Acoustic Investigation of Stress Patterns in Parkinson's Disease

***Tereza Tykalova, *,†Jan Rusz, *Roman Cmejla, †Hana Ruzickova, and †Evzen Ruzicka,** *†*Prague, Czech Republic*

Summary: Objectives. Although reduced stress is thought to be one of the most deviant speech dimensions in hypokinetic dysarthria associated with Parkinson's disease (PD), the mechanisms of stress production in PD have not been thoroughly explored by objective methods. The aim of the present study was to quantify the effect of PD on prosodic characteristics and to describe contrastive stress patterns in parkinsonian speech.

Methods. The ability of 20 male speakers with early PD and 16 age- and gender-matched healthy controls (HCs) to signal contrastive stress was investigated. Each participant was instructed to unnaturally emphasize five key words while reading a short block of text. Acoustic analyses were based on the measurement of pitch, intensity, and duration. In addition, an innovative measurement termed the stress pattern index (SPI) was designed to mirror the effect of all distinct acoustic cues exploited during stress production.

Results. Although PD patients demonstrated a reduced ability to convey contrastive stress, they could still notably increase pitch, intensity, and duration to emphasize a word within a sentence. No differences were revealed between PD and HC stress productions using the measurements of pitch, intensity, duration, and intensity range. However, restricted SPI and pitch range were evident in the PD group.

Conclusions. A reduced ability to express stress seems to be the distinctive pattern of hypokinetic dysarthria, even in the early stages of PD. Because PD patients were able to consciously improve their speech performance using multiple acoustic cues, the introduction of speech therapy may be rewarding.

Key Words: Parkinson's disease–Speech disorders–Reduced stress–Acoustic analysis–Prosody–Contrastive stress.

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the progressive loss of dopaminergic neurons in the substantia nigra, affecting 1–2% of persons over the age of 60 years.[1,2] In addition to cardinal motor manifestations, such as bradykinesia, rigidity, postural instability, and resting tremor, up to 90% of individuals with PD develop an alteration of speech termed hypokinetic dysarthria.[3,4] Moreover, these distinctive speech deficits may be one of the earliest symptoms and appear even several years before the diagnosis is established.[5] Speech deficits commonly reported to be experienced by PD individuals include monoloudness, monopitch, reduced stress, imprecise articulation, variability of speech rate, a breathy and harsh voice, disfluency, voice tremor, and other manifestations that can lead to overall reduced speech intelligibility.[6,7] These changes in speech production may have a significant, negative impact on social interactions and overall quality of the patient's life.[8]

Speech is a unique, complex, dynamic motor activity through which individuals express their thoughts and feelings.[7] From an acoustic perspective, speech can be surveyed with respect to five speech subsystems including respiration, phonation, resonance, articulation, and prosody. Prosody itself is an important aspect of language that is necessary for recovering the intended meanings of an utterance, that is, information that is unavailable in the orthographic transcription. In particular, prosody may serve a variety of functions, including signaling questions or lexical boundaries, conveying contrastive meanings, and expressing emotions and attitudes.[9–11] One of the techniques used by speakers to convey these suprasegmental features is word and sentence stress, representing the relative emphasis given to a certain syllable or word.

In acoustics, there are three main prosodic cues commonly associated with stress: pitch, intensity, and duration.[12–14] In the 1960s, studies of healthy speakers established pitch prominence as the primary marker of stress.[12,15–17] For example, Fry[15] measured pitch and duration changes in lexical stress pairs (eg, HOTdog vs hot DOG) and found pitch to be superior to duration. Another experiment conducted by Fry[13] to determine whether intensity or duration was a better cue to stress showed that duration, on the whole, was a more distinctive cue. On the other hand, some researchers have argued that duration and/or intensity also convey stress and might be at least as important as pitch.[9,10,18,19] Moreover, the way stress manifests itself in the speech stream is partially language dependent.[20,21] From this point of view, the prosodic characteristics of stress are somewhat ambiguous, even considering nonimpaired speech.

Although the manifestation of reduced stress has been well documented in several motor speech disorders,[7,22–24] the mechanisms of stress production in hypokinetic dysarthria of PD have not been thoroughly explored by objective methods. Ma et al[25] analyzed question-statement contrast in 14 Cantonese PD speakers and found that subjects with PD used similar acoustic cues as healthy adults; however, adequate contrast was not observed in all speakers. Cheang and Pell[26] reported

Accepted for publication July 2, 2013.

This research was partly supported by the Czech Science Foundation (GACR 102/12/2230), Czech Ministry of Health (NT 12288-5/2011 and NT14181-3/2013), Czech Ministry of Education (MSM 0021620849), and Charles University in Prague (PRVOUK-P26/LF1/4).

From the *Department of Circuit Theory, Czech Technical University in Prague, Faculty of Electrical Engineering, Prague, Czech Republic; and the †Department of Neurology and Centre of Clinical Neuroscience, Charles University in Prague, First Faculty of Medicine, Prague, Czech Republic

Address correspondence and reprint requests to Jan Rusz, Department of Circuit Theory, Czech Technical University in Prague, Faculty of Electrical Engineering, Technicka 2, 166 27, Prague 6, Czech Republic. E-mail: ruszjan@fel.cvut.cz

Journal of Voice, Vol. 28, No. 1, pp. 129.e1-129.e8

0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2013.07.001

that medicated patients in the early stages of PD exhibit various changes in the speaking tasks of lexical stress, contrastive stress, and emotional prosody. The acoustic results showed that the average amplitude measurement was the most robust parameter, as tokens elicited from PD speakers were lower in amplitude than tokens spoken by healthy participants in all three production tasks. Meanwhile, pitch was found to be aberrant among PD speakers for contrastive stress and emotional prosody; duration measures revealed anomalies between healthy and PD speakers merely in phonemic stress.[26] In addition, there is also some evidence that the ability of PD speakers to express intended stress or emotions through prosody is often poorly understood by listeners.[11]

Reduced stress is thought to be the second most deviant speech dimension in hypokinetic dysarthria[6]; therefore, the effect of PD on prosodic characteristics and the detailed description of stress patterns in parkinsonian speech are of principal concern in this investigation. To further examine this issue, we chose a speaking task of *contrastive stress* because lexical stress is not inherent in all languages. Contrastive stress refers to a production task in which the information conveyed is altered by the location of syllabic stress. As acoustic analysis has the potential to provide a cheap, precise, and noninvasive method for the evaluation and support of speech therapy, further aims of the present study were to verify the suitability of commonly used measurements for the evaluation of stress in PD and to design an innovative measurement that would reflect the effect of all main acoustic cues exploited during stress production. We hypothesized that PD subjects would have a reduced ability to convey contrastive stress and show abnormal patterns of pitch, amplitude, and duration in both stressed and nonstressed tokens.

PATIENTS AND METHODS

The participants of this study were originally recruited as a part of an earlier study.[27] No study of contrastive stress has been previously published on the current participants. A total of 36 male Czech native speakers volunteered for the study. The PD group consisted of 20 individuals with idiopathic PD, whose age ranged from 34 to 82 years (mean = 60.5; standard deviation [SD] = 11.3). The diagnosis of idiopathic PD was made in a specialized center and was based on accepted criteria.[28] All patients were recruited immediately after the diagnosis was established and before symptomatic treatment was started. Before the recording procedure, each patient underwent a neurologic examination including the Unified Parkinson's Disease Rating Scale part III (UPDRS III, an objective measure of parkinsonian motor signs, ranging from 0 to 108, where a higher score indicates more severe disability), and Hoehn and Yahr (HY) staging (ranging from 1 to 5, where a higher stage indicates more severe disability). In our patients, the UPDRS III score ranged from 5 to 32 (mean = 17.8; SD = 7.2) and the HY ranged from 1 to 3 (mean = 2.2; SD = 0.5). In addition, the estimated duration of PD manifestations before the examination was surveyed and ranged from 6 to 82 months (mean = 31.9; SD = 21.4). No patient had a history of speech, language, or hearing disorders unrelated to parkinsonian symptoms or underwent speech-language treatment before participation in this study. All subjects were free of depression and cognitive deficits that could interfere with the measurements.

The healthy control (HC) group consisted of 16 male participants of comparable age, ranging from 36 to 80 years (mean = 61.8; SD = 13.3). None of these individuals reported a history of neurologic difficulties or any disorders that may affect speech, language, or hearing. No significant differences in age distribution were detected between the PD and HC groups. The study was approved by the local ethics committee and all participants provided written, informed consent for the speaking task and recording procedure.

The recordings were obtained during one session with a speech therapist who conveyed instructions to the subjects. Each participant completed a series of speaking tasks as part of the larger protocol. There were no time limits during the recordings. All participants were asked to repeat their performance at any time if they or the examiner were not fully satisfied with their initial attempt. The performance of the task including contrastive stress was selected for further investigation. The task was designed to evaluate whether speakers could highlight the semantic importance of information in their utterances using prosody. During the recording, each patient read a short block of text composed of five similar sentences and was required to unnaturally emphasize certain "key words" included in the text (Table 1). The first part of each sentence was variable and determined the linguistic context to identify the key word in the second part of the sentence, which was uniform and highlighted one of five key words (eg, *Dnes jsme to již nestihli*, možná ZÍTRA navštívíme všechny své známé./*Today we did not have enough time* but TOMORROW we will visit all our acquaintances; the part that determined the linguistic context of the sentence is indicated in italics). To ensure a better understanding of the task, the key words were underlined and written in capital letters and the entire task was demonstrated by a speech therapist (H.R.). As a result, for five different key words (ie, zítra/tomorrow; navštívíme/visit; všechny/all; své/our; and známé/acquaintances), we elicited one emphasized and four normally read tokens that were subjected to further investigation.

The speech samples were recorded in a quiet room with a low level of ambient noise using an external condenser microphone placed approximately 15 cm from the subject's mouth and coupled to a Panasonic NV-GS 180 video camera (Panasonic Corporation, Osaka, Japan). The audio data were digitized from the videotape to a computer at a sampling rate of 48 kHz and 16-bit quantization using original *Panasonic* software (Panasonic Corporation, Osaka, Japan).

Acoustic analysis was completed using the widely used specialized speech software *PRAAT* [available at: www.praat.org (Phonetic Sciences, University of Amsterdam, The Netherlands)].[29] To ensure the correctness of the automatic detecting procedure, the results of automatic analysis were verified by the examiner (T.T.) and manually adjusted if necessary. For the entire duration of each token, three standard acoustic parameters were assigned: fundamental frequency (F_0) in hertz; intensity in decibels; and duration in milliseconds. F_0 as well as intensity were expressed as the mean and range, that is, the difference between the maximum and minimum values. Duration was