Psychology of Language and Communication 2022, Vol. 26, No. 1

sciendo

DOI: 10.2478/plc-2022-0006

Federica Biassoni[1], Martina Gnerre[1], Eleonora Malaspina[1], Sonia Di Tella[1, 2], Isabella Anzuino[1], Francesca Baglio[2], Maria Caterina Silveri[1]

[1] Department of Psychology, Catholic University, Largo A. Gemelli, 1, 20123 Milan, Italy.
[2] IRCCS Fondazione Don Carlo Gnocchi ONLUS, Via A. Capecelatro, 66, 20148 Milan, Italy.

How does prosodic deficit impact naïve listeners recognition of emotion? An analysis with speakers affected by Parkinson's disease

This study aimed to understand the impact of the prosodic deficit in Parkinson's disease (PD) on the communicative effectiveness of vocal expression of emotion. Fourteen patients with PD and 13 healthy control subjects (HC) uttered the phrase "non è possible, non ora" ("It is not possible, not now") six times reading different emotional narrations. Three experts evaluated the PD subjects' vocal production in terms of their communicative effectiveness. The PD patients were divided into two groups: PD+ (with residual effectiveness) and PD− (with impaired effectiveness). The vocal productions were administered to 30 naïve listeners. They were requested to label the emotion they recognized and to make judgments about their communicative effectiveness. The PD speakers were perceived as less effective than the HC speakers in conveying emotions (especially fear and anger). The PD− group was the most impaired in the expression of emotion, suggesting that speech disorders impact differently at the same stage of the disease with varying degrees of severity.

Key words: Parkinson's disease, emotional prosody, communicative effectiveness, speech pathology

Address for correspondence: Federica Biassoni, Department of Psychology, Catholic University, Largo A. Gemelli, 1, 20123 Milan, Italy.

E-mail: federica.biassoni@unicatt.it

This is an open access article licensed under the CC BY NC ND 4.0 License.

A Concise Framework of Speech and Voice Disorders in Parkinson's Disease

Idiopathic Parkinson’s disease (PD) is a slow, chronic neurodegenerative disorder caused by a basal ganglia circuitry control dysfunction characterized by motor and nonmotor clinical manifestations. Vocal impairment due to hypokinetic dysarthria is a motor disorder that affects between 75% and 90% of patients during the disease (Defazio et al., 2016). The Parkinsonian syndrome motor symptoms may impair all dimensions involved in spoken production, including respiration, phonation, articulation, and prosody (Gillivan-Murphy et al., 2019; Defazio et al., 2016). In individuals with PD who have dysarthria, six out of ten of the most deviant speech features are associated with prosody: low speech intensity (hypophonia or mono loudness), flat speech melody (flat pitch), short rushes of speech, altered rate, inappropriate silence, reduced stress (Galaz et al., 2016; Mahler et al., 2015). These vocal dysfunctions should be analyzed not only as motor problems related to single patients, but also as a multidimensional impairment that affects interpersonal interaction, emotional well-being, and, as a result, the patients’ social context and their quality of life. Consequently, it is important to explore the perception and the subjective experience of such communication deficits from the listeners’ point of view to better understand the barriers to successful communication it poses (Miller, 2017). Communication decline in PD is perceived as dysphonia, which is often described as harshness and breathiness by listeners (Gillivan-Murphy et al., 2019; Mahler et al., 2015). Listeners also perceive distorted or omitted sound and/or syllables and words slurred together, which may give the impression that the patient is speaking too fast (Harris et al., 2016) and may prevent the comprehension of the verbal content (Tjaden & Wilding, 2011). The increased number of pauses is also dysfunctional in comparison to those of control speakers (Miller, 2017; Tjaden & Wilding, 2011).

Emotional Prosody Impairment and its Consequences on the Effectiveness in Conveying Emotions

Under normal circumstances, listeners can adapt to relative changes in pitch, duration, and loudness, or speech prosody, to infer emotion or the affective state of a speaker, and in order to resolve syntactic ambiguities (Anolli & Ciceri, 1992). Indeed, prosody has several functions related to emotion expression and to pragmatics, and has an impact on the naturalness and intelligibility of speech, being ”the music of everyday speech” (Wennerstrom, 2001). When listening to the impaired speech of patients with PD, the listeners cope with some difficulties in understanding the tone, implications, and emotional flavor of an utterance that are signaled by the prosodic content. Analyzing the speech of PD patients, Cheang and Pell (2007) showed abnormalities in their ability to acoustically emphasize specific words within an utterance to underscore the salience of these word meanings in the verbal context. Speakers affected by PD are also restricted in their ability to modulate fundamental frequency (F0) and intensity in an emotional intonation task but not in a nonemotional phonation task, suggesting an emotional disturbance in addition to the motor impairment (Möbes et al., 2008). Furthermore, some studies highlighted different levels of effectiveness in the vocal emotion expression of different emotions. Borod et al. (1990) showed that negative emotions are better expressed by PD patients than positive ones, while Pell et al. (2006) found that PD patients are interpreted by listeners with greater difficulty compared to control speakers, particularly for the expression of anger and disgust. In this last study, the exception is utterances of sadness produced by PD speakers, which are frequently characterized as sounding either sad (in this case with a correct attribution) or neutral. Such data highlight how easy it is to wrongly attribute an emotional label to the vocal production of a subject affected by PD. In partial contrast with these studies, other studies on emotion expression by patients with PD suggest an arousal modulation deficit both in emotion expression and in emotion recognition processes rather than an impairment regarding a specific emotion. Namely, some findings argued that the arousal dimension (the level of autonomic activation that varies from low to high) proves to be a more relevant variable than the valence dimension (the level of pleasantness in a continuum from negative to positive) in discriminating the area of preserved effectiveness and the area of impairment. Research about emotion expression in healthy speakers revealed that the expression of high-arousal emotion is characterized by higher mean F0 and amplitude, with greater F0 and amplitude variation than in emotionally neutral utterances (Anolli & Ciceri, 1992). Interestingly, the same distinction between the arousal and the valence dimension needs to be taken into account when analyzing PD patients' ability to recognize emotional states from the facial expression (Wieser et al., 2006). In fact, patients with PD reported a lower level of physiological arousal than healthy subjects in the recognition of emotional stimuli demonstrating a reactivity reduction to highly arousing stimuli (Miller et al., 2009). As a proof of the arousal modulation deficit in the emotional prosody of patients with PD, it can be noted that dysarthria associated with PD may cause deficits in the suprasegmental production that further reduces the range of F0 variability and consequently the high arousal emotions expression, so the speech turns out monotonous. The monotonous and hypophonic nature of PD patients' speech erroneously suggests that they are aloof and emotionally detached (Benke et al.,1998; Pitcairn et al.,1990), as well as depressed, unmotivated, reluctant to engage in conversation (Jaywant & Pell, 2010), and also pragmatically "inappropriate." Blind analyses of speech samples from patients with PD and healthy subjects, where listeners were unaware who was or was not affected by PD, show that these patients were judged to be introverted, withdrawn and anxious, alongside other negative social characteristics (Jaywant & Pell, 2010; Pitcairn et al., 1990; Pell et al., 2006). Pell et al. explored how dysprosody altered a PD patient's overall communicative intent and indicated that dysprosody can lead to negative social relationships, as unfamiliar listeners had greater trouble identifying the intended meaning (Pell et al., 2006). Harris et al. (2016) showed that the negative effects of PD dysprosody on fruitful communication were reported by unfamiliar listeners. It has been suggested that, while having the capacity to produce the same sentences, speakers with PD are perceived to be less “intense” by naïve listeners during a spontaneous language generation task (Blonder et al., 1989). Previous studies also suggest a possible distorted perception of their own vocal abilities by the patients themselves (Kwan & Whitehill, 2011) and their frustrations in failing to convey their affective intentions (Gillivan-Murphy et al., 2019). Collectively, these findings consolidate that “disprosody is intrinsic to Idiopathic Parkinson’s disease” (Darkins et al.,1988) and that this phenomenon could act as a serious relational and social barrier for these individuals. It also means that negative stereotyping may take place due to the patients’ speech patterns. Despite a consensus that dysprosody is a persistent symptom, few studies have investigated how effectively the voice of patients affected by PD communicates emotion. Such studies also usually take into account expert assessment (Benke et al, 1998; Blonder et al., 1989, Caekebeke et al., 1991; Pell et al., 2006; Penner et al., 2001). For example, in Martens et al. (2011), according to the judgement of professional listeners no difference was found between the samples of healthy participants and those with PD in terms of emotional prosody.

The Present Study

The present study is part of a recent trend to investigate the emotional expression domain through prosody in the clinical population of patients with PD (Benke et al., 1998; Blonder et al., 1989; Caekebeke et al., 1991; Hsu, 2016; Miller, 2017; Penner et al., 2001). It adds to literature by providing data about the perception by naïve listeners of the effectiveness of PD patients expressing five different emotions: fear, anger, happiness, disgust, and sadness. Specifically, the present study aimed at adding to previous research in the field of the effectiveness of vocal emotional expression by PD patients in three ways. First, previous studies have highlighted the limitations of the instrumental (acoustic) measurement of voice, showing that the auditory perceptual assessments are more relevant in describing voice quality (Gerratt et al., 1991). The perceptual descriptions are part of the classical and traditional evaluation of voice quality, also used with patients with PD (Santos et al., 2010). According to such an approach (Oates, 2009), global perceptive evaluation involves higher cognitive mechanisms and consolidated emotional scripts. Therefore, perceptual evaluation comprises important elements, such as pragmatic functions and emotional meaning, that are not attainable through acoustic analysis. For this reason, in the present study, we chose to ask listeners to provide a global perceptual evaluation of the proposed stimuli. Second, the auditory rating has often been used to investigate dysphonic voices (Oates, 2009), but few studies have applied auditory perceptual judgements to collect a global evaluation of effectiveness in the vocal expression of emotions. Moreover, the present study involved naïve listeners. Indeed, if the goal is to evaluate the "hoarseness" or “breathiness” of a voice (auditory ratings), an expert listener should be questioned. Instead, when exploring the effectiveness of vocal expression in conveying an utterance’s emotional flavor and in conveying a specific emotion distinguished from other ones, naïve listeners are the privileged interlocutors, and global evaluation is the elective procedure, as showed by Scherer (1995). In a domain as complex as prosody, the global evaluation provides a powerful method within which to explore the perceptual patterns as perceived by nonexpert listeners who are not influenced by theoretical bias. Therefore, each vocal emotional profile should be analyzed through perceptual evaluations that rate the global effectiveness of emotion expression rather than specific variations in one acoustic parameter or another (such as the F0 variability or the rate of the speech). Such evaluation complements the accounts of the acoustic differences provided by expert auditory analysis. Third, in the present study, we described the global perceptual evaluation in terms of three properties of the stimuli: expressiveness, that is, the ability to convey an emotional flavor, the ability to discriminate between different emotions, and communicative effectiveness, that is, the ability to disclose a communicative intention (Sperber & Wilson, 2002). Furthermore, it has been noted that most studies aiming to investigate vocal emotional recognition have considered raw hit scores (Lima et al., 2013; Sauter et al., 2010; Schroder, 2003), which mainly fail in measuring accuracy. As a fourth element of innovation to this study and to overcome such a limitation, we chose to use the unbiased hit rate (Wagner, 1993) which takes bias and response habits into consideration.

Objectives

The current study aimed to understand the impact of the prosodic deficit associated with PD on the effectiveness of expressing emotions via the vocal channel, measured through naïve listeners' assessments. The emotional expression ability has been analyzed with regards to two components: the competence in conveying the emotional flavor and denotation of an utterance, and the ability to express different emotions in a distinguishable manner. The two components are reflected in the chosen measures: not only the degree of discrimination of different emotions, but also the global subjective evaluations (about expressiveness, ability to discriminate, and communicative effectiveness) by naïve listeners.

Moreover, the present study builds upon the idea that the prosodic deficit related to PD dysphonia represents a complex pattern of deficient components involving different acoustic parameters and their combination, which are shaped in different ways. Hence, a second objective of the present study was to verify that perceptual evaluation by naïve listeners distinguishes two different patterns of acoustic parameters (named PD+ and PD−), previously identified through the perceptual analysis of expert listeners. Namely, the present study aimed at verifying that such different patterns of acoustic parameters are perceived as different in their effectiveness of expressing emotions when assessed by naïve listeners in a perceptual evaluation task, analogously to what is highlighted by the expert listeners' assessments.

Hypotheses

Two hypotheses served as the basis of the present study. The first hypothesis is that, starting from a proven vocal emotional expression deficit in PD, PD speakers' utterances, compared to a healthy control (HC) group (a) will be decoded less accurately in terms of the emotion they express, and (b) will be perceived as less effective in expressing emotions according to a global evaluation.

Moreover, since the motor symptoms in PD affect the expression of typical emotion acoustic-perceptual characteristics, we assumed that the high arousal emotions (anger, fear, happiness) will be the most impaired (Wieser et al., 2006) and therefore less frequently recognized. In particular, in relation to anger, this study referred to *hot anger*, which is characterized by an intense outward display of anger and higher F0 and F0 variability, higher intensity, and increased speech rate (Johnstone et al; 200; Biassoni et al; 2016)

The second hypothesis was that within the domain of deficient patterns of acoustic parameters characterizing the emotional vocal expression of patients with PD, two different patterns can be distinguished with regard to their effectiveness in conveying emotions.

Method

Ethical Approval

The study was approved by the IRCCS Don Carlo Gnocchi Foundation Ethics Committee, and it was performed in accordance with the principles of the Helsinki Declaration. All the participants provided their written and informed consent.

Stimuli Production and Selection

Speaker Sample

Fourteen subjects with idiopathic PD were recruited (seven females and seven males, age: 69,93 ± 7,12) and 13 healthy control speakers (seven females, six males, age: 68,13 ± 8,27).

Subjects in the PD group were recruited from the Don Carlo Gnocchi Foundation Neurological Unit (Milan, Italy). Participants in the control group were recruited among healthy familial members or caretakers of patients from the neurology outpatient clinic. None of the participants had any previous neurological or psychiatric disorders. All subjects were native Italian speakers. Participants had an education threshold of eight years. All patients fulfilled the clinical diagnostic criteria of PD according to the United Kingdom Parkinson's Disease Society Brain Bank (Hughes et al., 1992) and a positive DATscan imaging. The PD severity was rated by the Modified Hoehn and Yahr (H&Y) Scale (Goetz et al, 2004), and the scoring for all participants was in the mild to moderate stages of the disease, with a scoring between Stages 1 and 2.5. The PD patients also underwent an examination of movement disorders with the Unified Parkinson Disease Rating Scale (UPDRS) – part III (Goetz et al; 2008). The absence of dementia in all speakers was confirmed, as PD and HC speakers underwent a series of neuropsychological tests to assess cognitive functioning. Specifically, the neuropsychological assessment of PD patients included: the Mini-Mental State Examination (MMSE) (Folstein et al; 1975), object and action oral naming, phonological fluency, semantic fluency, immediate and delayed recall of 15 words, free and cued selective reminding test (Frasson et al; 2011), The Rey-Osterrieth Complex Figure (copy and immediate recall (Caffarra et al., 2002), verbal span and Corsi's test (Monaco et al., 2013), Raven Colored Matrices (Raven, 2003), Trail making test (Reitan, 1971), attentional matrices (Spinnler & Tognoni, 1987), Stroop test (Stroop, 1935), and the modified Wisconsin Card Sorting test (Caffarra et al., 2004). Neuropsychological assessment and experimental tasks were administered in two different sessions so as not to strain the participants.

Neuropsychological assessment of the healthy control group was limited to the MMSE. All PD participants had regular, stable drug therapy with L-Dopa (alone or in association with dopamine agonistics, catechol-O-methyltransferase inhibitors, monoamine oxidase inhibitors, and anticholinergic drugs, anticholinergic drugs). Exclusion criteria for potential participants were: clinical signs satisfying the criteria of other neurological disorders, including possible atypical PD, secondary or iatrogenic PD, and major psychiatric illnesses, excluding the presence of mild-moderate depression and claustrophobia.

Two subjects were excluded because of the acoustic echo that may have been caused by the reflection of sound waves and acoustics between the loudspeaker and the microphone. All procedures were carried out with the adequate understanding and informed consent of the subjects. Each selected subject was then identified with an anonymous ID.

Stimuli Production Task

Regarding the elicitation of emotional prosody, all subjects (PD and HC) read six emotional texts, each one depicting a scenario representative of a

Table 1. *Demographic and Clinical Features of Speakers with Parkinson's Disease (PD) and Healthy Control (HC) Participants (M ± SD)*

Variable	PD (*n* = 14)	BC (*n* = 13)
Age	69.93 ± 7,12	68.13 ± 8,27
Sex (M+F)	7+7	6+7
Education (years)	12.20 ± 3.99	13.20 ± 3.69
UPDRS motor score	27.80 ± 10.60	
Disease duration (years)	41.80 ± 21.37	
LEDD	174.31 ± 159.18	
MMSE	27.64 ± 1.79	27.85 ± 1.32

Note. UPDRS = Unified Parkinson Disease Rating Scale; MMSE = Mini-Mental State Examination; LEDD = levodopa equivalent daily dose.

prototypical situation, selected from actual situations reported by a large number of respondents in response to the request to recall a specific emotion category (Scherer et al., 2001). Each text contained the same standard sentence that, when considered out of context, is not endowed with any emotional meaning or connotation. According to the emotion elicitation procedure (Anolli & Ciceri, 1992; Scherer et al., 1991), the sentence acquires specific emotional value thanks to the interaction with the context (pragmatic function of the context, Fatigante et al., 2016). The target utterance "non è possible, non ora" ("It is not possible, not now") appeared in five different emotional contexts: fear (F), anger (A), happiness (H), disgust (D), and sadness (S), and in one neutral expression (N). This sentence has already been used in previous research (see Anolli & Ciceri 1992; Anolli et al., 2008). The decision to employ a single standard sentence was motivated by the need to control for any possible influence or interference of segmental phonemic differences. Indeed, this emotion elicitation methodology, compared to that of spontaneous speech collection which is more sensitive to prosodic abnormalities, is more reliable precisely because it is standardized. The participants (both PD and HC) were asked to read the whole text, to identify with the main character, and then to speak the target utterance as they would speak in the depicted scenario. The same procedure was followed for the six scenarios (five emotional scenarios and one neutral situation). All utterances (total number: 162, six utterances for each one of the 27 speakers) were audio recorded (participants wore an AKG C 520 head-worn condenser microphone, selected for its ability to collect accurate speech signals). Recordings were obtained in a quiet room, and the presented emotional texts were administered in a randomized order.

Experts' Perceptual Evaluation and Selection of Stimuli

Perceptual evaluation is defined as the set of procedures that refer to the expert listener's ability to determine the quality of the voice, regardless of instrumental measurements (De Almeida et al., 2019). To reduce or eliminate biases, three experts listened to the sentences without knowing the speakers' diagnosis. The experts took part in two distinct tasks:

Emotion recognition task

The experts listened to the audio tracks of the utterances produced both by PD patients and HC subjects. Each expert assigned a label to the expressed emotions, choosing one from a list of eleven labels (including nostalgia, compassion, doubt, irony, and jealousy, which had been added to the list as distractor categories, in addition to the five targeted emotions and the neutral condition). According to Russell (1993), such a procedure prevents a forced-choice response format).

Global perceptual evaluation task

Three criteria were identified to evaluate the vocal performance in reference to the expression of emotion:

- expressiveness: the ability to use prosody and rhythm to convey an emotional flavor to the listener;
- ability to discriminate: competency in the use of a specific nonverbal vocal pattern to express a specific emotion and therefore to discriminate between different emotions. While *expressiveness* refers to the ability to be expressive, *ability to discriminate* refers to the ability to effectively use vocal features in order to distinctively express different emotions.
- communicative effectiveness: an overall judgment about the ability to realize the communicative intention in expressing a certain emotion. It comprises and integrates the global ability to communicate emotions, to distinguish between different emotions, as well as the ability to integrate the expression of emotion with individual intentions, will, and needs, and to engage in successful interactions with others (Hustad, 1999; Dykstra et al., 2015).

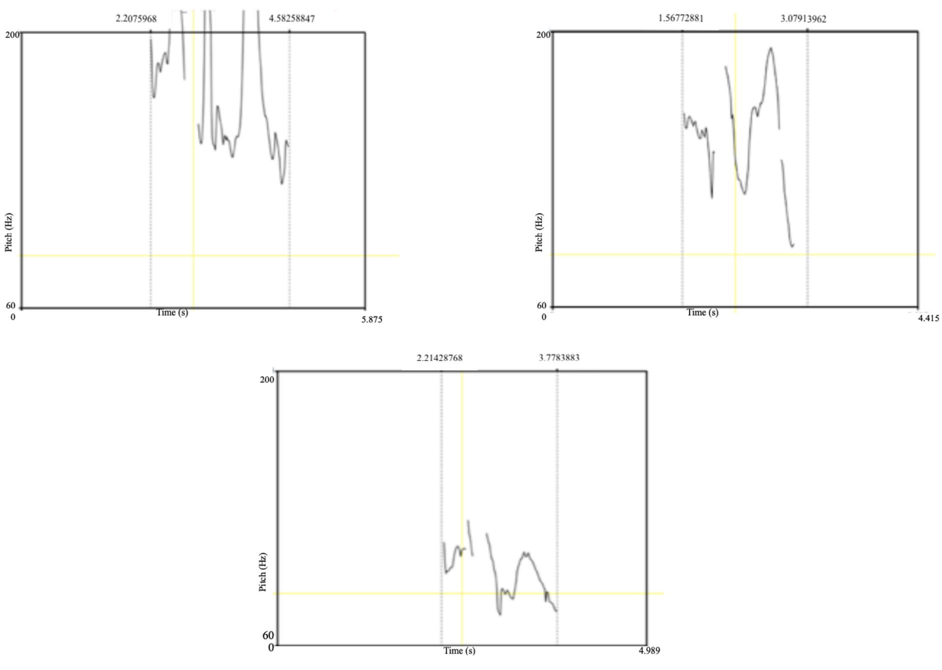
For every utterance, each expert rated the three criteria scoring on a 5-point Likert scale. All ratings (from the emotion recognition task and from the global perceptual evaluation task) and each emotional utterance were considered. Inter-rater reliability between the three independent judges was calculated using Krippendorff's α (see Tables 2 and 3).

Based on the ratings by the three experts, PD patients were assigned to two profiles on the basis of the patterns of acoustic parameters and related perceptual features characterizing their utterances. Hence, each profile corresponded to a low or high level of the perceived communicative effectiveness assessed by the expert listeners, and to a specific acoustic profile:

The first profile (PD−) was characterized by a feeble voice, lacking in modulation, tonality, expressiveness, and "sound body." Regarding the expert evaluation, the average score for communicative effectiveness is 1.

The second profile (PD+) was characterized by a more modulated voice, with a wider range of sonic nuances (as well as superior expressive and communicative power). Regarding the expert evaluation, the average score for communicative effectiveness is 2.25. The most and the least effective utterances in each category (PD+ and PD−) were selected as stimuli for the experiment with the naïve listeners together with an utterance from a healthy subject (HC). A total of 36 utterances were selected: six utterances per scenario (the most and least effective produced by PD+, the most and least effective produced by PD−, and the most and least effective produced by HC), for each one of the six scenarios (five emotional scenarios and a neutral one). For HC subjects, the average score of communicative effectiveness was 4. By way of example, Figure 1 shows the graphic representation of the pitch contour of the target utterance "non è possible, non ora" when produced by a healthy speaker, a PD+ patient, and a PD− patient. It shows that the first and second are more varied than the third, which appears flattened, suggesting one element of the different perceptual experience by the listeners.

Figure 1. *Pitch Contours of the Same Utterance Produced by Healthy Speakers (Upper Left), PD+ Patients (Upper Right) and PD- Patients (Bottom). The intonation Contours Were Extracted Using the PRAAT Software (Boersma & Weenink, 2018)*



Naïve Listeners' Perceptual Evaluation

Listeners' Sample

The sample was composed of 30 listeners, 15 men and 15 women. Their mean age was 24.2 years (*SD* = 5.8) and their mean number of years of education was 15.7 (*SD* = 2.1). All listeners were native Italian speakers and self-reported normal hearing. All listeners were university students, recruited through an announcement, and they took part in the research voluntarily. None of them knew the PD patients personally.

Naïve Listeners' Perceptual Evaluation

For the emotion recognition task, each participant listened to 36 audio tracks, previously selected through the experts' evaluation. Each track expressed a specific emotional state (or no emotion state): anger, happiness, fear, sadness, disgust, and neutral, and has been uttered by a PD patient with a PD+ profile, or by a PD patient with a PD− profile, or by a HC participant. The listeners were unaware of the speaker's condition (PD or HC). Participants were told that they would hear a total of 36 utterances, each expressing an emotion (or no emotion). Then they were asked to indicate the expressed emotion from a list of 11 labels (*anger*, *fear*, *happiness*, *disgust*, *sadness*, and the distractor emotions that included *nostalgia*, *compassion*, *doubt*, *irony*, and *jealousy*).

Moreover, the participants were requested to express a global judgement about expressiveness, ability to discriminate, and communicative effectiveness on a 5-point Likert scale (same as for the expert listeners). Each utterance was played up to two times. Participants were tested individually in a quiet setting and listened to the stimuli over headphones at a comfortable volume. Stimuli were presented randomly, and the speaker's identity was anonymized. Before the test session began, one practice trial was performed.

Design and Statistical Analysis

For the emotion recognition task, the main investigation consisted of a 3 × 6 mixed-design analysis of variance (ANOVA), with two independent variables: speaker's profile (PD+, PD−, and HC) as the between-subjects factor and emotion (happiness, sadness, fear, anger, disgust, and neutral) as the within-subject factor. The main dependent variable was recognition accuracy for each emotion. Each answer was coded as correct (in case the emotion was recognized correctly) or incorrect (in case the emotion was not recognized, both when the listener indicated a different emotion or provided no answer). To account for possible stimulus or response biases, emotion recognition performance was assessed using the unbiased hit rate (*Hu*, Wagner, 1993) prior to statistical analysis. The term *unbiased hit* rate refers to a precise measure of classification performance that accounts for false alarms and biases in the use of response categories by multiplying the raw hit rate by the positive predictive value (PPV, Wagner, 1993). The PPV is calculated as a proportion of predictions of a situation that correctly forecast the situation (Armistead, 2013). The *Hu* score varies between 0 and 1 (0 = *no correct responses*; 1 = *perfect performance*). The converted scores were entered into a mixed-design ANOVA. A pairwise multiple comparison procedure using Tukey's honestly significant difference (HSD) test to control for experiment-wise error rate was also conducted. The IBM SPSS Statistics 25.00 (SPSS, Corp., USA) software package was used for statistical analyses. Significant interactions were examined using Tukey's HSD (Tukey α = 0.05). The global error rate for each speaker's profile (PD−, PD+, and HC) was also calculated. It consistsed of a percentage of the listeners' errors in the emotion recognition task regardless of the specific emotion. It was the ratio of the total of recognition errors committed for each group and the maximum number of possible errors.

The selection rate (both correct and wrong answers for each emotion) in the three different speaker profiles (PD−, PD+, and HC) was provided. For the global perceptual evaluation, the main investigation consisted of a 3 × 6 mixed-design ANOVA with two independent variables: speaker's profile (PD+, PD−, and HC) as the between-subjects factor and emotion (happiness, sadness, fear, anger, disgust, and neutral) as the within-subject factor. The main dependent variables were expressiveness, ability to discriminate, and communicative effectiveness, each one measured through a 5-point Likert scale. For each statistical analysis, a *p* value of < .05 was considered statistically significant.

Results

The results from the experts' perceptual evaluation and from the naïve listeners perceptual evaluation are reported in succession.

Table 2. *Means, SDs, and Krippendorff's α of Unbiased Hit Rate (Hu) for Emotion Recognition by the Experts*

Emotions	PD-		PD+		HC	
	Hu (*M*, *SD*)	α	*Hu* (*M*, *SD*)	α	*Hu* (*M*, *SD*)	α
Anger	0.08 (0.07)	0.1	0.27 (0.02)	0.47	0.88 (0.03)	0.9
Disgust	0.01 (0.02)	0.1	0.04 (0.1)	0.1	0.3 (0.05)	0.5
Happiness	0.09 (0.01)	0.1	0.09 (0.0)	0.1	0.28 (0.2)	0.48
Neutral	0.18 (0.08)	0.3	0.21 (0.02)	0.33	0.36 (0.5)	0.56
Fear	0.07 (0.02)	0.1	0.23 (0.01)	0.41	0.86 (0.4)	0.86
Sadness	0.27 (0.04)	0.47	0.29 (0.07)	0.49	0.36 (0.07)	0.6

Note. Krippendorff's α benchmarks: > 0.67 (acceptable) and > 0.80 (adequate).

Table 3. *Means, SDs, and Krippendorff's α of the Global Perceptive Expert's Evaluations*

Global perceptive evaluations	PD-		PD+		HC	
	M (*SD*)	α	*M* (*SD*)	α	*M* (*SD*)	α
Expressiveness	1.64 (0.5)	0.68	2.76 (0.18)	0.63	4.1 (0.8)	0.68
Ability to discriminate	1.49 (0.3)	0.61	2.61 (0.8)	0.59	4.12 (0.4)	0.69
Communicative effectiveness	1.53 (0.7)	0.68	2.72 (0.4)	0.69	4.03 (0.8)	0.69

Note. Krippendorff's α benchmarks: > 0.67 (acceptable) and > 0.80 (adequate).

Emotion Recognition by Experts

The *Hu* for each emotion were calculated. Krippendorff's α scores were generated to assess the agreement among the three experts on the evaluations related to each emotion.

Global Perceptive Evaluations by Experts

Means, *SD*s, and Krippendorf's α scores were estimated for the the experts' global perceptive evaluations (see Table 3).

Emotion Recognition by Naïve Listeners

Although the *Hu* was lower for all the considered emotions when the speaker was a PD patient (both PD+ and PD−) compared to HC (see Table 4 for descriptive data), statistical significance was present only for anger and fear (see Table 5). Tukey's HSD test was used to probe the differences between the groups and showed that the mean in the HC group was statistically different from the mean of both the PD+ and PD− groups, but there was no significant difference between the latter two groups. The difference between the HC and PD+ groups was significant ($p < .001$). Similarly, the difference between the HC and PD− groups was significant ($p < .001$), whereas the difference between the PD+ and PD− groups was not statistically significant (p = .887). The main effects of the type of emotion ($F[5] = 15.13$, $p < .05$; $\eta^2 = 0.36$, $\eta_p^2 = 0.84$, observed power = 0.99) and of the speaker profile ($F[2] = 50.24$, $p < .05$, $\eta^2 = 0.48$, $\eta_p^2 = 0.84$, observed power = 1) were observed. Moreover, an interaction effect between the type of emotion and the speaker profile was observed ($F[10] = 5.43$, $p < .05$, $\eta^2 = 0.25$, $\eta_p^2 = 0.75$, observed power = 0.94).

Data showed that the effect of the speaker profile on the degree of emotion recognition was registered only for two emotions: anger and fear. Moreover, a significant difference in the degree of emotion recognition distinguished PD patients (both PD+ and PD−) from healthy controls but did not distinguish between PD+ and PD− patients (see Figure 2).

Noticeably, the global error rate, regardless of individual emotions, was significantly higher for the PD+ (76.94%) and PD− (79.72%) groups than it was

Table 4. *Means and SDs of Unbiased Hit Rate (Hu) for All Emotions*

Hu	PD-	PD+	HC
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Anger	0.07 (0.07)	0.21 (0.01)	0.59 (0.05)
Disgust	0.03 (0.02)	0.11 (0.0)	0.11 (0.03)
Happiness	0.09 (0.04)	0.03 (0.04)	0.28 (0.1)
Neutral	0.12 (0.08)	0.20 (0.03)	0.27 (0.2)
Fear	0.01 (0.00)	0.06 (0.04)	0.66 (0.1)
Sadness	0.10 (0.04)	0.06 (0.02)	0.25 (0.08)

Table 5. *ANOVA Comparing PD+, PD- and HC Groups*

Hu	*F*	*p*
Anger	14.47	**.02**
Disgust	2.99	.19
Happiness	1.915	.29
Neutral	5.44	.10
Fear	114.42	**.001**
Sadness	4.86	0.11

for the HC group (48.33%). There was a statistically significant difference between the three groups ($F[2,27] = 4.467$, $p = .021$). Tukey's post hoc test revealed that the errors committed when the speaker was in the HC group were statistically significantly lower than when the speaker was a PD+ group ($p = .016$) or a PD− group (p = .04) patient. There was no statistically significant difference between the two PD groups ($p = .75$). Indeed, the global error rate distinguished between PD speakers and HC speakers but not between the two PD profiles (PD− and PD+).

Table 6 shows the descriptive data for the selection rate for each emotion. A statistically significant difference emerged only for the neutral condition (F = 18.79, p = 0.01). Tukey's HSD test showed that the PD– group was statistically different form the other two groups ($p = .01$). For the three criteria used to rate the global perceptive evaluations, namely, expressiveness, ability to discriminate, and communicative effectiveness, a significant difference between patients with PD and HC participants emerged (expressiveness: $F[3] = 35{,}699$, $p < .05$, $\eta^2 = 0.55$, observed power = 1; ability to discriminate: $F = 43.169$, $p < .05$, $\eta^2 = 0.59$, observed power = 1; communicative effectiveness: $F = 35.37$; $p < .05$, eta squared = 0.550; observed power = 1). Tukey's post hoc test revealed that when the speaker was a HC subject, the score for the three perceptive evaluations (expressiveness, ability to discriminate, and communicative effectiveness) was significantly higher than for PD+ ($p < .05$) and PD− ($p < .05$) patients. There was a statistically significant difference between the two PD groups as well ($p < .05$).

Figure 2. *Interaction Graph Showing the Unbiased Hit Rate (Hu) for Each Emotion Based on the Three Identified Profiles*

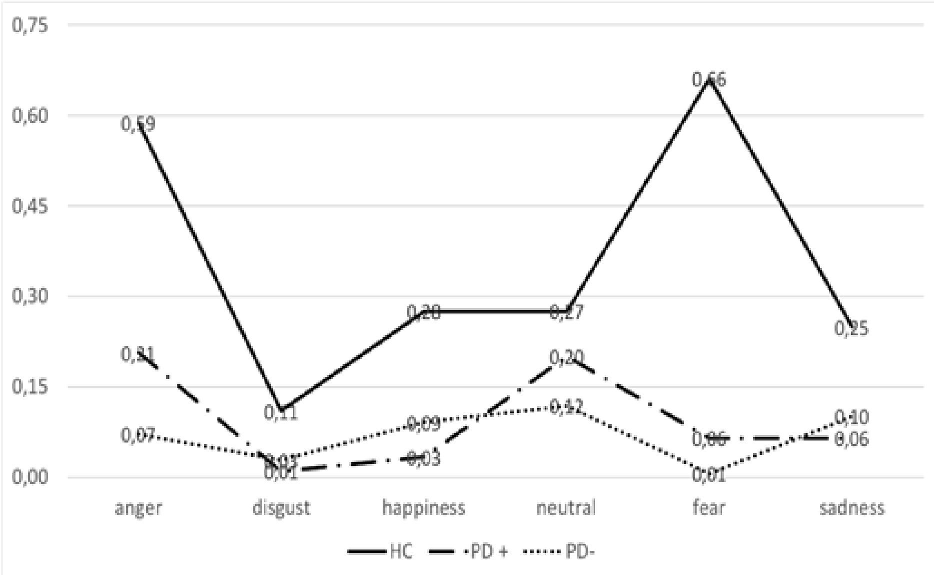


Table 6. *Selection Rates (Both Correct and Wrong Answers for Each Emotion)*

Emotion	PD-	PD+	HC
Anger	13%	14%	19%
Disgust	9%	8%	7%
Happiness	3%	3%	5%
Neutral	20%	7%	8%
Fear	8%	10%	15%
Sadness	15%	14%	14%
Distractor emotions	32%	44%	32%

Table 7. *Means and SDs of Global Perceptive Evaluations*

Global perceptive evaluations	PD-	PD+	HC
Communicative effectiveness	1.65 (0.94)	2.6 (0.76)	3.8 (0.92)
Discriminative ability	1.6 (0.77)	2.5 (0.90)	3.8 (1.04)
Expressive capacity	1.75 (1.01)	2.75 (0. 94)	3.95 (1.11)

Table 7 shows that the three variables displayed a very similar trend. With regard to the decoding task, there was a difference based only in terms of diagnosis (HC vs. PD), while the results in the global perceptive evaluation task highlighted a difference between all three vocal profiles (HC, PD+, and PD−). This result is consistent with the results of the expert evaluation by the three judges. Finally, the Pearson correlation coefficient showed a positive linear relation ($r = 0.88$, $p = < .05$) between the *Hu* and the scores from the evaluation of the communicative effectiveness.

Discussion and Conclusion

The present study investigated the experience of 30 naïve subjects listening to the emotional vocal expression of speakers affected by PD compared to the emotional vocal expression from HC speakers. PD patients were grouped into two profiles (PD + and PD−) depending on the communicative effectiveness of their vocal expressions as rated by three expert judges.

The results showed that the recognition of the emotions from the perspective of naïve subjects was significantly impaired when the speaker was a PD patient (regardless of the classification as PD+ or PD−) compared to HC participants. Such results are consistent with our hypothesis and with the literature highlighting the role of hypokinetic dysarthria in PD vocal impairment.

Though the expert listeners recognized differences between the two PD profiles (as the Hu and inter-rater agreement showed), these were not perceptible to the naïve listeners. We assume that the experts' professional background made them more prepared to grasp a wealth of emotionally relevant information from the perceptual clues in vocal expression. Indeed, it is possible that the expert listeners, through a perceptual evaluation, were able to better distinguish the acoustical properties (pitch, time patterns, etc.) they know are typical of the patterns featuring various emotions expression.

Instead, in the global evaluation task, the naïve listeners distinguished the two profiles (PD + and PD−). We assume that in the global evaluation task, the naïve listeners performed innate gestalt-oriented processing rather than an analysis of the single components, where the whole is phenomenologically more than its parts (Town et al., 2020). Moreover, the wider range of scores used in the global evaluation task (rated using a Likert scale) guaranteed a better differentiation of the different speakers' profiles.

The collected data are consistent with the literature as well (Caekebeke et al., 1991; Defazio et al., 2016; Harris et al., 2016; Schröder et al., 2010), namely, that the PD voice is devoid of an efficient "prosodic scaffold" and, therefore, the vocal emotional expression by PD patients is similar, as far as prosody is regarded, to neutral speech, since it is characterized by unmodulated rhythmic and pitch patterns. That is to say, the voice of PD patients is perceived as monotonous, static, lacking in expressiveness, and analogous to neutral speech. Such a similarity is confirmed by another result from the present study: that when listening to the speech expressed by a PD patient, participants selected the "neutral" labels with a higher frequency, often even when this label was incorrect. This data shows how easy it is to wrongly attribute an emotional label to the vocal production of PD patients. Noticeably, although the overall results suggest a general vocal emotional expression deficit involving all the examined emotions, especially anger and fear, it is possible to assume that the most preserved encoding patterns cover the expression of emotions defined by the combination of negative hedonic valence and low arousal, namely, sadness and disgust. Indeed, the expressions of these two emotions were not statistically different between HC and PD participants, although PD patients were not completely perceptually comparable to HC participants.

Pell et al. (2006) found that patients with PD have selective difficulties in expressing anger and disgust. The present study adds to the previous results supporting the hypothesis that the emotional states for which expression proves to be more difficult for PD patients are those featuring a combination of high arousal and negative valence, namely, anger and fear.

A possible explanation for these findings claims that the motor limitations in PD cause a distortion in the production of some typical acoustic-perceptual attributes essential to distinguishing different vocally expressed emotions. In this sense, the motor limitations of PD speech account for their greater prosodic inability to express sounds/acoustic features usually associated with fear and anger (e.g., the higher mean F0). However, perceptual analysis by the expert listeners highlighted a deficient vocal expression of happiness, too. The label “happiness” was rarely selected for any of the three conditions (HC, PD+, and PD−), and such a result seems to be consistent with the claim by Ekman (1994), that this emotion is poorly discriminated based on vocal parameters. Instead, it is more commonly recognized if expressed through the facial channel. Interestingly, even when uttered by HC speakers, the emotion of happiness was still recognized less frequently than other emotions. Therefore, arousal appears to perform a primary function in this emotional prosodic dysfunction, while the valence dimension may interact with arousal in different and independent ways. In conclusion, as far as emotion recognition is concerned, the results of the present study do not support the hypothesis of a specific emotion expression deficit, but rather support the hypothesis of an arousal-modulated deficit (Miller et al., 2009), according to which the dimension of valence intervenes in an interdependent way.

The judgments by naïve listeners about expressiveness, ability to discriminate, and perceived communicative effectiveness showed a significant difference between HC participants, whose utterances achieved the highest scores, he PD+ participants, and finally he PD− participants. This data is fully consistent with previous research on this topic, highlighting the speech of PD patients as largely intelligible and ineffective in terms of expressive and communicative competence (Dykstra et al., 2015; Gillivan-Murphy et al., 2019). In this sense, HC subjects were perceived to be more expressive, precise, and effective in communicating emotions than PD+ patients, and considerably more so than PD− patients. This data demonstrates how PD can result in a significant impairment in communicative processes. In particular, these impairments harm the ability to express emotions in a competent and effective way, thus increasing the risk of relational and social barriers from both the perspective of the patient and that of their conversational partners, which can exacerbate the patient's withdrawal from the relational and social dimension (Gillivan-Murphy et al., 2019; Miller, 2017). Measuring expressiveness added some meaningful data to the results from the emotion recognition task. Indeed, it allowed to expand the knowledge on the prototypical "lack of expressiveness" characterizing the emotional speech production in PD and to distinguish different levels of expressiveness within PD patients. This outcome is consistent with the hypothesis that the PD– group's vocal expressions can be described as feeble, without tone modulation, and void of sound nuances. In summary, though some evidence already showed the differences of emotional speech production between HC and PD speakers in terms of acoustic patterns (Caekebeke et al.,1991; Cheang & Pell, 2007; Benke et al.,1998; Pell et al., 2006), the present study explored the diversity between HC and PD speakers and highlighted the existence of differences within the field of PD−impaired speech in terms of communicative effectiveness from the perspective of listeners.

Study Limitations and Future Directions

The main limitation of this study was the relatively small sample size. A larger sample size would allow a deeper and more precise understanding of the research topic. A second limitation concerns some characteristics of the listeners that were not controlled for. For example, there is a possibility that the listeners were driven in their emotional decoding processing by some specific characteristics related to the speakers, such as vocal style. In addition, each listener may exhibit different cognitive, emotional, and decoding processing abilities which could impact their emotion recognition performance. Finally, the sets of utterances judged by the experts and naïve listeners differed, as out of the initial 162 utterances (judged by the experts), only 36 were chosen for the second part of the study (playing all the audio tracks in the naïve listeners task would have been too time-consuming for the participants).

Future studies may widen the knowledge on this topic by examining the relationship between *communicative effectiveness* and *communicative participation*, a social construct that refers to the nature (and the quality) of an individual's involvement in life situations (World Health Organization, 2001). Furthermore, future research on Parkinson's disease may benefit from investigating emotional speech in a natural context and from analyzing other emotions in addition to the fundamental ones considered in the present study. Finally, in the wake of recent research focusing on prosody in communication disorders (see Phi et al., 2019), future studies may also provide further emotional acoustic metrics to classify PD. The impact of impaired communication in PD, specifically the impairment affecting the prosodic dimension, is worthy of further investigation given the early overall impact that prosodic abnormalities have on PD patients' quality of life. The results of perceptual analysis from the present study may represent an initial baseline measure to guide the clinical understanding of the emotional voice characteristics of patients affected by PD and to track their vocal differences in terms of communicative effectiveness. By shedding light on the possibility of different degrees of impairment in communication effectiveness among PD patients, the present study opens up questions about the possible causes (excluding phonetic dysarthria) of such impairment. Since the subjects of the present study were all at the same stage of progress of the disease, these differences were evidently attributable to other factors, whose investigation is certainly one of the possible future directions for research.

Acknowledgments

The authors have no acknowledgments to include.

Conflict of Interest Disclosure

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors declare no sources of funding.

Research Ethics Statement

The study was approved by the IRCCS Don Carlo Gnocchi Foundation Ethics Committee, and it was performed in accordance with the principles of the Helsinki Declaration. All the participants provided their written and informed consent.

References

Armistead, T. W. (2013). HL Wagner's unbiased hit rate and the assessment of categorical forecasting accuracy. *Weather and Forecasting, 28*(3), 802–814. https://doi.org/10.1175/WAF-D-12-00047.1

Anolli, L., & Ciceri, R. (1992). *La voce delle emozioni*. Franco Angeli.

Anolli, L., Wang, L., Mantovani, F., & De Toni, A. (2008). The voice of emotion in Chinese and Italian young adults. *Journal of Cross-Cultural Psychology, 39*(5), 565–598. https://doi.org/10.1177/0022022108321178

Benke, T., Bösch, S., & Andree, B. (1998). A study of emotional processing in Parkinson's disease. *Brain and Cognition, 38*(1), 36–52. https://doi.org/10.1006/brcg.1998.1013

Biassoni, F., Balzarotti, S., Giamporcaro, M., & Ciceri, R. (2016). Hot or cold anger? Verbal and vocal expression of anger while driving in a simulated anger-provoking scenario. *Sage Open, 6*(3), 2158244016658084

Blonder, L. X., Gur, R. E., & Gur, R. C. (1989). The effects of right and left hemiparkinsonism on prosody. *Brain and Language, 36*(2), 193–207. https://doi.org/10.1016/0093-934X(89)90061-8

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program]. Version 6.0. 37.

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E., & Diller, L. (1990). Parameters of emotional processing in neuropsychiatric disorders: Conceptual issues and a battery of tests. *Journal of Communication Disorders, 23*(4–5), 247–271. https://doi.org/10.1016/0021-9924(90)90003-H

Caekebeke, J. F., Jennekens-Schinkel, A., Van der Linden, M. E., Buruma, O. J., & Roos, R. A. (1991). The interpretation of dysprosody in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 54*(2), 145–148. https://doi.org/10.1136/jnnp.54.2.145

Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2002). Rey-Osterrieth complex figure: normative values in an Italian population sample. *Neurological Sciences, 22*(6), 443–447. https://doi.org/10.1007/s100720200003

Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2004). Modified card sorting test: normative data. *Journal of Clinical and Experimental Neuropsychology, 26*(2), 246–250. https://doi.org/10.1076/jcen.26.2.246.28087

Cheang, H. S., & Pell, M. D. (2007). An acoustic investigation of Parkinsonian speech in linguistic and emotional contexts. *Journal of Neurolinguistics, 20*(3), 221–241. https://doi.org/10.1016/j.jneuroling.2006.07.001.

Darkins, A. W., Fromkin, V. A., & Benson, D. F. (1988). A characterization of the prosodic loss in Parkinson's disease. *Brain and Language, 34*(2), 315–327. https://doi.org/10.1016/0093-934X(88)90142-3

De Almeida, S. C., Mendes, A. P., & Kempster, G. B. (2019). The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) Psychometric Characteristics: II European Portuguese Version (II EP CAPE-V). *Journal of*

Voice, 33(4), 582-e5. https://doi.org/10.1016/j.jvoice.2018.02.013

Defazio, G., Guerrieri, M., Liuzzi, D., Gigante, A. F., & Di Nicola, V. (2016). Assessment of voice and speech symptoms in early Parkinson's disease by the Robertson dysarthria profile. *Neurological Sciences, 37*(3), 443–449. https://doi.org/10.1007/s10072-015-2422-8

Dykstra, A. D., Adams, S. G., & Jog, M. (2015). Examining the relationship between speech intensity and self-rated communicative effectiveness in individuals with Parkinson's disease and hypophonia. *Journal of Communication Disorders, 56*, 103–112. https://doi.org/10.1016/j.jcomdis.2015.06.012

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*, 268–287. https://doi.org/10.1037/0033-2909.115.2.268

Fatigante, M., Biassoni, F., Marazzini, F., & Diadori, P. (2016). Responsibility and culpability in apologies: Distinctive uses of " sorry" versus " I'm sorry" in apologizing. *Discourse Processes, 53*(1–2), 26–46. https://doi.org/10.1080/0163853X.2015.1056696

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": Apractical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Frasson, P., Ghiretti, R., Catricalà, E., Pomati, S., Marcone, A., Parisi, L., ... & Clerici, F. (2011). Free and cued selective reminding test: an Italian normative study. *Neurological Sciences, 32*(6), 1057–1062. https://doi.org/10.1007/s10072-011-0607-3

Galaz, Z., Mekyska, J., Mzourek, Z., Smekal, Z., Rektorova, I., Eliasova, I., Kostalova, M. & Berankova, D. (2016). Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. *Computer Methods and Programs in Biomedicine, 127*, 301–317. https://doi.org/10.1016/j.cmpb.2015.12.011

Gerratt, B. R., Till, J. A., Rosenbek, J. C., Wertz, R. T., & Boysen, A. E. (1991). Use and perceived value of perceptual and instrumental measures in dysarthria management. In C. A. Moore, K. M. Yorkston, & D. R. Beukelman (Eds.), *Dysarthria and apraxia of speech: Perspectives on management* (pp. 77–93). Brookes.

Gillivan-Murphy, P., Miller, N., & Carding, P. (2019). Voice treatment in Parkinson's disease: patient perspectives. *Research and Reviews in Parkinsonism, 9*, 29–42. https://doi.org/10.2147/JPRLS.S180183

Goetz, C. G., Poewe, W., Yahr, M. D., Seidl, L., Rascol, O., Sampaio, C., Stebbins, G. T., Counsell, C., Giladi, N., Holloway, R. G., Moore, C. G., Wenning, G. K., & Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. (2004). Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations. *Movement Disorders, 19*(9), 1020–1028. https://doi.org/10.1002/mds.20213

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M. B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., . . . Movement Disorder Society UPDRS Revision Task Force. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders, 23*(15), 2129–2170. https://doi.org/10.1002/mds.22340

Harris, R., Leenders, K. L., & de Jong, B. M. (2016). Speech dysprosody but no music 'dysprosody'in Parkinson's disease. *Brain and Language, 163*, 1–9. https://doi.org/10.1016/j.bandl.2016.08.008

Hsu, C. C. T. (2016). *Production and perception of emotional prosody: the case of autism and Parkinson's disease* [Doctoral dissertation]. University College London.

Hughes, A. J., Ben-Shlomo, Y., Daniel, S. E., & Lees, A. J. (1992). UK Parkinson's disease society brain bank clinical diagnostic criteria. *Journal of Neurology, Neurosurgery, and Psychiatry, 55*(181), e4. https://doi.org/10.1136/jnnp.55.3.181

Hustad, K. C. (1999). Optimizing communicative effectiveness: Bringing it together. In K. Yorkston, D. Beukelman, E. A. Strand, & K. R. Bell (Eds.), *Management of motor speech disorders in children and adults* (pp. 483–542). Pro-Ed.

Jaywant, A., & Pell, M. D. (2010). Listener impressions of speakers with Parkinson's disease. *Journal of the International Neuropsychological Society, 16*(1), 49–57. https://doi.org/10.1017/S1355617709990919.

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 220–235). Guilford.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* Sage.

Kwan, L. C., & Whitehill, T. L. (2011). Perception of speech by individuals with Parkinson's disease: A review. *Parkinson's Disease, 201*. https://doi.org/10.4061/2011/389767

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods, 45*, 1234–1245. https://doi.org/10.3758/s13428-013-0324-3

Mahler, L. A., Ramig, L. O., & Fox, C. (2015). Evidence-based treatment of voice and speech disorders in Parkinson disease. *Current Opinion in Otolaryngology & Head and Neck Surgery, 23*(3), 209–215. https://doi.org/10.1097/MOO.0000000000000151

Martens, H., Van Nuffelen, G., Cras, P., Pickut, B., De Letter, M., & De Bodt, M. (2011). Assessment of prosodic communicative efficiency in Parkinson's disease as judged by professional listeners. *Parkinson's Disease*, *2011*, 1–10. https://doi.org/10.4061/2011/129310

Miller, K. M., Okun, M. S., Marsiske, M., Fennell, E. B., & Bowers, D. (2009). Startle reflex hyporeactivity in Parkinson's disease: an emotion-specific or arousal-modulated deficit?. *Neuropsychologia, 47*(8–9), 1917–1927. https://doi.org/10.1016/j.neuropsychologia.2009.03.002

Miller, N. (2017). Communication changes in Parkinson's disease. *Practical Neurology, 17*(4), 266–274. https://doi.org/DOI:10.1136/practneurol-2017-001635

Möbes, J., Joppich, G., Stiebritz, F., Dengler, R., & Schröder, C. (2008). Emotional speech in Parkinson's disease. *Movement Disorders, 23*(6), 824–829. https://doi.org/10.1002/mds.21940

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurological Sciences, 34*(5), 749–754. https://doi.org/10.1007/s10072-012-1130-x

Oates, J. (2009). Auditory-perceptual evaluation of disordered voice quality. *Folia Phoniatrica et Logopaedica, 61*(1), 49–56. https://doi.org/10.1159/000200768

Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language, 97*(2), 123–134. https://doi.org/10.1016/j.bandl.2005.08.010

Penner, H., Miller, N., Hertrich, I., Ackermann, H., Schumm, F. (2001). Dysprosody in Parkinson's disease: An investigation of intonation patterns. *Clinical Linguistics & Phonetics, 15*(7), 551–566. https://doi.org/10.1080/02699200110078140

Phi, H., Janarthanan, S., Zhang, L., & Ghomi, R. H. (2019). Voice biomarker identification for effects of deep-brain stimulation on parkinson's disease. *arXiv*:1912.00866.

Pitcairn, T. K., Clemie, S., Gray, J. M., & Pentland, B. (1990). Impressions of parkinsonian patients from their recorded voices. *British Journal of Disorders of Communication, 25*(1), 85–92. https://doi.org/10.3109/13682829009011965

Raven, J., & Raven, J. (2003). Raven progressive matrices. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 223–237). Springer.

Reitan, R. M. (1971). Trail making test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33*(2), 575–581. https://doi.org/10.2466/pms.1971.33.2.575

Russell, J. A. (1993). Forced-choice response format in the study of facial expression. *Motivation and Emotion, 17*, 41–51. https://doi.org/10.1007/BF00995206

Santos, L. L. M., Reis, L. O. D., Bassi, I., Guzella, C., Cardoso, F., Reis, C., & Gama, A. C. C. (2010). Acoustic and hearing-perceptual voice analysis in individuals with idiopathic Parkinson's disease in" on" and" off" stages. *Arquivos de Neuro-Psiquiatria, 68*(5), 706–711.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences, 107*(6), 2408–2412. https://doi.org/10.1073/pnas.0908239106

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*(3), 235–248. https://doi.org/10.1016/S0892-1997(05)80231-0

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion, 15*(2), 123–148. https://doi.org/10.1007/BF00995674

Scherer, Klaus R., Banse, Rainer, Wallbott, Harald G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*(1), 76–92. https://doi.org/10.1177/0022022101032001009

Schroder, M. (2003). Experimental study of affect bursts. *Speech Communication, 40*, 99–116. https://doi.org/10.1016/S0167-6393(02)00078-X

Schröder, C., Nikolova, Z. T., & Dengler, R. (2010). Changes of emotional prosody in Parkinson's disease. *Journal of the Neurological Sciences, 289*(1–2), 32–35. https://doi.org/10.1016/j.jns.2009.08.038

Sperber, D., & Wilson D. (2002). Pragmatics, modularity and mind-reading. *Mind and Language, 17*, 3–23. https://doi.org/10.1111/1468-0017.00186

Spinnler H., & Tognoni G. (1987). Standardizzazione e taratura italiana di test neuropsicologici: Gruppo italiano per lo studio neuropsicologico dell'invecchiamento. *Italian Journal of Neurological Sciences 6*(Suppl. 8), 8–120.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651

Tjaden, K., & Wilding, G. (2011). Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and multiple sclerosis. *Journal of Communication Disorders, 44*(6), 655–665. https://doi.org/10.1016/j.jcomdis.2011.06.003

Town, S., Donovan, M. C., & Beach, E. (2020). A "gestalt" framework of emotions and organizing: integrating innate, constructed, and discursive ontologies. *Management Learning, 52*(5), 519–540. https://doi.org/10.1177/1350507620972238

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17*(1), 3–28. https://doi.org/10.1007/BF00987006

Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.

Wieser, M. J., Mühlberger, A., Alpers, G. W., Macht, M., Ellgring, H., & Pauli, P. (2006). Emotion processing in Parkinson's disease: dissociation between early neuronal processing and explicit ratings. *Clinical Neurophysiology, 117*(1), 94–102. https://doi.org/10.1016/j.clinph.2005.09.009

World Health Organization. (2001). *The World Health Report 2001: Mental health: New understanding, new hope*. World Health Organization.