Acoustic Measures of Voice and Physiologic Measures of Autonomic Arousal during Speech as a Function of Cognitive Load

***Megan K. MacPherson, †Defne Abur, and †‡§Cara E. Stepp,** **Tallahassee, FL, and †‡§Boston, MA*

Summary: Objectives/Hypothesis. This study aimed to determine the relationship among cognitive load condition and measures of autonomic arousal and voice production in healthy adults.

Study Design. A prospective study design was conducted.

Methods. Sixteen healthy young adults (eight men, eight women) produced a sentence containing an embedded Stroop task in each of two cognitive load conditions: congruent and incongruent. In both conditions, participants said the font color of the color words instead of the word text. In the incongruent condition, font color differed from the word text, creating an increase in cognitive load relative to the congruent condition in which font color and word text matched. Three physiologic measures of autonomic arousal (pulse volume amplitude, pulse period, and skin conductance response amplitude) and four acoustic measures of voice (sound pressure level, fundamental frequency, cepstral peak prominence, and low-to-high spectral energy ratio) were analyzed for eight sentence productions in each cognitive load condition per participant.

Results. A logistic regression model was constructed to predict the cognitive load condition (congruent or incongruent) using subject as a categorical predictor and the three autonomic measures and four acoustic measures as continuous predictors. It revealed that skin conductance response amplitude, cepstral peak prominence, and low-to-high spectral energy ratio were significantly associated with cognitive load condition.

Conclusions. During speech produced under increased cognitive load, healthy young adults show changes in physiologic markers of heightened autonomic arousal and acoustic measures of voice quality. Future work is necessary to examine these measures in older adults and individuals with voice disorders.

Key Words: Cepstral–Spectral–Autonomic arousal–Autonomic nervous system.

INTRODUCTION

Of the estimated 7% of the U.S. population impacted by voice disorders,[1,2] the most frequent diagnoses are classified as hyperfunctional voice disorders.[2] Although hyperfunctional voice disorders account for up to 40% of referrals to multidisciplinary voice clinics,[3] there is currently no agreement about their etiology. In addition to poor vocal hygiene and other voice-use factors,[4,5] psychological factors,[5–10] personality traits,[11–17] and autonomic nervous system dysfunction[16,18–20] have been implicated. With respect to the latter, investigations have primarily been based on signs and symptoms of autonomic dysfunction based on responses to questionnaires. However, there is growing experimental evidence that cognitive load and autonomic nervous system arousal affect speech motor control processes and detailed aspects of voice and speech motor performance in typical speakers.[21–24]

The autonomic nervous system is associated with the control of unconscious or involuntary physiologic functions. Traditionally, the sympathetic division has been associated with the preparation for and response to stressors (including higher cognitive demands), whereas the parasympathetic division has been associated with rest and repair functions.[25,26] Although they exert opposite influences, the sympathetic and parasympathetic divisions work together to facilitate bodily responses and the maintenance of homeostasis.[25,27] Thus, here, autonomic arousal is operationally defined as a physiologic state in which autonomic balance is shifted toward the sympathetic division relative to the parasympathetic division.

During speech production, autonomic arousal is increased relative to a resting state.[23,28–31] This arousal is thought to be related to similar changes in speech production that occur when cognitive demands are increased. With respect to speech articulation, higher cognitive demands have been shown to result in increased speech rate, increased spatiotemporal variability in labial kinematic patterns, and vowel centralization.[32–35] Overall, increased autonomic arousal has been found to detrimentally affect the speech motor system, leading to disfluency and decreased speech motor stability.[23,31,36] However, the effects of increased cognitive load and the related autonomic arousal on features of *vocal* motor control are less clear.

Whereas some studies have found that speakers show increased sound pressure level during speech produced with a concurrent cognitive load relative to typical speech,[22,34,37] other studies have shown decreases[38]; further, other work has found disparate sound pressure level changes in response to increased cognitive load within the same sample.[33] Likewise, increased,[39,40] decreased,[38] and subject-specific responses[37] in voice fundamental frequency have also been shown as a result of increased cognitive demands. The few studies that have examined acoustic measures of voice quality have suggested that increased

Accepted for publication October 27, 2016.

Early versions of this work were presented at the 44th and 45th Annual Symposiums of the Voice Foundation, Philadelphia, PA (2015 and 2016, respectively).

From the *Department of Communication Science and Disorders, Florida State University, Tallahassee, FL; †Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA; ‡Department of Biomedical Engineering, Boston University, Boston, MA; and the §Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, Boston, MA.

Address correspondence and reprint requests to Megan K. MacPherson, School of Communication Science and Disorders, Florida State University, 201 W. Bloxham Street, Tallahassee, FL 32306-1200. E-mail: mmacpherson@fsu.edu

Journal of Voice, Vol. 31, No. 4, pp. 504.e1–504.e9

0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.10.021

stress and/or cognitive demands are associated with a variety of vocal outcomes, some of which are conflicting: decreased time-domain perturbation measures (jitter and shimmer),[40] decreased nonharmonic noise,[40] increased energy in higher frequency harmonics[40] and thus decreased spectral tilt,[33] and shorter maximum phonation time (interpreted as increased breathiness and decreased glottal closure).[38] This wide range of results is likely a result of the variability in the number of subjects, the degree of control over the cognitive demands, and differences in the tasks and outcome measures.

Thus, the purpose of this study was to determine the relationship among the cognitive load condition in which a sentence was produced, physiologic measures of autonomic arousal, and acoustic measures of voice production in healthy young adults. We used an experimental paradigm that consisted of a sentence-level modification of Stroop's naming of color words task.[41] The Stroop task is a well-established mental stressor that has been used to study cognitive demands and autonomic arousal.[23,42–45] In the current study, a Stroop task was embedded into a sentence production task to experimentally manipulate cognitive load. Two cognitive load conditions were used: congruent, in which the color words were written in font colors that were the same as the semantic meaning of the text (eg, "red" written in red font), and incongruent, in which the color words were written in font colors that differed from the semantic meaning of the text (eg, "blue" written in red font). In both cognitive load conditions, participants were instructed to say aloud the name of the font color in which the color word was written, rather than the word itself. The incongruent condition represented an increase in cognitive load relative to the congruent condition. This increase in cognitive load has been primarily attributed to inhibitory and attentional executive processes that are taxed in the incongruent, or interference, condition.[44,46,47] Here, physiologic measures of autonomic arousal and acoustic measures of voice were examined as predictors of the congruent and incongruent cognitive load conditions.

METHOD

Participants

Participants were 16 individuals with typical voices who reported no history of voice, speech, language, or hearing disorders (8 men, 8 women). They were aged 22-32 years (M = 25.8 years; SD = 3.5 years). Thorough inclusionary screening was conducted to control for factors known to affect autonomic, cognitive, linguistic, and speech functions. Likewise, thorough exclusionary criteria were used to ensure that participants were free from factors that could affect these functions.

To be included in the study, all participants were required to pass a pure tone hearing screening at 25 dB HL at 500, 1000, 2000, 4000, and 6000 Hz, demonstrate normal oral motor function as assessed with the Oral Speech Mechanism Screening Examination-3rd Edition,[48] and demonstrate normal color vision as assessed with the Ishihara Color Blindness Test.[49] All participants were native speakers of North American English, a requirement incorporated to add further control for differences in reading ability and cognitive load associated with speech production. Additionally, all participants demonstrated age-appropriate cognitive and language skills, as assessed with the Cognitive Linguistic Quick Test,[50] and demonstrated reading abilities at an eighth-grade level or higher, as assessed with the Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests-Revised.[51] All participants successfully completed the Sentence Reading subtest of the Psycholinguistic Assessment of Language Processing in Aphasia,[52] which required them to read aloud sentences with varying syntactic and semantic properties with at least 80% accuracy.

Individuals were excluded if they were taking medication known to have an appreciable effect on motor or cognitive function (eg, medications for attention deficit disorder, anticonvulsants, and muscle relaxants). Because of potential effects on autonomic function, no individual reported a history of any of the following: autonomic failure, multiple system atrophy, diabetes, chronic obstructive pulmonary disease or emphysema, drug or alcohol abuse, or schizophrenia.[53–67] None of the participants had smoked within the past 5 years. None of the participants reported any of the following within the 6 months before the study: pregnancy or nursing; active depression, anxiety, or other psychiatric or psychological disorders; high or low blood pressure that was not under control with medication and/or lifestyle modifications; prediabetes or metabolic syndrome; sleep apnea or other diagnosed sleep disorders; dermatologic conditions (eg, eczema and psoriasis) affecting the hands; or loss of sensory or motor function in the upper extremities.[54,56,68–73] Finally, participants reported having abstained from the consumption of alcohol, caffeine, and large meals and having not experienced any heavy physical activity or stressful events, such as a vigorous workout or a class exam, for at least 3 hours before the experiment.[23]

Procedure

Participants were seated in front of a computer monitor that was used for stimulus presentation. The examiner attached the electrodes and transducer for physiologic data collection and positioned the microphone for collection of the acoustic signal. The examiner then explained the experimental task, model and practice sentences were completed, and the experimental task commenced. The experimental task consisted of a sentence-level modified Stroop paradigm in which cognitive load was manipulated through the use of congruent and incongruent Stroop conditions.[41] Autonomic and acoustic signals were collected over repeated sentence productions. Breaks and an informal picture description task were interspersed at regular intervals to decrease the potential for monotony and autonomic habituation.

The stimulus sentence used in this experiment was *Pammy and Bobby picked blue, pink, red, and brown poppies with their mommy*. The stimulus sentence had a Flesch-Kincaid grade level of 4.9, indicating that it should be understandable to individuals with a fourth-grade education level,[74,75] and it had a Flesch Reading Ease score of 83, indicating that it should be "easy" to read.

As previously mentioned, the stimulus sentence contained an embedded Stroop task. This task occurred on the four color words in the middle of the sentence ("and" was always presented in black font). The sentence occurred an equal number of times in each of the two cognitive load conditions: congruent and