the absence of any significant correlation between recognition of emotions from voices performances and MADRS scores indicates that the patients’ mood state cannot explain the modifications in recognition of emotions from voices performances in the post-operative group.

The present results appear to be in line with previous studies demonstrating impaired facial emotion recognition after STN surgery in PD patients (Biseul et al., 2005; Drapier et al., 2008; Dujardin et al., 2004; Le Jeune et al., 2008; Péron et al., in press; Schroeder et al., 2004). Therefore, the present study, along with the above, suggests that STN stimulation affects the recognition not only of facial emotions but also of vocal emotions, arguing in favour of multimodal alteration rather than changes specific to the visual modality. At first sight, the vocal impairment, which was only observed on continuous judgments in the post-operative condition in the present study, would seem to be qualitatively different from the facial impairment observed in previous studies. The
latter, however, only analyzed categorical judgments and did not investigate continuous ones. In this context, the discrepancy could simply be due to the poor sensitivity of the categorical methodology for the auditory modality (Scherer & Ekman, 2008). Further studies using continuous judgments to explore the recognition of facial expressions following STN DBS are needed to confirm this hypothesis.

Our findings appear to confirm a more general role of the STN in the processing of emotional information and enable us to overcome a major limitation of previous studies that have explored STN involvement in emotional processing. The latter used exclusively visual stimuli and, even though they controlled for an overall visuospatial deficit, it is impossible to exclude the possibility that the impaired recognition of facial expressions observed in the postoperative condition may have been due to oculomotor changes following STN DBS. The STN is now widely assumed to play an oculomotor role (Fawcett et al., 2007; Fawcett, Dostrovsky, Lozano, & Hutchison, 2005; Matsumura, Kojima, Gardiner, & Hikosaka, 1992). It has also been suggested that the ability to make normal use of information from the eye region of faces (through oculomotor exploration) is essential when judging emotions (Adolphs et al., 2005). The fact that changes in emotional processing are apparently present in the auditory modality, too, is an important argument against the “oculomotor” bias.

The performances of the post-operative group on this emotional prosody paradigm may have been a consequence of a systematic cognitive bias towards negative emotions (Mineka & Sutton, 1992; Wenzel & Finstrom, 2007). Compared with that of the other two groups, the performance of the post-operative group was characterized by a tendency to perceive negative emotions more strongly. This alteration may have been a consequence of changes in the relationships between the different brain regions, such as the amygdala and the orbitofrontal cortex, in the emotional attribution processes involved in decoding. Modulations in the neuronal activity of these two regions related to STN DBS were reported in a previous study (Le Jeune et al., 2008) investigating the decoding of emotional facial expressions. These modulations in neuronal activity may be related to an overall cognitive bias following STN DBS, whereby patients attribute more emotions to other people than healthy subjects do.

The present findings provide further support for the hypothesis that the STN is a key structure for regions known for their involvement in emotional processing, such as the amygdala and orbitofrontal cortices (Le Jeune et al., 2008; Temel, Blokland, Steinbusch, & Visser-Vandewalle, 2005). It is highly likely that changes in emotional processes are the result of dysfunction induced by STN stimulation affecting brain regions that are crucial to emotional processing, especially through the STN’s connections with the orbitofrontal and anterior cingulate circuits (Le Jeune et al., 2008), although the exact nature of the underlying mechanisms is not yet fully understood. Due to the small size of the STN, the stimulating current flow may well spill over from the STN’s targeted sensorimotor territory and affect other STN territories, too, in particular the limbic one, as well as the closely connected basal ganglia areas and the afferent or efferent connections to subcortical and cortical regions, particularly the amygdala, orbitofrontal cortices and anterior cingulate cortex, as supported by several neuroimaging studies (Hilker et al., 2004; Le Jeune et al., 2008, 2009; Schroeder et al., 2002, 2003). Another possible explanation, which in no way precludes the previous one, is that the functional partitioning of the STN is not absolute and there are probably interactions between subterritories (Joel & Weiner, 1997; Mallet et al., 2007). As some authors have suggested, the STN may play a role in the integration of the motor, cognitive and emotional components of behaviour, which may explain why stimulation of this core structure has effects on motor functions but also on emotional functions related to other brain regions and/or on associative functions (Mallet et al., 2007). Whether the STN acts as a behavioural regulator (Temel et al., 2005, 2006) or a behavioural integrator (Mallet et al., 2007), there is converging evidence that the STN plays a central role within a broadly distributed network of brain areas subtending emotional processing.

Our study suggests that the recognition of negative emotions expressed in the vocal modality is altered following STN stimulation in PD patients. The present results, specifically the differential impact of STN stimulation on the recognition of vocal emotions, now need to be confirmed with a larger sample of PD patients who have undergone STN stimulation, with a within-subjects design. It is also important to control for the natural course of the disease by including a group of matched nonsurgical PD patients assessed twice with the same interval as the surgical PD patient group (Péron et al., in press). Lastly, further evidence needs to be gathered from PD patients with STN DBS, in particular using functional neuroimaging to establish a correlation between deficits in decoding emotional prosody and changes in neuronal activation, particularly in those brain regions that are known to be involved in emotional processing, such as the orbitofrontal cortex and amygdala.

Conflict of interest

The authors report no conflicts of interest.

Acknowledgements

We would like to thank the patients and normal controls for contributing their time to this study, Mme Wiles-Portier for preparing the manuscript, and the Ear, Nose and Throat Department of Rennes University Hospital for conducting the hearing tests.

This study was funded by the Neurology Unit of Pontchaillou University Hospital in Rennes, France (Prof. G. Edan) and the Swiss Centre for Affective Sciences in Geneva, Switzerland.

Appendix A. Computer interface for the original paradigm of emotional prosody recognition

References


