



Published in final edited form as:

J Voice. 2017 July ; 31(4): 504.e1–504.e9. doi:10.1016/j.jvoice.2016.10.021.

Acoustic measures of voice and physiologic measures of autonomic arousal during speech as a function of cognitive load

Megan K. MacPherson¹, Defne Abur², and Cara E. Stepp^{2,3,4}

¹Department of Communication Science and Disorders, Florida State University

²Department of Speech, Language, and Hearing Sciences, Boston University

³Department of Biomedical Engineering, Boston University

⁴Department of Otolaryngology – Head and Neck Surgery, Boston University School of Medicine

Abstract

Objectives/Hypothesis—The purpose of this study was to determine the relationship among cognitive load condition and measures of autonomic arousal and voice production in healthy adults.

Study Design—Prospective.

Methods—Sixteen healthy young adults (eight females) 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 three variables were significantly associated with cognitive load condition: skin conductance response amplitude, cepstral peak prominence, and the low-to-high spectral energy ratio.

Conclusions—During speech produced under increased cognitive load, healthy young adults show changes in physiologic markers of heightened autonomic arousal and acoustic measures of

§Corresponding Author: Megan K. MacPherson – mmacpherson@fsu.edu, Address: Florida State University, School of Communication Science and Disorders, 201 W. Bloxham Street, Tallahassee, FL 32306-1200, Phone: (850) 644-8465 Fax: (850) 644-6639.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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

voice quality. Future work is necessary to examine these measures in older adults and individuals with voice disorders.

Keywords

cepstral; spectral; pulse amplitude; pulse period; skin conductance response; 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 disorders². Although hyperfunctional voice disorders account for up to 40% of referrals to multidisciplinary voice clinics³, there is not currently agreement about their etiology. In addition to poor vocal hygiene and other voice-use factors^{4, 5}, psychological factors⁵⁻¹⁰, personality traits¹¹⁻¹⁷, 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.

While 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}, others have shown decreases³⁸; 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}, decreased³⁸, and subject-specific responses³⁷ 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 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 non-harmonic noise⁴⁰, increased energy in higher frequency harmonics⁴⁰ and thus decreased spectral tilt³³, 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 employed 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 in order to experimentally manipulate cognitive load. Two cognitive load conditions were utilized: congruent, in which the color words were written in font colors that were the same as the semantic meaning of the text (e.g., "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 (e.g., "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 males, 8 females). They were aged 22 to 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 utilized to ensure that participants were free from factors that could affect these functions.

In order 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 (OSMSE-3)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 (CLQT)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 (WRMT-R).⁵¹ All participants successfully completed the Sentence Reading subtest of the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)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 (e.g., medications for attention deficit disorder, anticonvulsants, and muscle relaxants). Due to 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 five years. None of the participants reported any of the following within the six months prior to 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; pre-diabetes or metabolic syndrome; sleep apnea or other diagnosed sleep disorders; dermatological conditions (e.g., 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 three hours prior to 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 utilized 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 two cognitive load conditions: congruent and incongruent. In the congruent condition, color words were written in font colors that matched the semantic meaning of the text (e.g., “blue”

written in blue font). In the incongruent condition, the font colors in which the color words were written did not match the semantic meaning of the text (e.g., “pink” written in blue font). In both conditions, participants were instructed to say the font colors in which the color words were written, rather than reading the words. The font colors and thus the target production were the same in both conditions. Thus, in the above example, participants should have said “blue” in both cases. For the incongruent condition, the color words used in each sentence were pseudorandomly varied across repetitions, with the exception that “pink” was never presented in red font and “red” was never presented in pink font. Additionally, a color word was never written in the matching font color in the incongruent condition.

In each cognitive load condition, the stimulus sentence was presented as both target and foil sentences, with additional sentences with the same overall structure. The target sentences were those utilized for analysis. The foil sentences were the same as the target sentences, except that the orders of the embedded font colors and color words were changed. Foil sentences were not included in the analysis. The foil sentences and additional sentences were included in the experimental protocol to reduce participants’ ability to predict or remember which colors occurred with the target sentences and to reduce the potential for autonomic habituation. As with the target sentences, foil and additional sentences occurred in both the congruent and incongruent conditions.

Prior to beginning the experimental task, the experimenter verified that participants were able to clearly identify the colors used in the sentence stimuli, as presented on the computer monitor, as well as clearly read text from the computer monitor. The examiner explained the task and modeled the production of the stimulus sentence (shown on the monitor in all black font) once in order to familiarize participants with the sentences. The examiner then showed participants practice sentences and modeled what should be said for one practice sentence in each cognitive load condition. After the examiner's models, participants completed two practice sentences in each condition. Reinstruction and additional modeling were provided, if needed. The practice sentences were different from those used in the experimental task. During the experimental task, participants produced multiple spoken repetitions of the sentence stimuli at their habitual rate and loudness. Stimuli were visually presented one-at-a-time in Arial 48-point bold font on the 20-inch computer monitor located approximately 5 feet in front of participants.

Sentences (target, foil, and additional) were presented in sets of four blocks each (four sentences per block) with each set containing two blocks of congruent sentences and two blocks of incongruent sentences. Target sentences occurred once in each condition per set. Additionally, the target sentences always occurred in the first or second position in a given block, in order to best capture autonomic responsiveness. The order in which the congruent and incongruent blocks appeared was pseudorandomized across sets. Each sentence block was preceded by a 30-s rest period during which baseline autonomic data were collected. Within each block, each sentence production was followed by an 8-s rest period to allow for autonomic recovery. During rest periods, participants sat quietly and viewed fixation crosses on the computer monitor. A short break was taken after each set of sentences, during which time participants were able to move and ask questions. In order to decrease monotony, at every other break participants completed a brief (< 30 s) picture description task (not used in

analyses). At a minimum, target sentences were presented eight times in each cognitive load condition per participant, resulting in at least eight productions in each condition that were appropriate for both autonomic and acoustic analysis, after excluding from data analysis inaccurate or atypical sentence productions. The entire experimental protocol took approximately 1 hour for each participant to complete.

Instrumentation and Signal Processing

Autonomic and acoustic signals were collected simultaneously with a Biopac MP150 Data Acquisition System, amplifiers, and the AcqKnowledge program (Biopac Systems, Inc.). Two peripheral autonomic signals, pulse and skin conductance, were collected with the Biopac system via a transducer and electrodes attached to the right hand and the associated amplifiers. These analog autonomic signals were passed through the Biopac system to the ODAU II of a 3D Investigator Motion Capture System (Northern Digital) for recording. The acoustic signal was also directly recorded using parameters suitable for acoustic analyses, as described later.

The pulse signal was collected with the Biopac Pulse Plethysmogram Amplifier (PPG100C) and Photo Plethysmogram Transducer (TSD200). The transducer was attached with a Velcro strap to the palmar surface of the distal phalange of the third/ring finger on the right hand. Neither the position of the transducer nor the tension of the Velcro strap was altered once the experimental task had begun so that the deformation of the underlying capillary bed remained constant throughout the experiment. The transducer emits an infrared signal and measures the amount of the signal reflected back, which varies with the amount of blood flow in the underlying capillaries.⁷⁶ Reflectance increases with increased capillary blood volume. The pulse signal was recorded with a gain of 100 V and a sampling rate of 14925 samples/s. It was bandpass filtered from 0.5 – 3 Hz with a 100-order FIR filter and downsampled to 250 samples/s.

The skin conductance signal was collected with the constant voltage (0.5 V) Biopac GSR EDA Galvanic Skin Response Amplifier (GSR100C) and two disposable EDA electrodes (EL507). The electrodes were attached to the palmar surface of the medial phalanges of the first/index and second/middle fingers⁵⁶ of the right hand. The electrodes were in place for a minimum of five minutes before experimental data collection began. The skin conductance signal reflects the activity of the eccrine sweat glands (and thus, sweat secretion) underlying the electrodes. Electrical conductance between these electrodes increases with increased sweat gland activity. The tonic skin conductance level signal was recorded with a gain of 10mS/V and a sampling rate of 14925 samples/s. It was low-pass filtered at 1 Hz and downsampled to 250 samples/s. The phasic skin conductance response signal was then derived from the level signal through the application of a direct-form II, second-order section Chebyshev high pass filter with a cut-off frequency of 0.07 Hz.

The speech acoustic signal was transduced with a Countryman E6i omnidirectional earset microphone at a 6 cm mouth-to-microphone distance and 90 degree mouth-to-microphone angle. The acoustic signal was recorded with a Marantz PMD670 solid state recorder at a sampling rate of 44.1 kHz.

Autonomic Measures

All autonomic data were analyzed with a custom MATLAB (MathWorks, Natick, MA) program. All data were also visually inspected and any necessary changes were made to insure measurement accuracy. Autonomic signals were examined across the entire sentence, from a trigger signal marking the appearance of the stimulus sentence on the computer monitor to three seconds after the end of the acoustic signal. This allowed the autonomic signals to reflect the arousal associated with sentence planning and preparation and accommodated the relatively slow-moving time course of the autonomic signals (compared to the time course of speech movements).^{54, 56} Autonomic data from the 30-s rest periods preceding each sentence block were analyzed to provide baseline autonomic values prior to each of the target sentences. Within each 30-s rest period, the two 5-s windows in which the autonomic signals reflected that the system was most relaxed were selected. The average values of each of the applicable autonomic measures across the two 5-s windows were used to create baseline values for each sentence block per participant.

The autonomic measures used in this study were selected for two primary reasons. First, they provide information on both sympathetic and parasympathetic function as well as cardiovascular (pulse measures) and electrodermal (skin conductance measure) systems. Second, the selected measures are established within the speech production literature and have been shown to be sensitive to speech production demands.^{23, 31, 77}

Two autonomic dependent measures, pulse volume amplitude and pulse period, were derived from the pulse signal. *Pulse volume amplitude* was measured by identifying amplitude minima and maxima in the pulse signal with a peak-finding algorithm and calculating the difference between adjacent maximum and minimum amplitudes. The mean pulse volume amplitude for each sentence production was referenced to the baseline mean pulse volume amplitude for that block, and it was then expressed as a percent of the baseline value. Thus, pulse volume amplitude as a percent of baseline amplitude was utilized in the statistical analysis. Pulse volume amplitude is expected to decrease with increased sympathetic arousal, as increased sympathetic arousal is associated with peripheral vasoconstriction. *Pulse period* was calculated as the time between adjacent pulse peaks in the pulse signal. It was measured using the same pulse peaks identified to calculate pulse volume amplitude. The mean pulse period was calculated for each sentence production. It was referenced to the baseline mean pulse period for that block and expressed as a percent of the baseline mean. Thus, pulse period during sentence production as a percent of baseline pulse period was utilized in the statistical analysis. Pulse period is expected to decrease with increased sympathetic arousal and/or decreased parasympathetic activation, as the heart is innervated by both of these divisions of the autonomic nervous system.

The amplitude of the phasic *skin conductance response* (in μS) occurring during each sentence production was measured as the difference between the maximum and onset amplitudes of the skin conductance response.⁵⁶ An algorithm automatically identified the points of onset and maximum of the skin conductance response within each sentence production window. The skin conductance response is a phasic (short-term) measure of electrodermal activity associated with the eccrine sweat gland activity. As sympathetic arousal increases, the activity of the eccrine sweat glands in certain body locations (such as

the palmar surfaces of the hands) increases. This increased sweat gland activity leads to greater electrical conductance between the electrodes on the fingers and increases in the amplitude of the skin conductance response.^{54, 56, 78} Due to the nature of this measure (i.e., skin conductance responses should not occur during the measurement windows used for calculation of baseline values), it is not referenced to baseline values.

Acoustic Measures

The average *sound pressure level* of each sentence was calculated using a custom MATLAB program. Prior to calculations, pauses were removed from all participant data to prevent bias in the measure. To remove pauses, data were initially filtered using a Butterworth filter designed in MATLAB to remove high frequency content. An envelope was then taken of the resulting signal to determine an appropriate cut-off for instances of speech. After the cut-off was applied, data were visually checked to ensure that all instances of speech were maintained. The remaining sections of data were concatenated and were used to calculate the average root-mean-square (RMS) in Volts for each sentence. The RMS was then converted to decibels relative to each participant's average RMS during the congruent condition. This measure was included in order to act as a direct comparison to the previous studies that found variable results across differing levels of experimental control.
22, 33, 34, 37, 38

The *Computerized Speech Lab 4500* and the add-in program *Analysis of Dysphonia in Speech and Voice (ADSV)* (KayPENTAX) were used to perform subsequent acoustic analyses using a cepstral peak prominence (CPP) threshold value of 6 dB. The measures were calculated as described by Awan⁷⁹. The software performed a discrete Fourier transform on each production, which was divided into regions of low- and high-frequency energy using a 4000 Hz cutoff. The low-to-high spectral energy ratio in decibels is termed the *L/H ratio*. This spectral measure tends to be lower in dysphonic voices, which can be due to relatively higher energy in upper harmonics and/or increased spectral noise associated with perceptually breathy voices. ADSV computes cepstral-based measures by performing a second discrete Fourier transform on the power spectrum, followed by smoothing. The ratio in decibels of the maximum amplitude of the smoothed cepstrum and the expected value via a regression analysis is computed, forming the *cepstral peak prominence* (CPP). CPP values tend to be lower in dysphonic voices due to disturbed periodicity. The L/H ratio and CPP provide complementary acoustic features that are related to voice quality, which was of primary interest in this study. The frequency associated with the CPP is an estimate of the *fundamental frequency* of the voice. Values for each sentence in Hz were converted to semitones relative to each participant's average values during the congruent condition, allowing for comparisons across male and female participants. As in the case of sound pressure level, the fundamental frequency was included in order to compare results to previous studies that found variable results across differing levels of experimental control.
37-40

Statistical Analysis

With Minitab statistical software (Version 17; Minitab Inc., State College, PA), a binary 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. Significance for the predictor variables was set *a priori* to $p < .05$ and was tested using likelihood ratio tests. Because the skin conductance response amplitude data were left-skewed, a cube root transform was applied to the data before statistical analysis, resulting in a distribution that was not significantly different from a normal distribution ($p > 0.05$) via a Kolmogorov–Smirnov test.

Results

Average autonomic measures are shown as a function of cognitive load condition (congruent or incongruent) in Figure 1. As expected, pulse volume amplitude during sentence production (expressed as a percent of baseline amplitude) was on average lower during the incongruent condition ($M = 60.8\%$, $SD = 26.6\%$) relative to the congruent condition ($M = 64.4\%$, $SD = 28.0\%$). Additionally, pulse period during sentence production (expressed as a percent of baseline pulse period) was shorter on average during the incongruent condition ($M = 93.8\%$, $SD = 9.3\%$) relative to the congruent condition ($M = 95.6\%$, $SD = 11.5\%$). Also, as expected, skin conductance response amplitude was on average higher during the incongruent condition ($M = 0.886 \mu\text{S}$, $SD = 1.091 \mu\text{S}$) relative to the congruent condition ($M = 0.667 \mu\text{S}$, $SD = 0.794 \mu\text{S}$).

Average acoustic measures are shown as a function of cognitive load condition in Figure 2. No notable difference was seen in f_0 between the congruent ($M = 0.025 \text{ ST}$, $SD = 0.912 \text{ ST}$) and the incongruent ($M = 0.005 \text{ ST}$, $SD = 1.06 \text{ ST}$) conditions or in the sound pressure level between the congruent ($M = -0.623 \text{ dB}$, $SD = 2.98 \text{ dB}$) and the incongruent ($M = -.267 \text{ dB}$, $SD = 3.10 \text{ dB}$) conditions. On average, CPP was higher during the incongruent condition ($M = 7.11 \text{ dB}$, $SD = 1.11 \text{ dB}$) relative to the congruent condition ($M = 7.00 \text{ dB}$, $SD = 1.12 \text{ dB}$), whereas the L/H ratio was on average lower during the incongruent condition ($M = 37.8 \text{ dB}$, $SD = 2.78 \text{ dB}$) relative to the congruent condition ($M = 38.1 \text{ dB}$, $SD = 2.80 \text{ dB}$).

Results of the logistic regression (Table 1) revealed that three predictor variables were significantly associated with the cognitive load condition: skin conductance response amplitude ($P = 0.001$), CPP ($P = 0.050$), and L/H ratio ($P = 0.004$). Despite these significant effects, overall the model provided a relatively poor fit of the data ($R^2 = 0.08$). However, this was likely due to the stochastic nature of physiologic and acoustic data, rather than an inappropriate modeling choice, because a Pearson chi-squared goodness-of-fit test indicated a nonsignificant result ($\chi^2 = 265.3$, $P = 0.06$).

Discussion

This study sought to determine the relationship among the cognitive load condition (congruent or incongruent) in which a sentence was produced, physiologic measures of autonomic arousal, and acoustic measures of voice production. Overall, the findings indicate that measures of autonomic arousal and voice production were predictive of cognitive load condition, suggesting that these measures are physiologic and acoustic markers of increased cognitive load.

Autonomic Findings

Regarding the autonomic nervous system findings, the results of the logistic regression (Table 1) indicated that skin conductance response amplitude was a significant predictor of the cognitive load condition in which the target sentences were produced ($p = .001$), with significantly greater amplitudes associated with the incongruent (increased cognitive load) condition. Skin conductance response amplitude reflects activation of the sympathetic (fight or flight) branch of the autonomic nervous system.²⁵ While not significant, the two other autonomic measures, pulse volume amplitude and pulse period, also suggested increased sympathetic activation (and potentially reduced parasympathetic activation in the case of pulse period) in the incongruent condition compared to the congruent condition. Thus, the branch of the autonomic nervous system associated with response to stressors showed increased activation in the heightened cognitive load condition, when measures of vocal quality were also affected.

This study adds to the growing literature aimed at understanding the relationship between physiologic autonomic measures and vocal mechanism function. Helou and colleagues found increased intrinsic laryngeal muscle activation paired with increased autonomic arousal (as reflected by heart rate and blood pressure measures) during stress (a cold pressor task).²¹ Relatedly, Dietrich and Verdolini Abbott documented changes in systolic blood pressure between baseline speech and stress (public speaking) conditions.⁸⁰ Finally, a clinical case study of patient-physician interactions showed that changes in acoustic measures of voice were associated with independently detected intervals of increased skin conductance response amplitudes.⁸¹ Although laryngeal electromyography was not completed in the present study and our data did not yield significant cardiovascular results, our finding of significantly increased electrodermal autonomic arousal is in general agreement with all three studies, demonstrating increased autonomic arousal under more demanding or stressful conditions and concomitant changes in vocal parameters.

The current work extends prior findings related to autonomic influences on the quality of speech motor control to autonomic influences on vocal quality. More specifically, the finding that vocal quality is affected by heightened autonomic arousal is consistent with work from the speech motor control domain, which has shown that speech motor performance (primarily articulatory) is negatively impacted by heightened autonomic arousal.^{23, 31} Collectively, the current findings together with those from prior work suggest that the larger anatomic and physiologic mechanism for speech and voice production is impacted by heightened autonomic arousal associated with increased task demands.

Prior work on the relationship between autonomic function and voice function has largely utilized self-report measures of autonomic symptoms¹⁸⁻²⁰, finding that individuals with dysphonia or voice complaints report more symptoms of autonomic dysfunction than those in comparison groups. The present findings that healthy young adults without voice disorders or complaints demonstrate increased autonomic arousal together with changes in vocal quality suggest that examining physiologic measures of autonomic arousal in individuals with voice disorders may be a fruitful avenue for better understanding the mechanisms underlying these disorders.

Voice Findings

Results of the logistic regression (Table 1) revealed that two of the acoustic measures of voice production were significantly predictive of the cognitive load condition: ($p = 0.050$) and the L/H ratio ($p = 0.004$). Increased CPP values suggest that the voicing signal was more periodic under cognitive load. This is consistent with the decreases in time-domain perturbation measures (jitter and shimmer) seen in previous work.⁴⁰ Lower L/H ratio values during the increased cognitive demand could be due to an increase in the energy of upper harmonics or increased spectral noise associated with perceptually breathy voices. Thus, interpretation of changes in L/H ratio values is less obvious. Previous studies have shown increased energy in higher frequency harmonics⁴⁰ and decreased spectral tilt³³ with increased cognitive demands, which would be consistent with decreased L/H ratio values here that are a result of increased energy of upper harmonics. However, previous studies have also shown both *decreases* in spectral noise⁴⁰ as well as shorter maximum phonation time, interpreted as increased breathiness, which should accompany increases in spectral noise.³⁸ However, given the finding of increased CPP values, the lower L/H ratio values under cognitive load are most likely related to the creation of a more pressed voice, with increased (periodic) energy in the higher vocal harmonics. That said, physiologic interpretation of these changes in L/H ratio values is not truly possible without physiologic data. Future work in this area should incorporate methods such as electroglottography and high speed videoendoscopy in order to better understand the bases of changes in voice quality under increased cognitive load.

Neither sound pressure level nor fundamental frequency were significant predictors of the presence of a cognitively demanding task. This lack of a strong association between cognitive load condition and the core voice attributes of sound pressure level and fundamental frequency is not terribly surprising. Although a variety of studies have examined these two acoustic measures as a function of cognitive load, the results have been disparate.^{22, 33, 34, 37-40} Based on these previous studies and the present results, voice *quality* may be more strongly impacted by increased cognitive load than sound pressure level or fundamental frequency.

Limitations and Future Work

This study is an initial step to understanding the intersection between voice production, autonomic arousal, and cognitive load. It offers both strengths and limitations. A strength of the current study is the use of direct, physiologic measures of autonomic arousal derived from multiple autonomic signals recorded during voice production, rather than secondary measures derived from responses on questionnaires. However, the sample size utilized does not allow for generalization of the results to the greater population. Also, voice production measures were limited to (non-invasive) acoustic measures. Future studies should additionally incorporate physiologic measures of voice production to aid in interpretation. Although acoustic measures were found to be significant predictors of increased cognitive load, the small magnitudes of these changes are of unknown clinical significance. Finally, a primary strength of the study is its well-controlled nature, which is also a limitation. The Stroop task allowed manipulation of cognitive load in the experimental setting in a manner that was consistent across participants. However, it is unclear how the task might transfer to

daily life. Well-controlled laboratory studies are ultimately limited in their ability to characterize vocal behaviors, particularly in light of stressors. New techniques for ambulatory monitoring of voice use via neck-surface acceleration⁸² when paired with ambulatory sensing of autonomic arousal may eventually allow for translation of this work to the everyday lives of speakers, including older individuals (who may respond differently to task demands) and individuals with voice disorders.

Conclusion

A physiologic measure of autonomic arousal (skin conductance response amplitude) and two measures of voice quality (CPP and the L/H ratio) were significantly associated with whether the cognitive load condition was congruent or incongruent. These results suggest that voice quality is impacted by increases in cognitive load, and that the associated changes in voice quality may be driven by autonomic nervous system arousal. More research is necessary to determine the extent and time course of these interactions, and to extend these methods to a broader population including older individuals and individuals with voice disorders.

Acknowledgments

This work was supported by NIH grant DC015570 from the National Institute on Deafness and Other Communication Disorders (NIDCD). We thank Carolyn Michener for preliminary data analysis assistance, Elizabeth Heller Murray for advice about acoustical analysis, and Victoria McKenna for statistical consulting.

References

1. Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *Laryngoscope*. 2005; 115:1988–1995. [PubMed: 16319611]
2. Bhattacharyya N. The prevalence of voice problems among adults in the United States. *Laryngoscope*. 2014; 124:2359–2362. [PubMed: 24782443]
3. Roy N. Functional dysphonia. *Current Opinion in Otolaryngology & Head and Neck Surgery*. 2003; 11:144–148. [PubMed: 12923352]
4. Altman KW, Atkinson C, Lazarus C. Current and emerging concepts in muscle tension dysphonia: a 30-month review. *J. Voice*. 2005; 19:261–267. [PubMed: 15907440]
5. Van Houtte E, Van Lierde K, Claeys S. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *J. Voice*. 2011; 25:202–207. [PubMed: 20400263]
6. Morrison MD, Nichol H, Rammage LA. Diagnostic criteria in functional dysphonia. *Laryngoscope*. 1986; 96:1–8.
7. Roy N, Bless DM, Heisey D. Personality and voice disorders: a multitrait-multidisorder analysis. *J. Voice*. 2000; 14:521–548. [PubMed: 11130110]
8. Baker J, Ben-Tovim D, Butcher A, Esterman A, McLaughlin K. Psychosocial risk factors which may differentiate between women with Functional Voice Disorder, Organic Voice Disorder and a Control group. *Int J Speech Lang Pathol*. 2013; 15:547–563. [PubMed: 23075157]
9. Kinzl J, Biebl W, Rauegger H. Functional aphonia: psychosomatic aspects of diagnosis and therapy. *Folia Phoniatr. (Basel)*. 1988; 40:131–137. [PubMed: 3220342]
10. Gunther V, Mayr-Graft A, Miller C, Kinzl H. A comparative study of psychological aspects of recurring and non-recurring functional aphonias. *Eur. Arch. Otorhinolaryngol*. 1996; 253:240–244. [PubMed: 8737777]
11. Ng JH, Lo S, Lim F, Goh S, Kanagalingam J. Association between anxiety, type A personality, and treatment outcome of dysphonia due to benign causes. *Otolaryngology--Head and Neck Surgery*. 2013; 148:96–102. [PubMed: 23112274]

12. Roy N, Bless DM. Personality Traits and Psychological Factors in Voice Pathology A Foundation for Future Research. *Journal of Speech, Language, and Hearing Research*. 2000; 43:737–748.
13. Roy N, Bless DM, Heisey D. Personality and Voice Disorders A Superfactor Trait Analysis. *Journal of Speech, Language, and Hearing Research*. 2000; 43:749–768.
14. O'Hara J, Miller T, Carding P, Wilson J, Deary V. Relationship between fatigue, perfectionism, and functional dysphonia. *Otolaryngology--Head and Neck Surgery*. 2011
15. McHugh-Munier C, Scherer KR, Lehmann W, Scherer U. Coping strategies, personality, and voice quality in patients with vocal fold nodules and polyps. *J. Voice*. 1997; 11:452–461. [PubMed: 9422280]
16. van Mersbergen M, Patrick C, Glaze L. Functional dysphonia during mental imagery: testing the trait theory of voice disorders. *J. Speech Lang. Hear. Res.* 2008; 51:1405–1423. [PubMed: 18664709]
17. Willinger U, Aschauer HN. Personality, anxiety and functional dysphonia. *Personality and individual differences*. 2005; 39:1441–1449.
18. Demmink-Geertman L, Dejonckere PH. Nonorganic habitual dysphonia and autonomic dysfunction. *J. Voice*. 2002; 16:549–559. [PubMed: 12512642]
19. Park K, Behlau M. Signs and symptoms of autonomic dysfunction in dysphonia individuals. *Jornal da Sociedade Brasileira de Fonoaudiologia*. 2011; 23:164–169. [PubMed: 21829933]
20. Paes CF, Zambon FC, Behlau M. Signs and symptoms of the autonomic dysfunction in teachers. *Revista CEFAC*. 2014; 16:957–966.
21. Helou LB, Wang W, Ashmore RC, Rosen CA, Abbott KV. Intrinsic laryngeal muscle activity in response to autonomic nervous system activation. *Laryngoscope*. 2013; 123:2756–2765. [PubMed: 23620064]
22. Dromey C, Bates E. Speech interactions with linguistic, cognitive, and visuomotor tasks. *J. Speech Lang. Hear. Res.* 2005; 48:295–305. [PubMed: 15989393]
23. Kleinow J, Smith A. Potential interactions among linguistic, autonomic, and motor factors in speech. *Dev. Psychobiol.* 2006; 48:275–287. [PubMed: 16617462]
24. Smith A, Goffman L. Interaction of motor and language factors in the development of speech production.. In: Maasen B, Kent RD, Peters HFM, Van Lieshout PHHM, Hulstijn W, editors *Speech Motor Control in Normal and Disordered Speech*. Oxford University Press; Oxford, UK: 2004. 227–252.
25. Andreassi JL. *Psychophysiology: Human Behavior and Physiological Response*. 5th ed.. Larence Erlbaum Associates; Mahwah, NJ: 2007.
26. Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 3rd ed.. Lippincott Williams and Wilkins; Baltimore, MD: 2007.
27. Hamill RW, Shapiro RE, Vizzard MA. Peripheral autonomic nervous system.. In: Robertson D, Biaggioni I, Burnstock G, Low PA, Paton JFR, editors *Primer on the autonomic nervous system*. 3rd ed.. Elsevier; Amsterdam, The Netherlands: 2012. 17–26.
28. Linden W. A microanalysis of autonomic activity during human speech. *Psychosom. Med.* 1987; 49:562–578. [PubMed: 3423164]
29. Lipman RD, Grossman P, Bridges SE, Hamner JW, Taylor JA. Mental stress response, arterial stiffness, and baroreflex sensitivity in healthy aging. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2002; 57:B279–284. [PubMed: 12084798]
30. Lynch JJ, Long JM, Thomas SA, Malinow KL, Katcher AH. The effects of talking on the blood pressure of hypertensive and normotensive individuals. *Psychosom. Med.* 1981; 43:25–33. [PubMed: 7208754]
31. Weber CM, Smith A. Autonomic correlates of stuttering and speech assessed in a range of experimental tasks. *J. Speech Hear. Res.* 1990; 33:690–706. [PubMed: 2273884]
32. Huttunen KH, Keranen HI, Paakkonen RJ, Paivikki Eskelinen-Ronka R, Leino TK. Effect of cognitive load on articulation rate and formant frequencies during simulator flights. *J. Acoust. Soc. Am.* 2011; 129:1580–1593. [PubMed: 21428521]
33. Lively SE, Pisoni DB, Van Summers W, Bernacki RH. Effects of cognitive workload on speech production: acoustic analyses and perceptual consequences. *J. Acoust. Soc. Am.* 1993; 93:2962–2973. [PubMed: 8315159]

34. Dromey C, Shim E. The effects of divided attention on speech motor, verbal fluency, and manual task performance. *J. Speech Lang. Hear. Res.* 2008; 51:1171–1182. [PubMed: 18664706]
35. Dromey C, Benson A. Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *J. Speech Lang. Hear. Res.* 2003; 46:1234–1246. [PubMed: 14575355]
36. McCarron LT. Paralanguage and autonomic response patterns in psychotherapy. *Psychotherapy: Theory, Research, and Practice.* 1973; 10:229–230.
37. Huttunen K, Keranen H, Vayrynen E, Paakkonen R, Leino T. Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. *Appl. Ergon.* 2011; 42:348–357. [PubMed: 20832770]
38. Van Lierde K, Van Heule S, De Ley S, Mertens E, Claeys S. Effect of psychological stress on female vocal quality: a multiparameter approach. *Folia Phoniatr. Logop.* 2008; 61:105–111.
39. Scherer KR, Grandjean D, Johnstone T, Klasmeyer G, Banzinger T. Acoustic correlates of task load and stress. *INTERSPEECH.* 2002
40. Mendoza E, Carballo G. Acoustic analysis of induced vocal stress by means of cognitive workload tasks. *J. Voice.* 1998; 12:263–273. [PubMed: 9763177]
41. Stroop JR. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 1935; 18:643–662.
42. Kobayashi N, Yoshino A, Takahashi Y, Nomura S. Autonomic arousal in cognitive conflict resolution. *Auton. Neurosci.* 2007; 132:70–75. [PubMed: 17067858]
43. MacLeod CM. Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 1991; 109:163–203. [PubMed: 2034749]
44. MacLeod CM, MacDonald PA. Interdimensional interference in the Stroop effect: unconverging the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences.* 2001; 4:383–391.
45. Tulen JH, Moleman P, van Steenis HG, Boomsma F. Characterization of stress reactions to the Stroop Color Word Test. *Pharmacol. Biochem. Behav.* 1989; 32:9–15. [PubMed: 2734355]
46. Andres P, Guerrini C, Phillips LH, Perfect TJ. Differential effects of aging on executive and automatic inhibition. *Developmental neuropsychology.* 2008; 33:101–123. [PubMed: 18443972]
47. Mathis A, Schunck T, Erb G, Namer IJ, Luthringer R. The effect of aging on the inhibitory function in middle-aged subjects: a functional MRI study coupled with a color-matched Stroop task. *Int. J. Geriatr. Psychiatry.* 2009; 24:1062–1071. [PubMed: 19418476]
48. St. Louis KO, Ruscello DM. Oral Speech Mechanism Screening Examination. Third ed.. Pro-Ed; Austin, TX: 2000.
49. Ishihara S. Ishihara's Tests for Colour Deficiency. Kanehara Trading; Tokyo, Japan: 2011.
50. Helm-Estabrooks N. Cognitive Linguistic Quick Test. The Psychological Corporation; San Antonio, TX: 2001.
51. Woodcock RW. Woodcock Reading Mastery Tests - Revised (NU). Pearson; Bloomington, MN: 1998.
52. Kay J, Lesser R, Coltheart M. Psycholinguistic Assessment of Language Processing in Aphasia. Erlbaum; Hove, UK: 1992.
53. Katsilambros NL, Boulton AJ, Tentolouris N, Kokkinos A, Liatis S. Autonomic neuropathy in diabetes mellitus and obesity: an update. *Exp. Diabetes Res.* 2011; 2011:607309. [PubMed: 22203830]
54. Boucsein W. Electrodermal Activity. 2nd ed.. Springer; New York, NY: 2012.
55. Ayad F, Belhadj M, Paries J, Attali JR, Valensi P. Association between cardiac autonomic neuropathy and hypertension and its potential influence on diabetic complications. *Diabet. Med.* 2010; 27:804–811. [PubMed: 20636962]
56. Dawson ME, Schell AM, Filion DL. The electrodermal system.. In: Cacioppo JT, Tassinary LG, Berntson GG, editors *Handbook of Psychophysiology.* 3rd ed.. Cambridge University Press; New York, NY: 2007. 159–181.
57. Donadio V, Cortelli P, Elam M, et al. Autonomic innervation in multiple system atrophy and pure autonomic failure. *J. Neurol. Neurosurg. Psychiatry.* 2010; 81:1327–1335. [PubMed: 20660924]
58. Hall NR, O'Grady MP, Menzies RA. Neuroimmunopharmacologic effects of drugs of abuse. *Adv. Exp. Med. Biol.* 1991; 288:13–23. [PubMed: 1719750]

59. Maltzman I, Marinkovic K. Alcohol, alcoholism, and the autonomic nervous system: A critical account.. In: Begleiter H, Kissin B, editors *The pharmacology of alcohol and alcohol dependence*. Oxford University Press; New York, NY: 1996. 248–306.
60. Schmidt C, Herting B, Prieur S, et al. Autonomic dysfunction in different subtypes of multiple system atrophy. *Mov. Disord.* 2008; 23:1766–1772. [PubMed: 18661564]
61. Tug T, Terzi SM, Yoldas TK. Relationship between the frequency of autonomic dysfunction and the severity of chronic obstructive pulmonary disease. *Acta Neurol. Scand.* 2005; 112:183–188. [PubMed: 16097961]
62. Van Gestel AJ, Kohler M, Steier J, Teschler S, Russi EW, Teschler H. Cardiac autonomic dysfunction and health-related quality of life in patients with chronic obstructive pulmonary disease. *Respirology.* 2011; 16:939–946. [PubMed: 21564403]
63. Wang Y, Zhao X, O'Neil A, Turner A, Liu X, Berk M. Altered cardiac autonomic nervous function in depression. *BMC Psychiatry.* 2013; 13:187. [PubMed: 23842138]
64. Koschke M, Boettger MK, Schulz S, et al. Autonomy of autonomic dysfunction in major depression. *Psychosom. Med.* 2009; 71:852–860. [PubMed: 19779146]
65. Alvares GA, Quintana DS, Hickie IB, Guastella AJ. Autonomic nervous system dysfunction in psychiatric disorders and the impact of psychotropic medications: a systematic review and meta-analysis. *J. Psychiatry Neurosci.* 2016; 41:89–104. [PubMed: 26447819]
66. Eccles JA, Owens AP, Mathias CJ, Umeda S, Critchley HD. Neurovisceral phenotypes in the expression of psychiatric symptoms. *Frontiers in neuroscience.* 2015; 9:4. [PubMed: 25713509]
67. Iodice V, Sandroni P. Autonomic neuropathies. *Continuum (Minneapolis, Minn.).* 2014; 20:1373–1397. [PubMed: 25299288]
68. Cencetti S, Lagi A, Cipriani M, Fattorini L, Bandinelli G, Bernardi L. Autonomic control of the cerebral circulation during normal and impaired peripheral circulatory control. *Heart.* 1999; 82:365–372. [PubMed: 10455091]
69. Dawson ME, Schell AM, Braaten JR, Catania JJ. Diagnostic utility of autonomic measures for major depressive disorders. *Psychiatry Res.* 1985; 15:261–270. [PubMed: 3865244]
70. Fu Q, Levine BD. Autonomic circulatory control during pregnancy in humans. *Semin. Reprod. Med.* 2009; 27:330–337. [PubMed: 19530067]
71. Leung RS. Sleep-disordered breathing: autonomic mechanisms and arrhythmias. *Prog. Cardiovasc. Dis.* 2009; 51:324–338. [PubMed: 19110134]
72. Soares-Miranda L, Sandercock G, Vale S, et al. Metabolic syndrome, physical activity and cardiac autonomic function. *Diabetes Metab. Res. Rev.* 2012; 28:363–369. [PubMed: 22238216]
73. Tentolouris N, Argyrakopoulou G, Katsilambros N. Perturbed autonomic nervous system function in metabolic syndrome. *Neuromolecular Med.* 2008; 10:169–178. [PubMed: 18224460]
74. Flesch R. A new readability yardstick. *J. Appl. Psychol.* 1948; 32:221–233. [PubMed: 18867058]
75. Kincaid JP, Fishburne LRP, Rogers RL, Chissom BS. Derivation of new readability formulas (automated readability index, Fog count, Flesch reading ease formula) for Navy enlisted personnel. Vol No. RBR-8-75. Naval Technical Training Command Millington TN Research Branch; Millington, TN: 1975.
76. Berntson GG, Quigley KS, Lozano D. Cardiovascular psychophysiology.. In: Cacioppo JT, Tassinary LG, Berntson GG, editors *Handbook of Psychophysiology*. 3rd ed.. Cambridge University Press; New York, NY: 2007. 182–210.
77. Arnold HS, MacPherson MK, Smith A. Autonomic correlates of speech versus nonspeech tasks in children and adults. *J. Speech Lang. Hear. Res.* 2014; 57:1296–1307. [PubMed: 24686989]
78. Lykken DT, Venables PH. Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology.* 1971; 8:656–672. [PubMed: 5116830]
79. Awan SN. Analysis of Dysphonia in Speech and Voice (ADSV): an application guide. KayPentax; Montvale, NJ: 2011.
80. Dietrich M, Verdolini Abbott K. Psychobiological stress reactivity and personality in persons with high and low stressor-induced extralaryngeal reactivity. *J. Speech Lang. Hear. Res.* 2014; 57:2076–2089. [PubMed: 25087852]

81. Postma-Nilsenova M, Holt E, Heyn L, Groeneveld K, Finset A. A case study of vocal features associated with galvanic skin response to stressors in a clinical interaction. *Patient Educ. Couns.* 2016; 99:1349–1354. [PubMed: 26988238]
82. Mehta DD, Van Stan JH, Hillman RE. Relationships between vocal function measures derived from an acoustic microphone and a subglottal neck-surface accelerometer. *IEEE/ACM transactions on audio, speech, and language processing.* 2016; 24:659–668.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

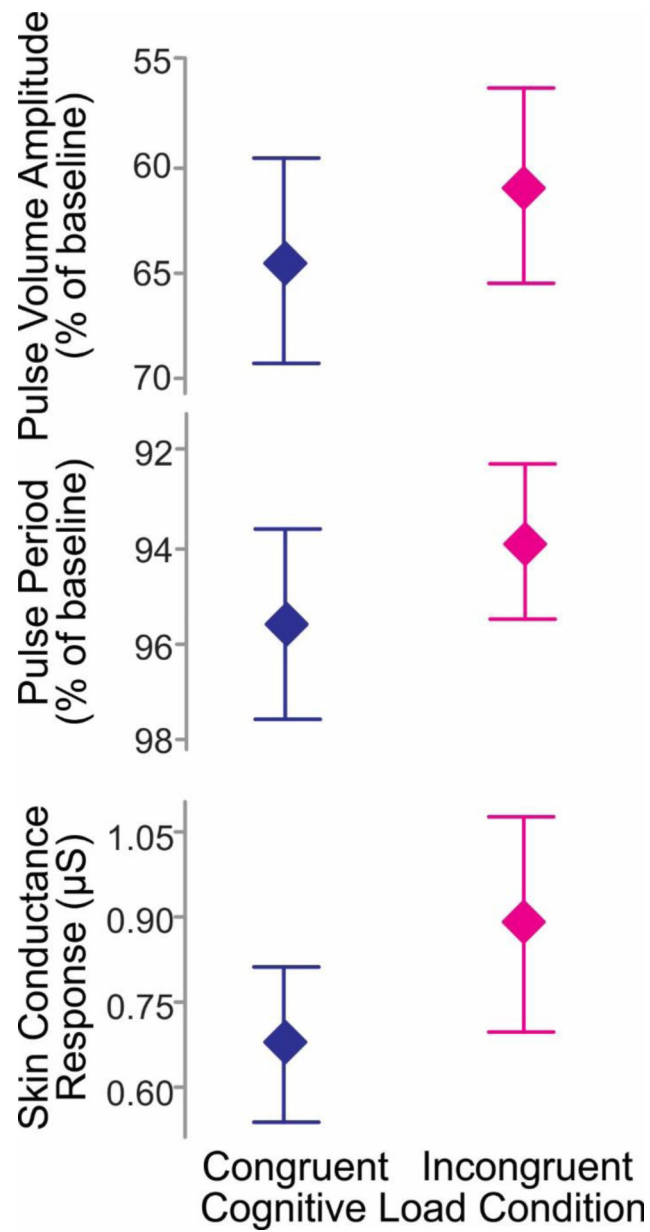


Figure 1. Autonomic measures as a function of cognitive load condition. Symbols represent means. Error bars show 95% confidence intervals. Note the reversed y-axis for upper two panels so that higher values on the y-axis reflect greater autonomic arousal.

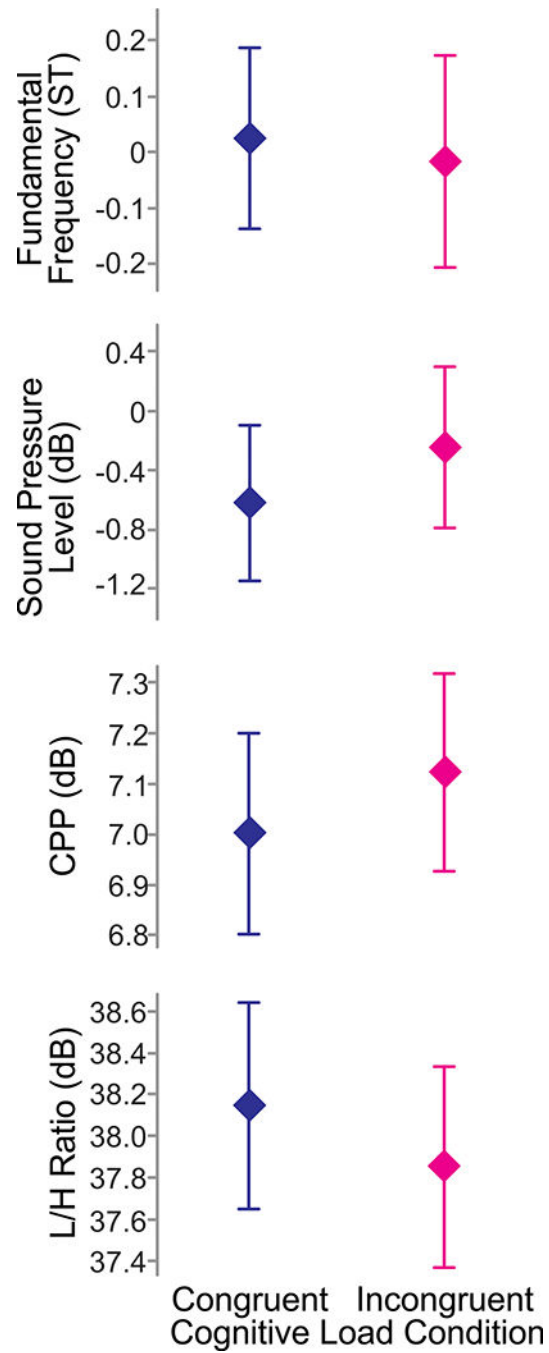


Figure 2. Acoustic measures as a function of cognitive load condition. Symbols represent means. Error bars show 95% confidence intervals.

Table 1

Results of the binary logistic regression predicting cognitive load condition.

Predictor Variables	DF	χ^2	<i>p</i>	Odds Ratio
Pulse volume amplitude (% of baseline)	1	0.40	0.526	0.996
Pulse period (% of baseline)	1	1.82	0.177	0.977
Skin conductance response amplitude ($\sqrt[3]{\mu S}$)	1	11.00	* 0.001	13.34
Fundamental frequency (ST)	1	2.40	0.121	1.08
Sound pressure level (dB)	1	0.03	0.858	.974
CPP (dB)	1	7.69	* 0.050	1.743
L/H ratio (dB)	1	8.08	* 0.004	.680
Subject	15	19.78	0.180	

* Indicates statistically significant ($p < 0.05$) predictor variables