Analogy Instruction and Speech Performance Under Psychological Stress

***,†Andy C. Y. Tse, *,†Andus W.-K. Wong, †Tara L. Whitehill, †Estella P.-M. Ma, and *Rich S. W. Masters,** **†Pokfulam, Hong Kong*

Summary: To examine the efficacy of explicit and implicit forms of instruction for speech motor performance under conditions of psychological stress. In experiment 1, 20 participants were asked to deliver a formal presentation to validate the modified Trier Social Stress Test (TSST). In experiment 2, 40 participants were instructed explicitly by verbal explanation or implicitly by analogy to speak with minimum pitch variation and were subjected to psychological stress using the modified TSST. Acoustic correlates of pitch height (mean fundamental frequency) and pitch variation (standard deviation of fundamental frequency) significantly increased in experiment 1 when participants delivered a speech under modified TSST condition. In experiment 2, explicitly instructed participants were unable to maintain minimum pitch variation under psychological pressure caused by the modified TSST, whereas analogy-instructed participants maintained minimal pitch variation. The findings are consistent with existing evidence that analogy instructions may result in characteristics of implicit motor learning, such as greater stability of performance under pressure. Analogy instructions may therefore benefit speech motor performance and might provide a useful clinical tool for treatment of speech-disordered populations.

Key Words: Analogy instruction–Speech motor performance–Pitch variation–Implicit motor learning.

INTRODUCTION

Evidence from the speech science literature suggests that psychological stress has a direct influence on speech production and voice quality.[1–5] For example, it is observed that people tended to display increased pitch height (ie, tonal height) when they spoke under stress.[6] Experiencing stress for a prolonged period can contribute to physical musculoskeletal disorders and functional voice disorders.[7–10] For instance, people with a general predisposition for high anxiety tended to develop hypertension in the laryngeal muscles, which resulted in dysphonia.[9] Moreover, stress also contributes to disrupted voice quality in individuals with Parkinson disease by lowering intonation and volume.[11–14] Due to the potential impact of psychological stress on speech, knowledge of vocal stress phenomena and methods of stress prevention may assist speech-language pathologists to design treatment programs for their clients.

From a physiological perspective, many vocal stress phenomena are associated with poor motor control and coordination in oral muscles such as the lips, jaw, and tongue.[5] For example, it is suggested that poorly coordinated tongue muscles result in vigorous and more rapid muscular activities when producing the target sound, which therefore affect the articulation of vowels and consonants.[8] Also, increasing muscle tension in the perilaryngeal region due to psychological stress causes tensing in the vocal folds, which can cause an elevation in the fundamental frequency (F_0) of speech (ie, a higher pitch voice).[15] Such vocal stress responses have also been reported in other studies.[4,16–18]

One way in which psychological stress is thought to disrupt movements is through reinvestment. According to the Theory of Reinvestment,[18,19] the processing of explicit movement-related knowledge makes high demands on working memory capacity.[20,21] Performance can be adversely affected if it must take place in a situation that further increases the demands on working memory (eg, psychological stress). Previous studies have suggested, however, that movements acquired implicitly, without processing of explicit movement-related knowledge, are more resistant to adverse situations such as psychological stress.[22–25] An essential feature of this type of learning, *implicit motor learning*, is that movement acquisition takes place without conscious accrual of, or awareness of, movement-related knowledge.[24] Performance is therefore less dependent on working memory and more resistant to decrement in adverse contexts such as psychological stress.

One paradigm that has been used to bring about implicit motor learning and which holds potential value in speech motor learning is analogy learning.[26–29] Analogy instructions are thought to encourage implicit motor learning by integrating many discrete "bits" of movement-related knowledge into a single memory representation, thus requiring little conscious processing and minimizing demands on working memory.[22,25,27,28] For example, table tennis novices were taught to use a topsin forehand stroke via a "right-angled triangle" analogy in which participants were instructed only to move the table tennis bat up the hypotenuse of the triangle as they hit the ball.[27] In contrast to explicitly instructed participants or controls, analogy-instructed participants displayed minimal explicit knowledge of the task and more robust performance under psychological stress.[23,29]

Recently, analogy instructions have also been investigated in speech motor performance. Researchers asked participants to modulate their intonation (lowering pitch variation) when reading aloud a passage. This was done by providing an analogy

Accepted for publication March 27, 2013.
The research was supported by the Sciences of Learning Strategic Research Theme of the University of Hong Kong.
From the *Institute of Human Performance, The University of Hong Kong, Pokfulam, Hong Kong; and the †Division of Speech and Hearing Sciences, The University of Hong Kong, Pokfulam, Hong Kong.
Address correspondence and reprint requests to Andy C.Y. Tse, Division of Speech and Hearing Sciences, Institute of Human Performance, The University of Hong Kong, Pokfulam, Hong Kong. E-mail: tcyandy@yahoo.com.hk
Journal of Voice, Vol. 28, No. 2, pp. 196-202
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2013.03.014

instruction ("speak like a calm sea") or by providing explicit instructions ("speak with no pitch variation. That is, speak without changes in the highness or lowness of your voice"). Evidence from acoustic measures indicated that analogy-instructed participants were more able to achieve the targeted speech motor performance (ie, speaking with minimum pitch variation) than explicitly instructed participants.[30] In a subsequent study, evidence from perceptual measures showed that analogy instructions resulted in speech that was perceived to be more natural than when explicit instructions were used.[31]

Given that stress-resistant characteristics are displayed by analogy learners when moving for sport, it is speculated that this advantage may generalize to moving for speech. The aim of the present study was to test this speculation. We first verified the commonly held assumption that psychological stress causes an increase in pitch height or pitch variation, denoted by an increase in mean F_0 (an acoustic correlate of pitch height) and an increase in standard deviation of fundamental frequency (SDF_0) (an acoustic correlate of pitch variation), respectively. We then examined whether speech performance modified by an analogy learning instruction remained stable under psychological stress compared with speech modified by an explicit learning instruction.

EXPERIMENT ONE

To examine the influence of psychological stress on pitch height and variation, we first modified the Trier Social Stress Test (TSST),[32] which is a standardized stress reactivity protocol commonly used to induce moderate psychological stress in a laboratory setting.[7,32] Participants were required to complete a series of speaking tasks in conditions of stress and no stress (verified by assessing changes in heart rate and anxiety).

Method

Participants. Twenty healthy native Cantonese speakers (10 females and 10 males; mean age 20.05 years, standard deviation [SD] 2.06) without any formal presentation or singing training (self-reported) were recruited from the University of Hong Kong. Participants were instructed to read aloud the standardized Cantonese passage "*North Wind and the Sun*"[33] using their habitual speaking voice. The passage is commonly used in speech research.[34] The mean F_0 and SDF_0 of males (mean F_0: 113.50 Hz; SDF_0: 32.83 Hz) and females (mean F_0: 194.67 Hz; SDF_0: 48.03 Hz) were comparable with our previous study.[30]

Apparatus and procedures. All speech samples were recorded by a handheld microphone (Shure Beta 58A; Shure Ltd, Niles, IL) positioned at a distance of 6 cm from the participant's mouth corner, through an external soundcard (M-audio Mobile PreUSB; M-Audio, Cumberland, RI) in two sound-attenuated rooms (room A and room B). The background noise in both rooms was less than 40.21 dBA. According to the classification scheme of the National Centre for Voice and Speech,[35] all speech samples were considered type I with periodic voice patterns displayed in the *Praat* (University of Amsterdam, The Netherlands) picture. The study comprised two speech conditions: baseline and modified TSST. Before beginning the study, all participants were fitted with a Polar heart rate monitor (model Polar E600; Polar Electro Oy, Kempele, Finland) and the nature of the study was explained to them, although no mention was made of the modified TSST.

Responsivity to stress was assessed by determining mean heart rate before and after introduction of the stress manipulation. A subjective measure of anxiety was also obtained by asking participants to complete an "anxiety thermometer"[36] by marking the appropriate point on a 10-cm visual analog scale ranging from 0 (meaning "not anxious at all") to 10 (meaning "extremely anxious") immediately before and after the stress manipulation.

Participants were first engaged in a 5-minute baseline condition in room A during which they were required to talk with the experimenter using their habitual speaking voice. Following this task, participants were given a rest interval of 3 minutes, during which self-rated anxiety score and heart rate were recorded. After the rest period, participants were told that they were required to deliver a formal 5 minutes presentation on a specified topic: "What can the government do to curb poverty in our society?" Participants were informed that the presentation would be recorded by video cameras and broadcasted on the official university Web site so that their performance could be assessed by senior academic staff.

Similar to the original TSST task, a duration of 10 minutes was provided to prepare for the formal presentation, following which participants were then guided to room B where the modified TSST took place. Room B contained two video cameras and two tape recorders, which were installed at the left- and right-hand sides of an interview table. A second experimenter (the interviewer) was seated at the table wearing a white laboratory coat. The participant was required to stand at a distance of 2 m from the interviewer to deliver the speech. The experimenter recorded heart rate and self-rated anxiety score and left the room, after which the participant delivered the speech to the interviewer. Consistent with the TSST protocol,[33] if participants finished their speech in less than 5 minutes, the interviewer responded in a standardized way by first staring at the participant silently for 20 seconds and then stating that the participant had time remaining and must continue to be financially rewarded on completion of the experiment. Finally, the participant was debriefed about the purpose of the study and informed that the speech would not be broadcasted online. A payment of USD 15 was made to each participant.

Data analysis. *Praat* software version 5.3.05[37] was used to measure the mean F_0 and SDF_0 of the speech samples. As females tend to have a higher pitch than males,[38] the absolute values of F_0 and SDF_0 (in the unit of Hertz) were converted to logarithmic scale (in the unit of semitone relative to an arbitrary musical note A1 or 55 Hz) to permit gender comparisons. All spectrograms were examined for mistracking errors of pitch by visually inspecting the narrowband spectrogram using the pitch analysis function provided by the *Praat* software. Any mistrackings were smoothed out using the *Praat* software. As manual segment extraction of each speech sample involved