Physiology & Behavior 163 (2016) 234–238

Contents lists available at ScienceDirect

Physiology & Behavior

journal homepage: www.elsevier.com/locate/phb

Individual differences in cortisol stress response predict increases in voice pitch during exam stress

Katarzyna Pisanski [a,*], Judyta Nowak [b], Piotr Sorokowski [a]

[a] Institute of Psychology, University of Wrocław, Poland
[b] Department of Human Biology, University of Wrocław, Poland

HIGHLIGHTS

- The finding that voice pitch increases under stress is often not replicated.
- We tested whether this is tied to variation in the severity of the stress response.
- Voice recordings and saliva were collected at baseline and during exam stress.
- Increases in cortisol levels predicted voice pitch under stress but not at baseline.
- Researchers examining speech under stress should control for the stress response.

ARTICLE INFO

Article history:
Received 23 March 2016
Received in revised form 11 May 2016
Accepted 12 May 2016
Available online 14 May 2016

Keywords:
Speech under stress
Fundamental frequency
Cortisol
Psychological stress
Exam stress

ABSTRACT

Despite a long history of empirical research, the potential vocal markers of stress remain unclear. Previous studies examining speech under stress most consistently report an increase in voice pitch (the acoustic correlate of fundamental frequency, *F*0), however numerous studies have failed to replicate this finding. In the present study we tested the prediction that these inconsistencies are tied to variation in the severity of the stress response, wherein voice changes may be observed predominantly among individuals who show a cortisol stress response (i.e., an increase in free cortisol levels) above a critical threshold. Voice recordings and saliva samples were collected from university psychology students at baseline and again immediately prior to an oral examination. Voice recordings included both read and spontaneous speech, from which we measured mean, minimum, maximum, and the standard deviation in *F*0. We observed an increase in mean and minimum *F*0 under stress in both read and spontaneous speech, whereas maximum *F*0 and its standard deviation showed no systematic changes under stress. Our results confirmed that free cortisol levels increased by an average of 74% (ranging from 0 to 270%) under stress. Critically, increases in cortisol concentrations significantly predicted increases in mean *F*0 under stress for both speech types, but did not predict variation in *F*0 at baseline. On average, stress-induced increases in voice pitch occurred only when free cortisol levels more than doubled their baseline concentrations. Our results suggest that researchers examining speech under stress should control for individual differences in the magnitude of the stress response.

1. Introduction

Studies examining the acoustic correlates of stress have produced widely mixed results (reviewed in [12,18]). In large part these inconsistencies stem from differences in how stress is defined [28], difficulties related to inducing and measuring genuine stress responses in the laboratory [2,22,31,33,37], and substantial variation in the magnitude of the stress response across individuals [2,17,21,35]. Exam stress, the most significant source of stress experienced by students, peaking just prior to an academic examination [30], offers an ecologically valid context in which to examine the effects of psychological stress on the voice. Psychological stress increases the activity of the hypothalamic adrenal (HPA) axis and circulating stress hormone levels, particularly cortisol [7,22], and is known to have multiple and apparent effects on the body including increased breathing rate, muscle tension, and changes in salivation rate, that may in turn affect vocal production ([18,32,33]).

Voice pitch (the acoustic correlate of fundamental frequency, *F*0) is inversely related to the rate of vocal fold vibration, and thus increases as the vocal folds stretch and become tenser or when sub-glottal pressure and vocal intensity increase [16,36], as often occurs under

* Corresponding author at: Mammal Vocal Communication and Cognition Laboratory, School of Psychology, University of Sussex, Brighton, United Kingdom.
E-mail address: K.Pisanski@sussex.ac.uk (K. Pisanski).

http://dx.doi.org/10.1016/j.physbeh.2016.05.018
0031-9384/

psychological stress. Indeed, an increase in voice pitch is the longest and most commonly reported finding in previous studies examining speech under stress (reviewed in [12,18]). However, many studies have failed to replicate this finding (e.g., [8,15,17,35,37,38]). Others report an increase in minimum voice pitch or a decrease in its standard deviation (*F*0 *sd*), with no systematic change in mean pitch [29,37]. Previous studies also consistently indicate that voice pitch increases more in natural than artificially induced stress scenarios [18,31]. It is possible that there are no reliable acoustic markers of psychological stress [35]. Alternatively, as has been suggested by many researchers, these inconsistencies in previous work may be tied to variation in the severity of the stress response, wherein vocal changes may be observed only in particular contexts (e.g., natural stress scenarios) or among individuals surpassing a critical threshold of stress [1,17,18,33]. Despite this parsimonious possibility, only a handful of studies examining speech under stress have controlled for individual differences in stress levels [15,34] or in participants' propensity toward anxiety [1,37].

The aim of the present study was to test whether individual differences in the stress response, quantified by free cortisol levels measured from saliva, predict the magnitude of voice pitch changes under stress. To test this prediction we collected voice recordings and saliva samples from undergraduate psychology students two weeks prior to (baseline session) and immediately prior to (stress session) an academic oral examination, and compared changes in free cortisol levels and voice pitch at both times. Voice recordings included both read and spontaneous speech from which we measured mean, minimum, maximum, and the standard deviation in fundamental frequency.

2. Methods and materials

2.1. Participants

Thirty-four students took part i 15 or more n the study (aged 21–32, $M = 22.7$, $sd = 2.0$, all female). All students were registered in an upper-level psychology course and completed an oral examination as a course requisite. Students who volunteered to take part in the study were informed that voice recordings and saliva samples would be collected approximately two weeks before and on the day of the oral examination, and that their choice to take part in the study would in no way affect their exam grade. Participants completed a prescreening questionnaire. On this basis, data from 3 participants were excluded from the study for one or more of the following reasons: the participant was taking steroids such as allergy medications containing corticosteroids, smoked 15 or more cigarettes per day, was pregnant or breastfeeding, and/or had a disease of the mouth or chronic illness (see [21]). All participants provided written informed consent to participate.

2.2. Procedure

Voice recordings and saliva samples were collected from each participant in two separate sessions (baseline and stress) following a similar procedure (see Data collection). Baseline sessions took place approximately two weeks ($M = 12.3$, $sd = 2.2$ days) before the students' oral examination was scheduled to take place. During the baseline session, each participant provided a voice recording, completed a demographic questionnaire, and provided a baseline saliva sample 15 min afterward. Stress sessions took place on the day of the oral exam. Upon arrival, participants blindly drew three exam questions out of a jar, and were given 10 min to prepare their oral responses to these questions. As previous studies indicate that exam stress peaks immediately before, and declines during the exam [11,26], participants' voices were recorded immediately before the oral exam. Oral exams took place in a private room with the instructor and lasted approximately 15 min, immediately after which each participant provided a saliva sample.

It takes approximately 15 min for the production of free cortisol by the adrenal glands to manifest itself in saliva following psychological stress [19], therefore saliva collection followed voice recording by 15 min. Cortisol responses to social stress are largely unaffected by time of day [4,24]. However, all sessions were scheduled in the afternoon to control for the awakening cortisol response and diurnal fluctuations in cortisol levels [9,23]. Within participants, baseline and stress sessions were also scheduled within 0–90 min ($M = 19.4$, $sd = 40.1$ min) of the same time of day. Participants confirmed having not consumed any food, caffeine, nicotine, vitamins or medication, having not brushed their teeth, and having not engaged in any form of rigorous exercise within 2 h of each session [21].

The study was performed in accordance with the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki) and was approved by the University of Wroclaw Institutional Review Board.

2.3. Data collection

2.3.1. Voice recording

In both baseline and stress sessions, voice recordings were conducted in the same quiet room using an M-Audio condenser microphone with a cardioid pick-up pattern and at a distance of 5–10 cm. Participants were asked to describe their studies (e.g., subject area or major, year of degree, and department). This constituted the spontaneous speech condition. Participants were also asked to familiarize themselves with and subsequently read the first five sentences of the Rainbow Passage ([10]; Polish translation, see Appendix A). This constituted the read speech condition. Audio was digitally encoded using an M-Audio Fast Track ultra interface at a sampling rate of 44.1 kHz and 16-bit amplitude quantization and stored onto a computer as WAV files.

2.3.2. Saliva collection

Saliva samples were collected in duplicate or triplicate from each participant during each session into 2 ml polypropylene microtubes (SARSTEDT®) using the passive drool method [13]. Sugar-free gum was used to stimulate saliva flow (Salimetrics LLC, State College, PA, USA). Saliva samples were immediately frozen at −25 °C until being transported to the Institute of Genetics and Microbiology, where the samples were kept at a temperature of −70 °C until analysis.

2.4. Data analysis

2.4.1. Voice measurement

Acoustic measurements were performed in Praat [3] while blind to speaker identity and session. For each voice recording (4 per participant), we measured voice pitch as mean fundamental frequency (*F*0 mean), minimum fundamental frequency (*F*0 min), maximum fundamental frequency (*F*0 max), and the standard deviation in fundamental frequency (*F*0 *sd*). All measures were taken from voiced speech segments only. Spontaneous speech recordings were analyzed in full (duration $M = 4.9$, $sd = 1.4$ s). To analyze a comparable duration of read speech, we selected the central sentence of the Rainbow Passage (see Appendix A). All *F*0 parameters were measured using Praat's autocorrelation algorithm with a search range set to 100–600 Hz [3].

2.4.2. Hormone measurement

Hormone concentrations were measured at the Institute of Genetics and Microbiology. Prior to analysis, frozen saliva samples were brought to room temperature and then centrifuged for 10 min in order to separate mucins. Clear colorless supernatant from samples taken from the same participant and condition were transferred into a separate sampling device, mixed on an orbital shaker and used for testing. Active free cortisol levels were measured using enzyme-linked immunosorbent assay (ELISA) and commercial kits (DEMEDITEC®, Germany, cat. number DES6611). Mixed saliva samples were assayed in duplicate following the manufacturer's instructions. Concentrations were calculated in relation to the standard curve and are expressed in ng/ml. The