

## Research Article

# Investigating Cognitive Load and Autonomic Arousal During Voice Production and Vocal Auditory-Motor Adaptation

Nicole E. Tomassi,<sup>a,b</sup>  Dea M. Turashvili,<sup>c</sup> Alyssa Williams,<sup>b</sup> Bridget Walsh,<sup>d</sup>  Emily P. Stephen,<sup>a,e</sup> and Cara E. Stepp<sup>a,b,c,f</sup> 

<sup>a</sup>Graduate Program for Neuroscience, Boston University, MA <sup>b</sup>Department of Speech, Language and Hearing Sciences, Boston University, MA <sup>c</sup>Department of Biomedical Engineering, Boston University, MA <sup>d</sup>Department of Communicative Sciences and Disorders, Michigan State University, East Lansing <sup>e</sup>Department of Mathematics and Statistics, Boston University, MA <sup>f</sup>Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, MA

## ARTICLE INFO

## Article History:

Received August 29, 2024

Revision received November 21, 2024

Accepted December 9, 2024

Editor-in-Chief: Erin M. Picou

Editor: Nancy Pearl Solomon

[https://doi.org/10.1044/2024\\_JSLHR-24-00601](https://doi.org/10.1044/2024_JSLHR-24-00601)

## ABSTRACT

**Purpose:** Cognitive load and autonomic arousal are hypothesized to affect voice production, yet the nature of these relationships is unclear. The purpose of this study was to assess how cognitive load and autonomic arousal differentially affect voice production and vocal motor control.

**Method:** Physiological measures of autonomic arousal were recorded from 30 adults under different cognitive loads elicited by a Stroop task. In Experiment 1, voice acoustic measures were measured during speech production. In Experiment 2, fundamental frequency ( $f_0$ ) responses to predictably altered auditory feedback (sensorimotor adaptation) were measured. Mixed linear-effects models assessed relationships between variables. Changes between cognitive loads were compared among the two experiments.

**Results:** In Experiment 1, increased cognitive load was associated with increases in sound pressure level, whereas increases in autonomic arousal measures (i.e., decreases in skin conductance rise time, pulse amplitude, and period) were related to decreases in cepstral peak prominence. Increased autonomic arousal (i.e., decreased pulse amplitude) was related to increased adaptation in Experiment 2. Participants who responded to increased cognitive load by decreasing  $f_0$  during Experiment 1 showed more adaptation in Experiment 2.

**Conclusion:** Differential effects of cognitive load and autonomic arousal emphasize the importance of individual physiological variability when assessing how stress affects the voice.

Efficient speech production requires the coordination of both higher order, neural control mechanisms and lower order, physiological mechanisms of the body. The brain and body must work together to adapt to natural environmental stressors in order to ensure speech is produced effectively. These stressors add an additional layer of complexity to the already complicated speech circuit by inducing bodily responses that impact speech subsystems. These responses are a result of an activated autonomic

nervous system (ANS), and studies have sought to understand how these physiological responses affect vocal production (i.e., Helou et al., 2020; MacPherson et al., 2017). However, there is a gap in understanding how the ANS affects the higher order, neural control processes of the voice (vocal motor control) and the coordination between vocal motor control and voice production.

The ANS controls and regulates physiological functions involving involuntary processes. This is especially relevant for the voice subsystem of speech as the neural substrates of vocal control include subcortical and limbic structures that are also responsible for regulating autonomic activity (Critchley, 2005). When the ANS is aroused,

Correspondence to Nicole E. Tomassi: [ntomassi@bu.edu](mailto:ntomassi@bu.edu). **Disclosure:** The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

the sympathetic branch is activated and the body's physiological functions change in response to the arousing stimuli (Bear et al., 2020; Hamill et al., 2012). Sympathetic innervation affects many aspects of the body, including voice-specific organs. ANS innervation of the larynx is supported by anatomical, immunohistochemical (Hisa et al., 1982, 1999; Ibanez et al., 2010; Ramaswamy et al., 1994), and physiological evidence (increased laryngeal muscle activation during ANS arousal; Helou et al., 2013).

Investigating autonomic arousal allows researchers to understand how well the body responds to stress and how those responses may impact other cognitive and motor processes, including vocal control and production. To measure sympathetic activation, researchers often target the cardiovascular and electrodermal system because they are accessible noninvasively. The signal collected to estimate cardiovascular activity detects changes in blood volume, which results from vasoconstriction during autonomic arousal (Jennings et al., 1980). The electrodermal signal measures changes in conductivity due to changes in sweat gland activity, which also fluctuate with autonomic arousal (Boucsein, 2012). To activate autonomic arousal in a laboratory setting, various methods such as cold pressor tasks (e.g., Kyle & McNeil, 2014; Lovallo, 1975), electrical stimulation (e.g., Kyle & McNeil, 2014), and the Valsalva maneuver (e.g., Hilz & Dütsch, 2006; Low, 1993) have been used in the literature. Modulating cognitive load is another common way to elicit ANS arousal and can easily be integrated into speech tasks. Cognitive load in this context refers to the amount of neural resources used to complete a task. Studies from the voice literature have modulated cognitive load in various ways including word list recall (Boyer et al., 2018), continuous attentive verbalization (Mendoza & Carballo, 1998), public speaking (Perrine & Scherer, 2020; Van Lierde et al., 2009), simulated aviation (Huttunen et al., 2011), dual tasking (Dromey & Bates, 2005; Dromey & Benson, 2003; Dromey & Shim, 2008; Lively et al., 1993; Scheerer, Tumber, & Jones, 2016), mental arithmetic (Perrine & Scherer, 2020), and Stroop tasks (Christodoulides, 2016; Dahl & Stepp, 2023; MacPherson et al., 2017). The common goal of these studies was to understand how voice mechanisms adjust when cognitive resources are taxed.

### ***Voice Production During Cognitive Loading and ANS Arousal***

Despite a large body of literature seeking to understand the impacts of cognitive load on voice production, the findings are highly variable. Specifically, with higher levels of cognitive load, fundamental frequency ( $f_0$ ) has been shown to increase (Boyer et al., 2018; Mendoza & Carballo, 1998; Perrine & Scherer, 2020), decrease (Dahl

& Stepp, 2023; Van Lierde et al., 2009), and remain the same (MacPherson et al., 2017). In some studies, this change differs at the individual level—varying by participant and potentially obscuring trends (Huttunen et al., 2011; Lively et al., 1993; Perrine & Scherer, 2020). Manipulating cognitive load also induced variable changes in sound pressure level (SPL). Several studies reveal an increase in SPL (Dahl & Stepp, 2023; Dromey & Bates, 2005; Dromey & Shim, 2008; Huttunen et al., 2011), while another study reported a decrease (Van Lierde et al., 2009). Similar to the  $f_0$  finding, Lively et al. (1993) also found changes in SPL varying at the participant level. There have been fewer studies investigating the impacts of cognitive load on vocal quality features. For example, a few studies have found differences in measures of periodicity in the acoustic signal including increased jitter and shimmer (Mendoza & Carballo, 1998), as well as increased cepstral peak prominence (CPP; MacPherson et al., 2017) with increased cognitive load. One study found that relative fundamental frequency decreased with increased cognitive load (Dahl & Stepp, 2023). Other studies have found changes in the relative strength of harmonics in the signal with increased cognitive load (Lively et al., 1993; MacPherson et al., 2017). Lastly, one study found changes in glottal source features such as increased glottal closure time with increased cognitive load (Christodoulides, 2016).

The theoretical underpinning of the effects of cognitive load on voice production involves the competition of neural demands and the intention of communication. Modulating cognitive load involves increasing the need for particular executive functions such as inhibition, working memory, and attention, necessitating a shift in the allocation of neural resources (Crystal, 1987; Dromey & Bates, 2005; Dromey & Benson, 2003). This has been shown in many goal-directed tasks that require cognitive-motor interaction and results in the degradation of one or more outputs (e.g., Al-Yahya et al., 2011; Bayot et al., 2018; Bhatt & Kannan, 2022; Keefe, 1985; Klotzbier & Schott, 2017; Kriegel, 2022). Unlike other goal-directed behavior, speech production adds an additional complexity to consider: the intention of communication. Successful communication relies both on the speaker to produce efficient speech as well as the listener to interpret the speech correctly. Thus, speakers must adjust their speech during unideal speaking situations (such as under cognitive load) to maximize efficiency by balancing the energy reserves of the speaker with the interpretation needs of the listener (Lindblom, 1990).

The above theoretical rationale supports the idea that cognitive load impacts voice production, but it remains unclear why the direction of change among the findings are so varied. A possible reason could be the

differences in experimental tasks, taxing different aspects of cognitive function and thus resulting in differing distributions of neural resources and voice adjustments (Hsu et al., 2021). Attempting to parse apart behavioral outcomes related to specific cognitive functions without neural data, however, is difficult because many cognitive loading tasks tax multiple cognitive processes. An alternative theory is that these different cognitive tasks induce different amounts of autonomic arousal, likely varying among participants, and result in inconsistent downstream biomechanical changes in the voice. Autonomic arousal and cognitive functions are inherently not independent of each other. They share overlapping neuroanatomical regions (Critchley, 2005) and are thus mechanically coupled. The relationship is complex and bidirectional such that ANS arousal can occur as a result of increasing cognitive load, but also cognitive functions are impacted by the physiological state. The degree to which this coupling applies across the range of cognitive functions is not well understood. In other words, it is not clear whether some cognitive functions elicit (or are impacted by) ANS arousal more than others. Because of these challenges, the current article uses cognitive load broadly as an increase in overall neural resources, rather than specifying exact cognitive processes, and discusses autonomic responses within this framework.

Only a few studies have modulated cognitive load while collecting both autonomic data and voice acoustics (Abur et al., 2023; Heaton et al., 2020; MacPherson et al., 2017; Perrine & Scherer, 2020). In some of these studies, cognitive load was treated as the responding variable to changes in autonomic activity and voice acoustics. For example, Heaton et al. (2020) assessed whether parameters of speech and/or autonomic activity could be used to detect cognitive fatigue with applications to inform military practices and found that electrodermal activity and vocal creak predicted cognitive performance. MacPherson et al. (2017) assessed ANS measures and voice acoustics as a function of cognitive load and found that skin conductance response (SCR) amplitude, CPP, and low-high (LH) ratio statistically predicted cognitive load. Using cognitive load as the dependent variable with autonomic activity and voice acoustics as independent variables obscures the connection between autonomic arousal and voice. Instead, by reframing these relationships and using cognitive load as a modulator to induce ANS arousal, we can understand how differences in physiological responses relate to changes in the voice. This would help in determining if the variability observed in the literature is pointing to a “trait” effect. In other words, are the changes in voice more prominent in those who experience heightened levels of ANS arousal, compared to those with more stable ANS arousal levels?

One study approached this question by using salivary cortisol levels, to determine how stress affected acoustic and aerodynamic measures of voice before, during, and after a cognitive stressor (Perrine & Scherer, 2020). Interestingly, they found increases in  $f_0$  before and after the stressor, but only in speakers who exhibited a biological response (heightened cortisol levels) to the stressor, supporting the possibility of a “trait” effect (Perrine & Scherer, 2020). This “trait” effect finding requires further support and replication to conclusively determine whether ANS arousal uniquely impacts the voice. Further, this study solely focused on downstream, or lower order voice biomechanisms, and it remains unclear how the higher order vocal motor control mechanisms are impacted by ANS arousal. Understanding how individual differences in physiological responses to stressors contribute to differences in lower and higher order voice mechanisms could assist in developing personalized voice therapy approaches, tailored stress management techniques, and more effective communication skills.

### ***Vocal Motor Control During Cognitive Loading and ANS Arousal***

Most of these prior studies have focused exclusively on the effects of ANS arousal or cognitive load on the downstream processes of voice production. Assessing these effects on vocal motor control, however, would provide a greater understanding of the neurological processes involved in how individuals monitor their voice to maintain this production. Following a popular model originally developed for articulation, the DIVA (Directions Into Velocities of Articulators) model describes two control systems working in parallel to regulate fluent speech (Tourville & Guenther, 2011). The feedforward control system uses stored motor programs to plan and drive efficient speech. Experimentally studying the functionality of feedforward control typically involves presenting persistent and expected errors to auditory feedback and observing how the system updates and integrates these motor programs to correct for future errors (i.e., sensorimotor adaptation; Jones & Munhall, 2000; Patel et al., 2011).

Sensorimotor adaptation is a critical component of goal-directed behavior. From the extensive literature investigating the neural mechanisms of sensorimotor adaptation, three distinct brain regions have surfaced: the anterior cingulate cortex, the cerebellum, and the basal ganglia (see review: Seidler et al., 2013). The anterior cingulate cortex and the cerebellum are also important for autonomic regulation and cognitive function; moreover, the basal ganglia also play a role in cognitive function (Middleton & Strick, 1994). The overlapping functions among these brain regions provide reasoning as to why there may be

relationships among sensorimotor adaptation, autonomic arousal, and cognition. However, only one prior study has described the experimental effects of cognitive load on sensorimotor adaptation of voice. Specifically, Scheerer, Tumber, and Jones (2016) investigated the impacts of attention on sensorimotor adaptation of voice. This study used persistent auditory errors of voice  $f_0$  to determine how the feedforward control system would adapt in differing attentional demand conditions. Scheerer, Tumber, and Jones found decreases in the ability to adapt vocal motor commands in response to persistent changes to  $f_0$  when attentional demands were greater. Because autonomic activity was not recorded in their study, it is unclear if the reported effects were due to changes in cognitive processes, or whether ANS arousal contributed to changes recorded in  $f_0$ . There are likely influences from both task demands as well as physiological responses to the demands that impact sensorimotor adaptation processes.

### **Current Study**

To better understand the effects of cognitive load and autonomic arousal on the voice, it is imperative to take a comprehensive approach and investigate how these systems relate to (a) the lower order voice output, (b) the higher order planning processes involved in monitoring and adapting the voice, and (c) the relationship between these lower and higher order voice mechanisms. Voice production is a fine-tuned behavior that requires resources from both lower and higher order levels. Any cognitive process that is taxed takes resources away from the control of voice, whether it is downstream control of laryngeal position and tension to ensure proper suprasegmental features are produced or upstream control of auditory error-monitoring to ensure acoustic targets align with intended goals. Downstream processes are more likely to be affected by the autonomic arousal that occurs as a result from cognitive loading as they require bodily resources that are innervated by the ANS. Higher order vocal motor control processes are more likely to be directly affected by cognitive loading as they require neural mechanisms drawing from multiple areas and networks that overlap with cognition. Since suprasegmental aspects of speech can fluctuate a fair amount and still convey the intended message, it is likely that the resources dedicated to monitoring and maintaining these features are flexible. Thus, we might expect vocal motor control mechanisms (i.e., sensorimotor adaptation) to be impacted during cognitive loading. These aims fill the gap in the extant literature, allowing us to understand how the brain and body work together to adapt to stressful situations. Behavioral studies investigating vocal motor control through altered auditory feedback paradigms rarely discuss the connection between the higher order processes being targeted by these

paradigms and the lower order, acoustic output in typical speech production. Speakers whose higher order feedforward control mechanisms are more flexible and thus affected by cognitive load and/or autonomic arousal may also be more likely to adjust their lower order voice mechanisms during typical running speech as they experience a potential disturbance in trying to convey their message. Observing a response to stress in downstream voice production as well as upstream vocal motor control would illustrate a strong coordination between these processes and would provide support for more holistic clinical interventions.

Thus, the purpose of this study was to describe the roles of ANS arousal and cognitive load on voice production and vocal motor control through both acoustic measures of voice and sensorimotor adaptation. Two experiments were performed to address these goals—one involving the sentence-level speech tasks targeting lower order voice production (Experiment 1) and the other involving single-word utterances with altered auditory feedback to target higher order vocal motor adaptation (Experiment 2). Both experiments involved a low- and high-cognitive-load condition to modulate autonomic arousal. Autonomic activity was simultaneously collected in both experiments, using cardiovascular and electrodermal measures. In Experiment 1, ANS measures and cognitive load were used as predictors to determine if changes in acoustic measures of voice were associated with heightened autonomic activity and cognitive demand. We hypothesized that physiological responses to cognitive loading, rather than the demands of the task itself, would affect voice production. Thus, we predicted that ANS arousal would lead to changes in acoustic features of voice production including an increase in average SPL,  $f_0$ , smoothed cepstral peak prominence (CPPS), and LH ratio. In Experiment 2, ANS measures and cognitive load were used as predictors to determine if changes in sensorimotor adaptation were related to heightened autonomic activity and cognitive demand. Based on a prior finding that increased attentional demands were related to decreased voice adaptation (Scheerer, Tumber, & Jones, 2016), we hypothesized that both cognitive load and ANS arousal measures would negatively affect adaptive responses to  $f_0$  perturbations. Lastly, we hypothesized that the between-condition changes in adaptive responses observed in Experiment 2 would be directly related to the between-condition changes in acoustic measures of voice observed in Experiment 1, reflecting the individual, trait-nature of the biomechanisms of the voice.

## **Method**

### **Participants**

All participants completed written consent in compliance with the Boston University institutional review

board. Participants were native speakers of American English (15 cisgender males, 15 cisgender females;  $M_{\text{age}} = 22.75$ ,  $SD = 2.62$ ) with no history of speech, language, hearing, neurological, or autonomic disorders, and no prior training in singing. To avoid potential confounding effects of ANS function or cognitive ability, participants self-reported they were nonsmokers, had no history of drug or alcohol abuse, and were not diagnosed with psychological disorders, sleep apnea, chronically low/high blood pressure, dermatological conditions, or hyperhidrosis. None of the participants reported taking prescription medications known to affect cognitive function. To control for ANS fluctuations, participants abstained from the consumption of caffeine, alcohol, and heavy meals and had not endured physical activity or stressful events at least 3 hr prior to the experiment. Although this criterion was verbally confirmed prior to beginning the experiment, the type of physical activity or stressful events had not been specified and was left up to the discretion of the participants.

All participants passed a pure-tone hearing screening using a Grason-Stadler GSI-18 audiometer and over-ear headphones at 25 dB HL at frequencies of 125, 250, 500, 1000, 2000, 4000, and 8000 Hz (American Speech-Language-Hearing Association, 2020). All participants passed a color vision screening with the Ishihara Color Blindness Test (Ishihara, 1996) to ensure they would be able to complete task requirements involving color identification. Cognition was assessed using the Montreal Cognitive Assessment. All participants scored over 26 ( $M = 28.3$ ,  $SD = 1.5$ ), confirming cognitive function was typical (Nasreddine et al., 2005) and would not pose as a potential confound. Participants also completed a Composite Autonomic Symptom Score (COMPASS)-31 exam, a survey of ANS symptoms that scores symptom severity on a weighted scale from 0 to 100 (Sletten et al., 2012). Scores between 20–40 and 40–100 have been used in the literature to indicate moderate and severe autonomic symptoms (Kedor et al., 2022; Sletten et al., 2012), respectively. COMPASS-31 scores for the current study ranged from 0 to 21.75 ( $M = 7.6$ ,  $SD = 7.8$ ). All surveys, screenings, and experiments were completed in one, 3-hr session.

## Experimental Design

### Experiment 1: Sentence Production

Experiment 1 consisted of sentences with embedded Stroop conditions (Stroop, 1935), a well-established technique used to manipulate cognitive load during speech and other motor tasks (e.g., Abur et al., 2023; Hoshikawa & Yamamoto, 1997; Janssen et al., 2019; MacPherson et al., 2017). Each trial was a single sentence containing three colored words in the middle (see Appendix). Participants read

the sentence aloud and said the color of the font, rather than the word, when they approached a colored word. The experiment was designed using four sets of four blocks (see Figure 1, Panel A). Each block consisted of three distinct sentences totaling 48 sentences. The sentences were matched for word length (11 words) and each sentence appeared an equal number of times (16). The sentences were pseudorandomized such that each sentence appeared in the block only once. In each set, the four blocks had two congruent and two incongruent blocks. The color of the font matched the words in the congruent (low cognitive load) block, but not in the incongruent (high cognitive load) block. The incongruent blocks always followed the congruent blocks because once the system is aroused, it remains aroused for a period of time (SCR recovery time  $\geq 20$  s; Boucsein, 2012).

Each block began with a 30-s period for ANS measurements to be recorded at rest. This time frame was found to be sufficient for ANS measures to return to an interstimulus resting state (Weber & Smith, 1990). During this time, participants were instructed to relax and breathe calmly. Each sentence was followed by an 8-s intertrial interval to allow the ANS to respond to the stimulus (SCR latency  $\sim 1$ –3 s, SCR rise time  $\sim 1$ –3 s, SCR half-recovery time  $\sim 2$ –10 s; Boucsein, 2012; Dawson et al., 2017). Between each block, a “foil” sentence containing no colored words was presented to prevent ANS habituation.

### Experiment 2: Sensorimotor Adaptation

During Experiment 2 participants were exposed to sustained, predictable changes in auditory feedback. Recordings started 30 s before the first stimulus (for baseline ANS measures to be collected) and there was an intertrial interval of 1 s. The stimuli (see Figure 1, Panel C) consisted of single words (“red,” “yellow,” “black”) that either matched the color of the font in all trials (congruent condition) or matched in some trials but not in others (incongruent condition). During the incongruent conditions, 33% of the trials were incongruent and the remaining 66% were congruent to avoid autonomic habituation. Presentation of the incongruent stimuli was semirandomized such that incongruent trials were not sequential. Participants were asked to produce the color of the font and sustain the first vowel of that color (/æ/ in black or /ɜ/ in red and yellow) for approximately 1 s.

Participants completed tasks with shifted feedback for both cognitive load conditions (low/congruent and high/incongruent) as well as control tasks in which participants received unperturbed feedback for all trials for a total of four tasks ( $f_0$  shift  $\times$  2 cognitive load conditions + control  $\times$  2 cognitive load conditions). The order of tasks was counterbalanced across participants. The “shift-on” tasks had four phases of varying auditory feedback (see

**Figure 1.** (Panel A: Experiment 1 design) Each participant completed four sets of four blocks. The sets contained two congruent and two incongruent blocks. The congruent blocks consisted of three sentences in which the imbedded colors matched the color of the phonemic description. The incongruent blocks consisted of three sentences in which the imbedded colors did not match the phonemic description. (Panel B: Experiment 2 perturbation design) Participants completed adaptation tasks in which they were exposed to gradual changes in auditory feedback. The baseline phase consisted of unperturbed feedback, the ramp phase was shifted downward by 3.13 cents of fundamental frequency ( $f_0$ ) with each successive trial, the hold phase maintained a downward shift of 100 cents of  $f_0$ , and the after-effect phase returned to unperturbed feedback. (Panel C: Experiment 2 stimuli) Participants said the color of the word rather than the phoneme. They were asked to sustain the first vowel of the color for 1 s.

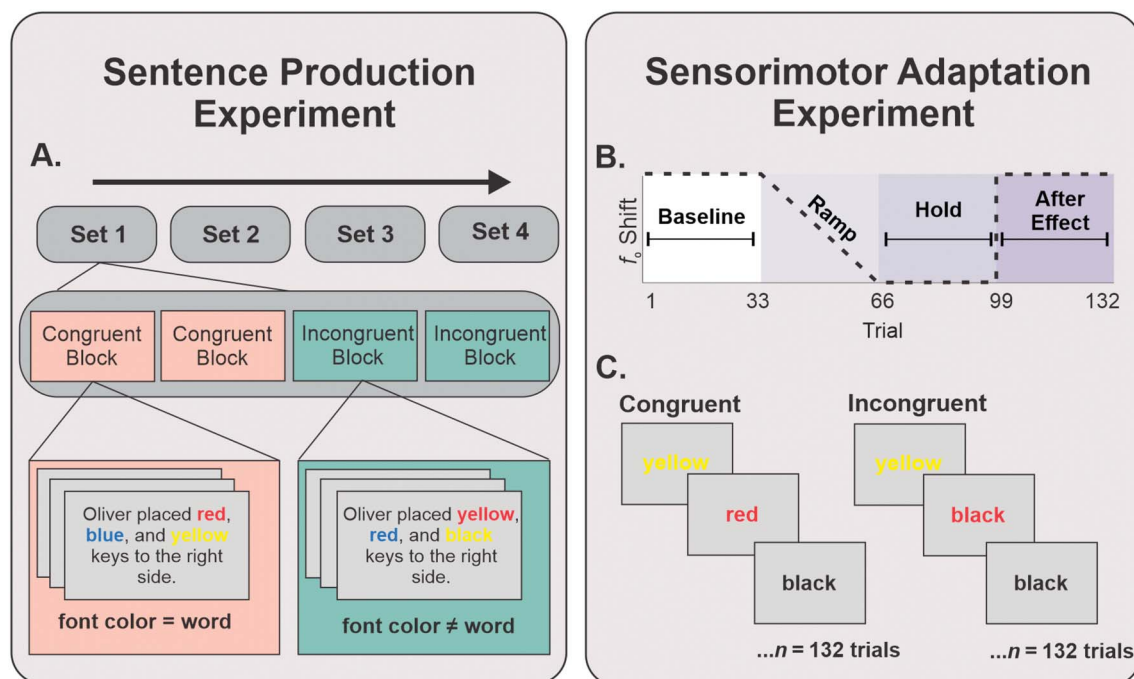


Figure 1, Panel B): baseline, ramp, hold, and after-effect. Each phase consisted of 33 trials for a total of 132 trials.

During the “shift-on” tasks,  $f_0$  was shifted downward over the course of the four phases (see Figure 1, Panel B). The shift direction was chosen to avoid a floor effect from an upward shift, as participants typically use lower ranges of their voice profile (Berg et al., 2017). The first phase (baseline) consisted of 33 trials of unperturbed auditory feedback. The second phase (ramp) consisted of 33 trials gradually shifted downward by 3.13 cents with each successive trial, reaching 100 cents below the participant’s true  $f_0$ . The following 33 trials in the hold phase maintained this level of feedback. The last phase (after-effect) consisted of 33 trials in which no perturbation was applied.

### Instrumentation, Signal Processing, and Data Analysis

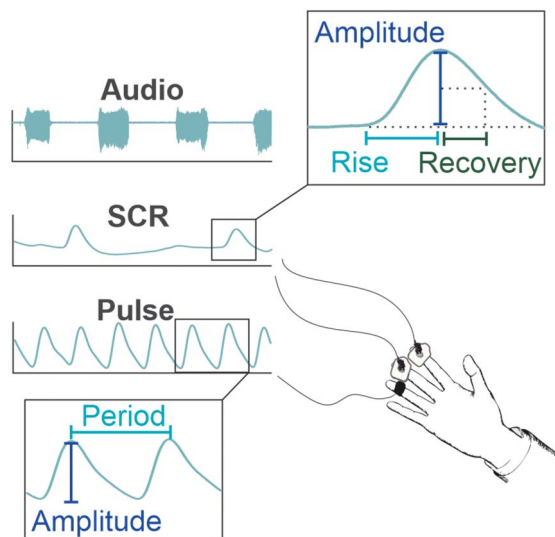
#### EDA Collection and Analysis

Electrodermal activity (EDA) signals were conditioned using a Biopac MP150 Data Acquisition System (Biopac

Systems, Inc.) and sampled at 14700 Hz using a National Instrument data acquisition card. Skin conductance was sensed through two EDA electrodes (EL507, Biopac Systems, Inc.) attached to the palmar surface of the first and second distal phalange on the right hand (see Figure 2). The skin conductance signal estimates the level of activity of the eccrine sweat glands (Dawson et al., 2017) and increases during autonomic arousal. The tonic skin conductance level was amplified with a gain of 10  $\mu\text{S}/\text{V}$  via a constant voltage (0.5 V) Biopac GSR EDA Galvanic Skin Response Amplifier (GSR100C). The SCR was derived from the tonic signal using a second-order Chebyshev high-pass filter with a cutoff frequency of 0.07 Hz.

Electrodermal measures included phasic and temporal measures of activity. Phasic measures included *skin conductance response amplitude* (SCR amplitude) and the *percent of event responses* (%events), which both increase with increased ANS arousal (Boucsein, 2012; Dawson et al., 2017) and are frequently used in the literature. Temporal measures included *SCR rise time* and *half-recovery time*; although less frequently used in the literature, these can provide complementary insight into sympathetic

**Figure 2.** Autonomic data collection and analysis. Placement of electrodes, capturing skin conductance response (SCR) data, and the phototransducer, capturing pulse data, on fingers is illustrated. The signals that are collected from these sensors as well as the measures extracted from the signals are described.



activity.<sup>1</sup> To derive many of the electrodermal measures, points of peak onset and maximum peak were determined manually from the filtered signal by author D.M.T. Automatic detection of peaks in electrodermal data is challenging due to the slow and highly variable nature of the responses. Unlike other more regular and cyclic signals that can be tracked with automated signal processing tools, the EDA signal historically requires more manual efforts. To assess rater reliability, 20% of the data was reanalyzed by author D.M.T. and an additional technician. Both intra- and interrater reliability were compared using intraclass correlation coefficient ( $r = .88$  and  $r = .97$ , respectively). The SCR amplitude was calculated as the difference between these points. Based on the contemporary literature, an amplitude threshold of  $0.02 \mu\text{S}$  was set to determine the presence of a SCR. The percent of event responses was calculated as the number of SCR amplitudes relative to the total number of evoking stimuli, or trials (48 for Experiment 1 and 132 for Experiment 2). Rise time was calculated as the time between the onset and peak of the SCR. Half-recovery time was calculated

<sup>1</sup>Jindrová et al. (2020) found that SCR rise times, relative to SCR amplitudes, were particularly sensitive to differences in levels of arousing stimuli, with shorter rise times associated with higher arousal. The meaning and implications of half-recovery time have been debated in the literature with one dominant argument suggesting that it may be associated with attentional “openness” (Boucsein, 2012).

as the time between the peak of the signal and when the signal returns to half of the amplitude height.

### Cardiovascular Collection and Analysis

Cardiovascular signals were conditioned and sampled using the same instrumentation and frequency as the EDA signals. Pulse was sensed through a Photo Plethysmogram Transducer (TSD200, Biopac Systems, Inc.) attached to the palmar surface of the third finger distal phalange on the right hand (see Figure 2). The transducer emits an infrared signal, which is then reflected back based on the amount of blood flow in the finger-bed capillaries (Cacioppo et al., 2000). The reflected signal increases with increased blood flow, varying indirectly with autonomic arousal due to vasoconstriction slowing blood flow. The signal was amplified with a gain factor of 100 via a Biopac Pulse Plethysmogram Amplifier (PPG100C, Biopac Systems, Inc.). Post-acquisition processing of the pulse signal consisted of bandpass filtering from 0.5 to 3 Hz with a 100-order finite impulse response filter and downsampling to 250 Hz.

From the filtered signal, *pulse volume amplitude* and *pulse period* were extracted as the measures of cardiovascular activity. These measures were chosen based on strong evidence from the literature that they are sensitive to ANS function and change in response to speech (Kleinow & Smith, 2006; Weber & Smith, 1990) and are understood to decrease with increased ANS arousal (Jennings et al., 1980). To extract these measures, the filtered signal was sent through an automated peak-detection algorithm (find-peaks.m, MATLAB release Version 2018a); peaks were manually verified by authors N.E.T. and D.M.T. Two 5-s windows were chosen in which the pulse signal exhibited the greatest amplitude—indicating the lowest level of arousal during the 30-s rest before each experiment. Pulse volume amplitude was calculated as the difference between adjacent maximum and minimum peaks. Pulse period was calculated as the time between adjacent peaks. To account for the natural between-participant variability that is often observed in these measures, they were normalized as a percent of the average during the two 5-s rest windows.

### Audio Collection and Analysis

Audio signals were pre-amplified using an RME Quadmic II and sampled using an RME Fireface UCX sound card at a rate of 44100 Hz and 32-bit resolution. A SHURE MX153 omnidirectional microphone was positioned  $45^\circ$  from the midline and 7 cm from the lips. SONAR Artist acoustic software was used for recording audio during Experiment 1. For Experiment 2, participants wore Sennheiser HD 280 Pro over-ear headphones. A Bruel & Kjaer coupler and sound level meter was used to calibrate auditory feedback to +5 dB relative to the

microphone to overcome the bone-conducted auditory feedback while minimizing unnecessarily high SPLs (Weerathunge et al., 2020). Manipulations of  $f_o$  were accomplished using the Eventide Eclipse V4 Harmonizer. Processing delay between speech production and delivery of auditory feedback was 11 ms (Heller Murray et al., 2019), which is less than the amount individuals can perceive (30 ms; Yates, 1963).

For data analyses of Experiment 1, pauses were removed, and measures were extracted from the continuous speech signal of the full sentences. To remove pauses, the signal was low-pass filtered and the envelope was used to determine a voicing threshold. This threshold was applied to the unfiltered signal to remove within-trial pauses. The final signal was manually examined (by author A.W.) to ensure all speech was maintained. Custom Praat (Versions 5 and 6.0.40; Boersma & Weenink, 2016) and MATLAB (Release 2018; MathWorks) scripts were used to calculate the acoustic measures, which have been widely used as objective correlates of auditory-perceptual features. These included SPL as a correlate of vocal loudness, mean  $f_o$  as a correlate of vocal pitch, and CPPS and LH ratio as correlates of voice quality. To calculate average SPL, the root-mean-square (RMS) value was determined for each sentence. RMS values were then converted to decibels relative to each participant's average RMS of the congruent condition. The mean  $f_o$  of each sentence was extracted and converted to semitones relative to each participant's mean  $f_o$  of the congruent condition. CPPS, defined as the most prominent peak in the cepstrum (the inverse transform of the logarithm of the signal spectrum), was calculated in Praat as described in Watts et al. (2017). The LH ratio was computed as the distribution of spectral energy in the signal using a 4000-Hz cutoff.

For  $f_o$  adaptation analysis in Experiment 2, estimation of  $f_o$  was performed offline using an autocorrelation method via Praat scripts. The mean  $f_o$  of every trial was calculated via custom MATLAB scripts over a time window of 40–120 ms (Lester-Smith et al., 2020) after vowel onset, to avoid coarticulation effects and feedback-based adjustments (Burnett & Larson, 2002; Larson et al., 2001; Tourville et al., 2008), and converted to cents relative to the mean  $f_o$  of the baseline phase. The control conditions were then subtracted from the shift conditions to normalize for prosody and natural variability over time (e.g., Chen et al., 2007). The normalized mean  $f_o$  of each phase for all participants was used in the statistical models. The normalized  $f_o$  during the after-effect phase was examined at the group level to determine if adaptation occurred. Regardless of if adaptation had occurred, all responses were included in the analyses.

## Statistical Analysis

Minitab 21.2 was used for all statistical analyses and an alpha value of .05 was set a priori. Before testing our specific study hypotheses, we assessed experimental fidelity by determining if the cognitive loading task effectively elicited ANS arousal in both experiments (see Appendix). Since voice and autonomic measures are known to be sensitive and variable, we assessed the reliability of these measures within participants in Experiment 1. We used Pearson's correlation coefficient to compare within-participant data in Sets 1 and 3 as well as Sets 2 and 4 to determine if these measures were reliable. To determine if sensorimotor integration occurred in the Experiment 2, a paired  $t$  test was applied to the  $f_o$  responses during the first trial of the after-effect phase to assess differences from baseline (Miller et al., 2023). Cohen's  $d$  was used to quantify the effect size with 0.2, 0.5, and 0.8 interpreted as small, medium, and large benchmarks, respectively (Cohen, 1977). To interpret the main study findings, partial eta squared was used to determine effect sizes, with .01, .09, and .25 as thresholds for small, medium, and large effects, respectively (Witte & Witte, 2017).

We hypothesized that ANS arousal would relate to voice production; to test this, linear mixed-effects models were used to understand the role of ANS arousal and cognitive load on acoustic measures of voice during running speech (Experiment 1). Separate models were used for each acoustic outcome of interest ( $f_o$ , SPL, LH ratio, and CPPS). All autonomic measures were incorporated as continuous variables. Participants were incorporated as random factors and cognitive load condition (incongruent/congruent) was a categorical predictor in each model.

We also hypothesized that ANS arousal would relate to vocal motor planning; to address this, a linear mixed-effects model was used to assess the role of ANS arousal and cognitive load on  $f_o$  responses to altered auditory feedback (Experiment 2). All autonomic measures were included as continuous variables. Cognitive load condition and the phase of auditory feedback perturbation (baseline/ramp/hold/after-effect) were entered as categorical predictors. Participants were incorporated as random factors. Interactions among the fixed variables were also assessed. Since cognitive load and ANS arousal are related, it is possible that the inclusion of both of these factors in the models of Experiments 1 and 2 could hinder the effects on the outcome variables. To ensure this would not be the case and to test for any mediation effects, additional statistical analyses were conducted and reported in the Appendix.

Finally, we hypothesized a relationship among the between-condition change in vocal production (Experiment 1)



and vocal motor control mechanisms (Experiment 2), which were assessed via correlation analyses. Specifically, the between-condition changes in acoustic measures ( $f_o$ , SPL, LH ratio, and CPPS) during Experiment 1 were compared with between-condition changes in adaptive responses (i.e.,  $f_o$  response during first trial of the after-effect phase) during Experiment 2 using Spearman's rank correlation coefficients. Using Grubb's outlier test, we determined an outlier in the change in  $f_o$  and CPPS data from Experiment 1. This motivated the use of Spearman's rank correlation coefficient as opposed to Pearson's correlation, as it is less sensitive to outliers. Bonferroni correction was used to account for multiple comparisons and an adjusted alpha value was set as .0125.

## Results

### **Experiment 1: Sentence Production**

Group-averaged autonomic data and voice acoustic data are described in Figure 3. All measures of autonomic arousal as well as voice outcomes were found to be reliable within participants. Results of the mixed-effects models assessing the effects of cognitive load and autonomic arousal on each voice acoustic measure are found in Table 1. There was a statistically significant effect of condition on SPL with a medium-large effect size such that increased cognitive load was associated with increased SPL. Further, there were statistically significant, positive associations of pulse volume amplitude, pulse period, and SCR rise time with CPPS, with medium-large effect sizes. No other variables or interactions were significantly related to any of the acoustic measures.

### **Experiment 2: Sensorimotor Adaptation**

Group-averaged autonomic data (Panel A) and the  $f_o$  responses to the altered auditory feedback (Panel B) are shown in Figure 4. These combined responses were assessed to determine if speakers effectively adapted to the altered auditory feedback paradigm. Mean  $f_o$  response during the first trial of the after-effect phase was 44.2 cents ( $SD = 88.2$  cents) and was statistically different from the baseline phase with a medium effect size. Results of the mixed-effects model assessing the effects of cognitive load and autonomic arousal on  $f_o$  responses to altered auditory feedback are found in Table 2. Pulse volume amplitude was significantly related to  $f_o$  responses with a medium effect size such that lower pulse volume amplitudes (i.e., greater ANS arousal) were related to larger  $f_o$  responses (greater compensation to the frequency shift). No other factors or interactions were statistically significant.

### **Relationship Between Experiment 1 and Experiment 2**

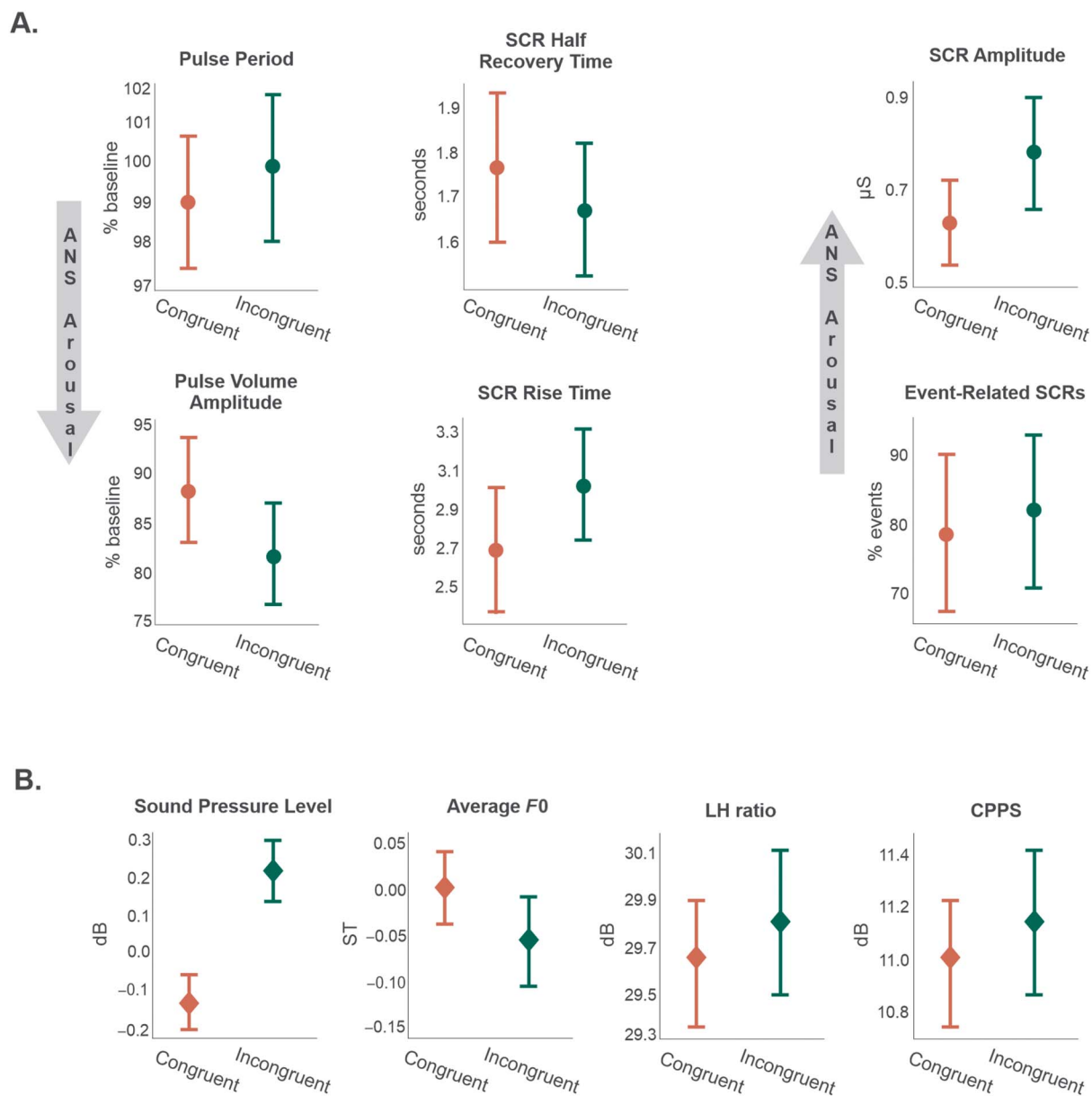
Correlation analyses revealed a statistically significant relationship among the between-condition change in  $f_o$  responses during the after-effect phase in Experiment 2 and the between-condition change in  $f_o$  during Experiment 1 (see Figure 5). Specifically, there was a negative relationship such that increased  $f_o$  responses (i.e., greater adaptation) between cognitive load conditions in Experiment 2 were associated with decreased  $f_o$  responses between cognitive load conditions in Experiment 1. No other acoustic measures from Experiment 1 were statistically related to  $f_o$  responses during the after-effect phase in Experiment 2.

## Discussion

### **Effects of Autonomic Arousal and Cognitive Load on Voice Production**

The purpose of Experiment 1 was to differentiate the effects of autonomic arousal and cognitive load on voice production. We predicted physiological responses to cognitive load, rather than the demands of the task itself, would induce changes to voice acoustics (i.e., significant effect of autonomic arousal, but not condition). Autonomic activity was only found to be statistically associated with CPPS, with decreases in pulse volume amplitude, pulse period, and SCR rise time associated with decreases in CPPS. Decreases in these ANS measures indicate increased arousal, whereas decreased CPPS values are typically associated with higher levels of noise in the acoustic signal (Patel et al., 2018). Thus, this suggests that those who experienced higher levels of autonomic arousal also had decreased periodicity, or more irregular vocal fold vibrations. This result is particularly interesting because it reveals complexities in the mechanisms of how the body responds to ANS arousal and cognitive load. MacPherson et al. (2017) found increased CPPS with high cognitive load and they interpreted this finding as the voice signal being more periodic, or stable, with increased cognitive load. Although we did not find a condition effect, we noticed a trend in this direction in the experimental validation procedures (i.e., when comparing between conditions, CPPS was higher in the incongruent condition; see Figure 3, Panel B). Further, although cognitive load tends to increase CPPS and elicit ANS arousal at the group level (see Figure 3, Panel A), the statistical model suggests that individuals with higher ANS arousal actually demonstrate decreased CPPS. In general, CPPS is often oversimplified in the literature when discussing voice quality. Lower CPPS is typically found in people with voice disorders; however, this is widely dependent on the symptom

**Figure 3.** Experiment 1 results. This figure represents the autonomic (Panel A) and acoustic (Panel B) measures between cognitive load conditions. ANS arousal arrows indicate the direction of change in the respective measures that represents an activation of sympathetic activity. Error bars represent 95% confidence intervals. ANS = autonomic nervous system; SCR = skin conductance response;  $f_0$  = fundamental frequency; LH = low-high; CPPS = smoothed cepstral peak prominence;  $\mu$ S = microsiemens; dB = decibels; ST = semitones.



presentation and various symptoms can change the overall noise level in the signal (which is ultimately what this measure captures). For example, lower CPPS can be caused by increased breathiness whereas higher CPPS can be brought about with a pressed voice. Both symptoms are found in voice disorders but will present differently in CPPS values. The differences in how CPPS changes relative to cognitive load and autonomic arousal provides additional evidence of the nuances of this measure and support that this measure is individual-specific.

SPL was the only acoustic measure that was statistically related to cognitive load. This was unsurprising due to the variability observed in the literature regarding the effects of cognitive load on acoustic measures of voice. Our results corroborate findings from earlier studies of Dahl and Stepp (2023), Dromey and Bates (2005), Dromey and Shim (2008), and Huttunen et al. (2011) in that increased cognitive loads were associated with increased SPLs. This finding could be due to a number of reasons. It could be the case that increased cognitive load

**Table 1.** Results of the mixed-effects models for acoustic outcomes of Experiment 1.

Dependent variable	Effect	<i>df</i>	<i>F</i>	<i>p</i>	$\eta_p^2$	Effect size
$f_o$ (ST)	Condition	1	0.07	.79		
	SCR Amplitude	1	0.00	.98		
	SCR rise time	1	2.79	.10		
	SCR half-recovery	1	0.13	.72		
	Event-related SCRs	1	0.77	.39		
	Pulse volume amplitude	1	0.68	.41		
	Pulse period	1	1.47	.23		
LH ratio (dB)	Condition	1	0.22	.64		
	SCR amplitude	1	0.56	.46		
	SCR rise time	1	0.30	.58		
	SCR half-recovery	1	1.42	.24		
	Event-related SCRs	1	0.52	.48		
	Pulse volume amplitude	1	0.00	.98		
	Pulse period	1	0.08	.78		
CPPS (dB)	Condition	1	0.45	.51		
	SCR amplitude	1	1.01	.32		
	<b>SCR rise time</b>	<b>1</b>	<b>5.68</b>	<b>.02</b>	<b>.17</b>	<b>Medium-large</b>
	SCR half-recovery	1	0.73	.40		
	Event-related SCRs	1	0.20	.66		
	<b>Pulse volume amplitude</b>	<b>1</b>	<b>5.86</b>	<b>.02</b>	<b>.13</b>	<b>Medium-large</b>
	<b>Pulse period</b>	<b>1</b>	<b>6.66</b>	<b>.01</b>	<b>.16</b>	<b>Medium-large</b>
SPL (dB)	<b>Condition</b>	<b>1</b>	<b>4.85</b>	<b>.04</b>	<b>.15</b>	<b>Medium-large</b>
	SCR amplitude	1	0.43	.51		
	SCR rise time	1	0.37	.55		
	SCR half-recovery	1	0.59	.44		
	Event-related SCRs	1	0.19	.67		
	Pulse volume amplitude	1	0.06	.81		
	Pulse period	1	0.18	.68		

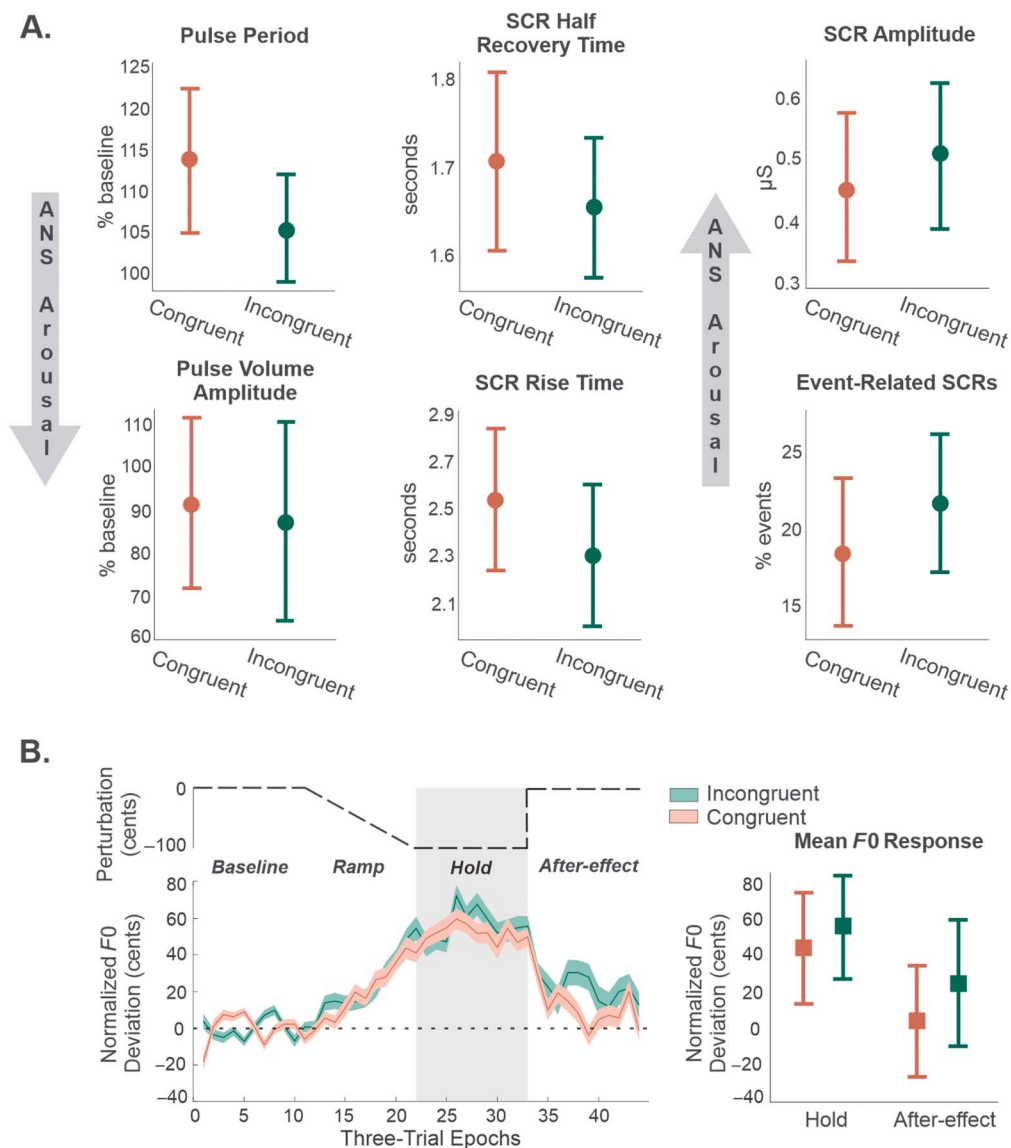
Note. Statistically significant *p* values < .05 are in bold.  $f_o$  = fundamental frequency; ST = semitones; SCR = skin conductance response; LH = low-high; dB = decibels; CPPS = smoothed cepstral peak prominence; SPL = sound pressure level.

diverted resources away from the speakers' ability to monitor their vocal parameters for efficient energy balancing and thus vocal intensity increased. Alternatively, participants might engage in compensatory strategies, such as speaking louder, to ensure their message is being conveyed despite cognitive demands. From a physiological perspective, the cognitive load could have impacted speech breathing patterns resulting in a change in vocal intensity; however, if this were the case, we would likely have observed an effect on the cardiovascular measures due to the coupling nature of the respiratory and cardiovascular systems (e.g., Schäfer et al., 1999; Taylor et al., 1999).

Cognitive load was not associated with changes in  $f_o$ , which aligns with results from MacPherson et al. (2017), whose study design was most similar to the current study in that they used imbedded Stroop words in sentences to modulate cognitive load. Other studies have reported significant effects of cognitive load on  $f_o$ , yet results of these studies are mixed—some studies have found increased  $f_o$

(Boyer et al., 2018; Mendoza & Carballo, 1998; Perrine & Scherer, 2020), whereas others have found decreased  $f_o$  (Dahl & Stepp, 2023; Van Lierde et al., 2009). Discrepancies between results are likely due to differences in cognitive tasks. These studies employ tasks that tax differing executive functions and thus could result in differing allocation of neural resources toward voice production. The variability in findings could also be attributed to differences in sex ratios in the studied participants, altering the distributions of biological ranges in  $f_o$ . For example, in studies that reported decreased  $f_o$  with increased cognitive load, the samples were comprised of all female speakers (Dahl & Stepp, 2023; Van Lierde et al., 2009). Whereas in most of the studies that found increased  $f_o$  with increased cognitive load, the samples had both female and male speakers (Boyer et al., 2018; Mendoza & Carballo, 1998). It is important to include both sexes in voice research to ensure a full range of fundamental frequencies are being assessed, and to make the research generalizable and applicable for the wider population.

**Figure 4.** Experiment 2 results. Panel A represents the autonomic measures between cognitive load condition. ANS arousal arrows indicate the direction of change in the respective measures that represents an activation of sympathetic activity (or ANS arousal). Panel B represents the acoustic behavioral responses of  $f_0$  over the course of the adaptation experiment (left). Average  $f_0$  response during the full hold and after effect phases (right). Error bars and shading represent 95% confidence intervals. ANS = autonomic nervous system; SCR = skin conductance response;  $f_0$  = fundamental frequency;  $\mu$ S = microsiemens.



We did not find a condition effect on either acoustic correlate of voice quality (CPPS, LH ratio) and because these measures are highly understudied in the literature with respect to changes in cognitive load, more work is warranted to understand if these results are typical. In general, the minimal condition effects on voice parameters could be due to the inclusion of the ANS metrics in the model which are likely accounting for overlapping variability. The inclusion of ANS metrics in this study was to attempt to parse the individual variability observed in the literature from studies that did not include physiological

data. Our results suggest that including bodily responses to cognitive loading tasks may be important in deciphering which vocal parameters are sensitive to changes specific to the task and which are sensitive to changes in the body.

### **Effects of Autonomic Arousal and Cognitive Load on Vocal Motor Control**

The purpose of Experiment 2 was to investigate the effects of autonomic arousal and cognitive load on vocal

**Table 2.** Results of the mixed-effects model for  $f_0$  responses in Experiment 2.

Dependent variable	Effect	df	F	p	$\eta_p^2$	Effect size
$f_0$ sensorimotor adaptation responses	Condition	1	1.18	.28		
	SCR amplitude	1	2.25	.14		
	SCR rise time	1	2.14	.15		
	SCR half-recovery	1	3.19	.07		
	Event-Related SCRs	1	11.75	.52		
	<b>Pulse volume amplitude</b>	<b>1</b>	<b>13.11</b>	<b>&lt; .01</b>	<b>.09</b>	<b>Medium</b>
	Pulse period	1	0.29	.58		
	<b>Phase</b>	<b>3</b>	<b>11.50</b>	<b>&lt; .01</b>	<b>.15</b>	<b>Medium-large</b>
Phase $\times$ Condition	3	0.23	.87			

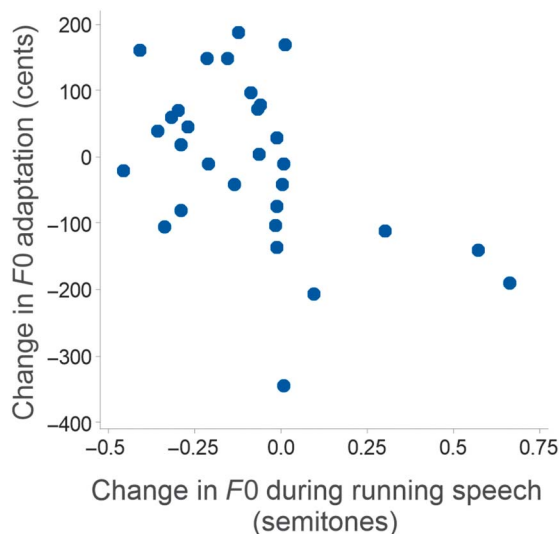
Note. Statistically significant  $p$  values  $< .05$  are in bold.  $f_0$  = fundamental frequency; SCR = skin conductance response.

motor control. We predicted both the physiological responses and the cognitive load itself would relate to voice adaptation. To the authors' knowledge, this study was the first to report the use of Stroop stimuli during an altered auditory feedback experiment and it was shown that participants updated their motor programs to this type of cognitively demanding stimuli. This result verifies the use of cognitively challenging stimuli in voice sensorimotor adaptation experiments, increasing the ecological validity for these types of experiments in future studies. Results of the mixed-effects model revealed pulse volume amplitude as a significant predictor of  $f_0$  responses to altered auditory feedback. The direction of this relationship was surprising: lower pulse volume amplitudes, indicative of *increased ANS arousal*, were associated with *larger  $f_0$  responses* (i.e., greater adaptation) to altered

auditory feedback. This result contradicts our hypothesis that increased ANS arousal would result in *decreased  $f_0$*  responses to altered auditory feedback.

When the ANS is aroused, the body is selective in how it responds to the external environment and diverts resources to the most important target (Mather & Sutherland, 2011). Moreover, research has shown that ANS arousal often relates to decreased voluntary control and increased responsivity to notable stimuli (Alexander et al., 2007; Arnsten, 2009; Liston et al., 2009; Wass et al., 2019). Thus, increased ANS arousal may increase adaptation due to heightened sympathetic activity increasing reactivity to the salient stimuli presented in altered auditory feedback experiments. In the present task, there were two external changes that probed responses, one requiring more overt cognitive resources (Stroop task) and one that was meant to elicit involuntary voice changes through altered auditory feedback. Based on our results, increased ANS arousal likely shifted resources toward the most important perceived change to our participants (i.e., their voice), which was observed as an increase in involuntary voice adaptation mechanisms.

**Figure 5.** Voice production versus vocal motor control. Relationship among the between-condition change in fundamental frequency ( $f_0$ ) during Experiment 1 and between-condition change in  $f_0$  adaptation during Experiment 2. Spearman's  $\rho = -0.46$  and  $p$  value = .01.



We observed an association between a cardiovascular measure and voice adaptation, but not any associations with electrodermal measures. One reason could be related to the neural resources responsible for regulating the cardiovascular and electrodermal systems. Cardiovascular measures are largely controlled by the medulla oblongata (Gordan et al., 2015), which is involved in receiving sensory inputs (such as auditory information) from the cerebellum—a critical region for motor learning (Ito, 2000). In contrast, electrodermal measures have been found to correlate with activity of neural areas associated with more voluntary, cognitive and emotional processes such as the anterior cingulate cortex, and prefrontal cortex (Critchley, 2005; Critchley et al., 2000). Thus, it is possible that the autonomic results in the current study reflect a differing responsivity to stimuli in the following manner: the cardiovascular activity demonstrates the aroused

system attending to the altered auditory feedback, whereas the electrodermal activity reflects the aroused system attending to the cognitive task. This discrepancy in responses can be observed in Figure 4, Panel A: the electrodermal responses show larger changes in most measures between cognitive load conditions compared with the cardiovascular responses.

The lack of a condition effect on  $f_0$  responses contradicted our original study hypothesis that increased cognitive load would diminish sensorimotor adaptation. This hypothesis was formulated based on findings from the limb literature (Redding & Wallace, 1985; Taylor & Thoroughman, 2007, 2008) and one prior study of voice that found increases in attentional demand decreased  $f_0$  adaptation (Scheerer, Tumber, & Jones, 2016). It is possible that voice adaptation is less sensitive to increased cognitive load, compared to limb adaptation, due to the finer precision required to meet motor targets and acoustic goals. It is not surprising that our null condition effect finding differs from Scheerer, Tumber, and Jones (2016), as there are substantial methodological differences between the studies. First, the current study incorporated data from all participants, whereas Scheerer, Tumber, and Jones excluded almost half (14/30 participants) due to noncompensatory responses in the low attentional load condition. This practice of removing responses based on the category of behavioral response (i.e., “compensator,” “follower,” “non-responder”) is not uncommon in the literature (e.g., Burnett et al., 1998; Scheerer, Jacobson, & Jones, 2016). However, a recent investigation pooled data from a large sample of studies and found all of these responses are from the same normal distribution and that “followers” are simply at one tail of the distribution (Miller et al., 2023). The authors of this mega-analysis suggested that the removal of these types of responses could lead to false conclusions (Miller et al., 2023). Another possibility for the differences in results of the current study compared to Scheerer, Tumber, and Jones are the differences in timing between conditions. Participants in the current study completed both conditions in one, 3-hr session, whereas participants in Scheerer, Tumber, and Jones completed the conditions 1 week apart. The test-retest reliability of  $f_0$  adaptation paradigms is questionable (Kapsner-Smith et al., 2024) and thus poses a risk for comparing results from two different time points. The time scale between sessions, combined with the practice of removing “following” responses, indicate that results from Scheerer, Tumber, and Jones may not be comparable to the current study. Another factor to consider regarding the results from Experiment 2 includes task difficulty. Relative to the more naturalistic sentence production task in Experiment 1, the task in Experiment 2 required participants to remember to sustain the color of the word while

also being exposed to altered auditory feedback. The difficulty of the task could have interfered with the cognitive processes being targeted by the Stroop design. However, since the purpose of this study was to investigate cognitive load broadly, the challenge of the task actually contributes to the cognitive load and would have amplified an effect if present.

### ***Relationship Between Vocal Production and Vocal Motor Control During Cognitive Load***

We found a significant negative correlation among the between-condition change in  $f_0$  during Experiment 1 and the between-condition change in  $f_0$  during the after-effect phase in the Experiment 2. In other words, with increased cognitive load, participants who decreased their  $f_0$  during running speech adapted *more* to altered auditory feedback. We did not find relationships among the between-condition changes in any other acoustic metric in Experiment 1 and the between-condition  $f_0$  changes in adaptation during Experiment 2. This suggests that  $f_0$  is uniquely sensitive to changes in cognitive load, such that these changes are comparable across tasks with different stimuli. To the authors’ knowledge, this is the first investigation to report a relationship between measures of  $f_0$  adaptation and  $f_0$  in running speech. We interpreted this result as a relationship between the upstream planning/control mechanisms and the downstream production of voice and during increased cognitive load. This finding supports our hypothesis and fits into theoretical frameworks: speakers whose ability to update their feedforward control was affected by increased cognitive load also experienced production-level changes during typical running speech, requiring adjustments to offset potential signal disruptions.

It is known that speech production requires planning and cognitive programming such that a relationship exists between the lexical planning of speech and the produced output of that sequence. The novel finding in the current study is that the motor planning involved in adaptation relates to the motor productions of unrelated speech tasks, as shown in the current study with different stimuli. This indicates some degree of dynamic interaction and flexibility between the feedforward control necessary for vocal motor planning and production. In feedforward models, stored commands are necessary for motor planning and their stability (or lack thereof) is an indication of the robustness of these models. These stored motor programs allow for the maintenance and delivery of speech despite circumstances that might degrade these processes, such as cognitively demanding situations. Thus, the relationship found in the current study suggests that the stability of one’s internal motor programming and planning mechanisms

can potentially be inferred from their ability to maintain their voice under stress.

### **Clinical Relevance**

Although the study motivation and aims were not developed to have a direct impact on clinical practice, the findings have potential clinical relevance. Voice quality and loudness are perceptual targets often used in voice therapy. In Experiment 1, we found a relationship between cognitive load and SPL and a relationship between autonomic arousal and CPPS. These findings suggest that these acoustic targets are potentially influenced by physiological responses and cognitive loading; thus, the context in which therapy tasks are implemented is important to consider. For example, in therapies with a hierarchical structure that gradually increase in difficulty (potentially inducing stress and cognitive demand), voice changes may not be linear. This could ultimately impact skill transfer from less cognitively demanding tasks to more complex ones. Regarding the findings from Experiment 2, the relationship between autonomic arousal and adaptation is relevant particularly for motor learning-based approaches. These findings suggest that an environment or task that induces autonomic arousal could potentially promote motor learning. It is likely the case that a threshold of autonomic arousal exists for optimal vocal motor learning and that performance would decrease if this threshold were crossed. Optimal physiological states are also likely person specific, and more research is warranted to understand how to harness this state to enhance therapeutic practices.

### **Future Directions/Limitations**

There are a number of limitations to consider when interpreting the results of the current study. First, the participants had limited time for their ANS to recover between conditions, and although the time allotted was similar to rest periods of other studies, which has been shown to be sufficient for ANS activity to return to a baseline state (Weber & Smith, 1990), it still must be considered. Another limitation of the current study is the sample of participants with typical speech and autonomic function. It is possible that relationships between the ANS, cognition, and vocal motor control are affected in clinical populations in which a dysfunction in one of the domains may elicit compensatory actions in another. As mentioned briefly in the discussion within the realm of task difficulty in Experiment 2, it is possible that other outside factors that were unrelated to the cognitive processes targeted by the Stroop design contributed to the results. However, the questions driving the study were focused on cognitive load in a broad sense. Since confounding variables are common in human behavioral

experiments, we did not limit our interpretation of the results to the cognitive processes specific to the Stroop effect. Instead, we framed the results in terms of overall cognitive load and discussed the accompanying physiological responses. Future studies should investigate these relationships in populations with autonomic dysfunction and atypical voice mechanisms. Additionally, this study used one type of cognitive task to elicit autonomic arousal. It would be interesting to determine how different cognitive tasks impact downstream physiological and vocal processes. Replicating this study but with the addition of multiple tasks that tax different aspects of cognition would allow for a deeper understanding of how cognition and autonomic arousal affect the voice. The wider breadth of data would also help in explaining some of the variability observed in the literature.

### **Conclusions**

The current study investigated the effects of autonomic arousal and cognitive load on voice production and vocal motor control in adults with typical speech. Results suggest that downstream voice production and upstream vocal motor control both change in the face of cognitive load and ANS arousal, but in different ways. Through the use of physiological measures, this study disassociated the effects of cognitive loading and autonomic arousal on the voice.

### **Data Availability Statement**

The data set for the current study is not publicly available due to the inability to de-identify participant voice recordings.

### **Acknowledgments**

This study was funded by the National Institute on Deafness and Other Communication Disorders Grants DC016270 (awarded to C.E.S. and F.H.G.) and DC015446 (awarded to R.E.H.), the Association for Computing Machinery's Special Interest Group of High Performance Computing Fellowship (N.E.T.), and the Acoustical Society of America's Raymond H. Stetson Scholarship in Phonetics and Speech Science (N.E.T.).

### **References**

Abur, D., MacPherson, M. K., Shembel, A. C., & Stepp, C. E. (2023). Acoustic measures of voice and physiologic measures

- of autonomic arousal during speech as a function of cognitive load in older adults. *Journal of Voice*, 37(2), 194–202. <https://doi.org/10.1016/j.jvoice.2020.12.027>
- Alexander, J. K., Hillier, A., Smith, R. M., Tivarus, M. E., & Beversdorf, D. Q. (2007). Beta-adrenergic modulation of cognitive flexibility during stress. *Journal of Cognitive Neuroscience*, 19(3), 468–478. <https://doi.org/10.1162/jocn.2007.19.3.468>
- Al-Yahya, E., Dawes, H., Smith, L., Dennis, A., Howells, K., & Cockburn, J. (2011). Cognitive motor interference while walking: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews*, 35(3), 715–728. <https://doi.org/10.1016/j.neubiorev.2010.08.008>
- American Speech-Language-Hearing Association. (2020). *Adult hearing screening* [Practice portal]. <https://www.asha.org/practice-portal/professional-issues/adult-hearing-screening/>
- Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, 10(6), 410–422. <https://doi.org/10.1038/nrn2648>
- Bayot, M., Dujardin, K., Tard, C., Defebvre, L., Bonnet, C. T., Allart, E., & Delval, A. (2018). The interaction between cognition and motor control: A theoretical framework for dual-task interference effects on posture, gait initiation, gait and turning. *Journal of Clinical Neurophysiology*, 48(6), 361–375. <https://doi.org/10.1016/j.neucli.2018.10.003>
- Bear, M., Connors, B., & Paradiso, M. A. (2020). *Neuroscience: Exploring the brain, enhanced edition*. Jones & Bartlett Learning
- Berg, M., Fuchs, M., Wirkner, K., Loeffler, M., Engel, C., & Berger, T. (2017). The speaking voice in the general population: Normative data and associations to sociodemographic and lifestyle factors. *Journal of Voice*, 31(2), 257.e13–257.e24. <https://doi.org/10.1016/j.jvoice.2016.06.001>
- Bhatt, T., & Kannan, L. (2022). Dual-task training for improving cognitive-motor interference. *Innovation in Aging*, 6(Suppl. 1), 394–394. <https://doi.org/10.1093/geroni/igac059.1551>
- Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* [Computer program]. <http://www.praat.org>
- Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4614-1126-0>
- Boyer, S., Paubel, P.-V., Ruiz, R., El Yagoubi, R., & Daurat, A. (2018). Human voice as a measure of mental load level. *Journal of Speech, Language, and Hearing Research*, 61(11), 2722–2734. [https://doi.org/10.1044/2018\\_JSLHR-S-18-0066](https://doi.org/10.1044/2018_JSLHR-S-18-0066)
- Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice  $f_0$  responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, 103(6), 3153–3161. <https://doi.org/10.1121/1.423073>
- Burnett, T. A., & Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America*, 112(3), 1058–1063. <https://doi.org/10.1121/1.1487844>
- Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). *Handbook of psychophysiology* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781107415782>
- Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F responses to pitch-shifted voice feedback during English speech. *The Journal of the Acoustical Society of America*, 121(2), 1157–1163. <https://doi.org/10.1121/1.2404624>
- Christodoulides, G. (2016). *Effects of cognitive load on speech production and perception*. UCL-Université Catholique de Louvain.
- Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology*, 493(1), 154–166. <https://doi.org/10.1002/cne.20749>
- Critchley, H. D., Elliott, R., Mathias, C. J., & Dolan, R. J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 20(8), 3033–3040. <https://doi.org/10.1523/JNEUROSCI.20-08-03033.2000>
- Crystal, D. (1987). Towards a “bucket” theory of language disability: Taking account of interaction between linguistic levels. *Clinical Linguistics & Phonetics*, 1(1), 7–22. <https://doi.org/10.1080/02699208708985001>
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Academic Press.
- Dahl, K. L., & Stepp, C. E. (2023). Effects of cognitive stress on voice acoustics in individuals with hyperfunctional voice disorders. *American Journal of Speech-Language Pathology*, 32(1), 264–274. [https://doi.org/10.1044/2022\\_AJSLP-22-00204](https://doi.org/10.1044/2022_AJSLP-22-00204)
- Dawson, M. E., Schell, A. M., & Fillion, D. L. (2017). The electrodermal system. In G. G. Berntson, J. T. Cacioppo, & L. G. Tassinary (Eds.), *Handbook of psychophysiology* (pp. 217–243). Cambridge University Press.
- Dromey, C., & Bates, E. (2005). Speech interactions with linguistic, cognitive, and visuomotor tasks. *Journal of Speech, Language, and Hearing Research*, 48(2), 295–305. [https://doi.org/10.1044/1092-4388\(2005\)020](https://doi.org/10.1044/1092-4388(2005)020)
- Dromey, C., & Benson, A. (2003). Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research*, 46(5), 1234–1246. [https://doi.org/10.1044/1092-4388\(2003\)096](https://doi.org/10.1044/1092-4388(2003)096)
- Dromey, C., & Shim, E. (2008). The effects of divided attention on speech motor, verbal fluency, and manual task performance. *Journal of Speech, Language, and Hearing Research*, 51(5), 1171–1182. [https://doi.org/10.1044/1092-4388\(2008\)06-0221](https://doi.org/10.1044/1092-4388(2008)06-0221)
- Gordan, R., Gwathmey, J. K., & Xie, L. H. (2015). Autonomic and endocrine control of cardiovascular function. *World Journal of Cardiology*, 7(4), 204–214. <https://doi.org/10.4330/wjc.v7.i4.204>
- Hamill, R. W., Shapiro, R. E., & Vizzard, M. A. (2012). Peripheral autonomic nervous system. In D. Robertson, I. Biaggioni, G. Burnstock, P. A. Low, & J. F. R. Paton (Eds.), *Primer on the autonomic nervous system* (pp. 17–26). Elsevier. <https://doi.org/10.1016/B978-0-12-386525-0.00004-4>
- Heaton, K. J., Williamson, J. R., Lammert, A. C., Finkelstein, K. R., Haven, C. C., Sturim, D., Smalt, C. J., & Quatieri, T. F. (2020). Predicting changes in performance due to cognitive fatigue: A multimodal approach based on speech motor coordination and electrodermal activity. *The Clinical Neuropsychologist*, 34(6), 1190–1214. <https://doi.org/10.1080/13854046.2020.1787522>
- Heller Murray, E. S., Lupiani, A. A., Kolin, K. R., Segina, R. K., & Stepp, C. E. (2019). Pitch shifting with the commercially available eventide eclipse: Intended and unintended changes to the speech signal. *Journal of Speech, Language, and Hearing Research*, 62(7), 2270–2279. [https://doi.org/10.1044/2019\\_JSLHR-S-18-0408](https://doi.org/10.1044/2019_JSLHR-S-18-0408)
- Helou, L. B., Jennings, J. R., Rosen, C. A., Wang, W., & Verdolini Abbott, K. (2020). Intrinsic laryngeal muscle response to a public speech preparation stressor: Personality and autonomic predictors. *Journal of Speech, Language, and Hearing Research*, 63(9), 2940–2951. [https://doi.org/10.1044/2020\\_JSLHR-19-00402](https://doi.org/10.1044/2020_JSLHR-19-00402)
- Helou, L. B., Wang, W., Ashmore, R. C., Rosen, C. A., & Verdolini Abbott, K. (2013). Intrinsic laryngeal muscle activity in response to autonomic nervous system activation. *Laryngoscope*, 123(11), 2756–2765. <https://doi.org/10.1002/lary.24109>
- Hilz, M. J., & Dütsch, M. (2006). Quantitative studies of autonomic function. *Muscle & Nerve*, 33(1), 6–20. <https://doi.org/10.1002/mus.20365>
- Hisa, Y., Bamba, H., Koike, S., Shogaki, K., Tadaki, N., & Uno, T. (1999). Neurotransmitters and neuromodulators involved in



- laryngeal innervation. *Annals Otolaryngology, Rhinology & Laryngology*, 108(Suppl. 7), 3–14. <https://doi.org/10.1177/00034894991080s702>
- Hisa, Y., Matsui, T., Fukui, K., Iyata, Y., & Mizukoshi, O.** (1982). Ultrastructural and fluorescence histochemical studies on the sympathetic innervation of the canine laryngeal glands. *Acta Oto-Laryngologica*, 93(1–6), 119–122. <https://doi.org/10.3109/00016488209130861>
- Hoshikawa, Y., & Yamamoto, Y.** (1997). Effects of Stroop color-word conflict test on the autonomic nervous system responses. *American Journal of Physiology-Heart and Circulatory Physiology*, 272(3), H1113–H1121. <https://doi.org/10.1152/ajpheart.1997.272.3.H1113>
- Hsu, N. S., Kuchinsky, S. E., & Novick, J. M.** (2021). Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience*, 36(2), 211–239. <https://doi.org/10.1080/23273798.2020.1836379>
- Huttunen, K., Keranen, H., Vayrynen, E., Paakkonen, R., & Leino, T.** (2011). Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. *Applied Ergonomics*, 42(2), 348–357. <https://doi.org/10.1016/j.apergo.2010.08.005>
- Ibanez, M., Valderrama-Canales, F. J., Maranillo, E., Vazquez, T., Pascual-Font, A., McHanwell, S., & Sanudo, J.** (2010). Human laryngeal ganglia contain both sympathetic and parasympathetic cell types. *Clinical Anatomy*, 23(6), 673–682. <https://doi.org/10.1002/ca.20956>
- Ishihara, S.** (1996). *Ishihara's tests for colour-deficiency*. Kanehara.
- Ito, M.** (2000). Mechanisms of motor learning in the cerebellum. *Brain Research*, 886(1–2), 237–245. [https://doi.org/10.1016/S0006-8993\(00\)03142-5](https://doi.org/10.1016/S0006-8993(00)03142-5)
- Janssen, S., Heijs, J., Van Der Meijs, W., Nonnekes, J., Bittner, M., Dorresteyn, L., Bloem, B. R., van Wezel, R. J. A., & Heida, T.** (2019). Validation of the Auditory Stroop Task to increase cognitive load in walking tasks in healthy elderly and persons with Parkinson's disease. *PLOS ONE*, 14(8), Article e0220735. <https://doi.org/10.1371/journal.pone.0220735>
- Jennings, J., Tahmouh, A., & Redmond, D.** (1980). Non-invasive measurement of peripheral vascular activity. In I. Martin & P. H. Venables (Eds.), *Techniques in psychophysiology* (pp. 69–136). Wiley.
- Jindrová, M., Kocourek, M., & Telenský, P.** (2020). *Skin conductance rise time and amplitude discern between different degrees of emotional arousal induced by affective pictures presented on a computer screen*. bioRxiv. <https://doi.org/10.1101/2020.05.12.090829>
- Jones, J. A., & Munhall, K. G.** (2000). Perceptual calibration of  $f_0$  production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America*, 108(3), 1246–1251. <https://doi.org/10.1121/1.1288414>
- Kapsner-Smith, M. R., Abur, D., Eadie, T. L., & Stepp, C. E.** (2024). Test-retest reliability of behavioral assays of feedforward and feedback auditory-motor control of voice and articulation. *Journal of Speech, Language, and Hearing Research*, 67(1), 34–48. [https://doi.org/10.1044/2023\\_JSLHR-23-00038](https://doi.org/10.1044/2023_JSLHR-23-00038)
- Kedor, C., Freitag, H., Meyer-Arndt, L., Wittke, K., Hanitsch, L. G., Zoller, T., Steinbeis, F., Haffke, M., Rudolf, G., Heidecker, B., Bobbert, T., Spranger, J., Volk, H.-D., Skurk, C., Konietzschke, F., Paul, F., Behrends, U., & Scheibenbogen, C.** (2022). A prospective observational study of post-COVID-19 chronic fatigue syndrome following the first pandemic wave in Germany and biomarkers associated with symptom severity. *Nature Communications*, 13(1), Article 5104. <https://doi.org/10.1038/s41467-022-32507-6>
- Keefe, K.** (1985). Motor and cognitive interference effects on unimanual tapping rates. *Brain and Cognition*, 4(2), 165–170. [https://doi.org/10.1016/0278-2626\(85\)90068-5](https://doi.org/10.1016/0278-2626(85)90068-5)
- Khawaja, M. A., Ruiz, N., & Chen, F.** (2007). Potential speech features for cognitive load measurement. *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces*, 57–60. <https://doi.org/10.1145/1324892.1324902>
- Kleinow, J., & Smith, A.** (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *Developmental Psychobiology*, 48(4), 275–287. <https://doi.org/10.1002/dev.20141>
- Klotzbier, T., & Schott, N.** (2017). Cognitive-motor interference during walking in older adults with probable mild cognitive impairment. *Frontiers in Aging Neuroscience*, 9, Article 350. <https://doi.org/10.3389/fnagi.2017.00350>
- Kriegel, Z.** (2022). *Influence of cognitive interference on speech* [Doctoral dissertation, Bowling Green State University]. OhioLINK Electronic Theses and Dissertations Center. [http://rave.ohiolink.edu/etdc/view?acc\\_num=bgsu1656092260918137](http://rave.ohiolink.edu/etdc/view?acc_num=bgsu1656092260918137)
- Kyle, B. N., & McNeil, D. W.** (2014). Autonomic arousal and experimentally induced pain: A critical review of the literature. *Pain Research and Management*, 19(3), 159–167. <https://doi.org/10.1155/2014/536859>
- Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., & Hain, T. C.** (2001). Comparison of voice  $f_0$  responses to pitch-shift onset and offset conditions. *The Journal of the Acoustical Society of America*, 110(6), 2845–2848. <https://doi.org/10.1121/1.1417527>
- Lester-Smith, R. A., Daliri, A., Enos, N., Abur, D., Lupiani, A. A., Letcher, S., & Stepp, C. E.** (2020). The relation of articulatory and vocal auditory-motor control in typical speakers. *Journal of Speech, Language, and Hearing Research*, 63(11), 3628–3642. [https://doi.org/10.1044/2020\\_JSLHR-20-00192](https://doi.org/10.1044/2020_JSLHR-20-00192)
- Lindblom, B.** (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (Vol. 55, pp. 403–439). Springer. [https://doi.org/10.1007/978-94-009-2037-8\\_16](https://doi.org/10.1007/978-94-009-2037-8_16)
- Liston, C., McEwen, B. S., & Casey, B. J.** (2009). Psychosocial stress reversibly disrupts prefrontal processing and attentional control. *Proceedings of the National Academy of Sciences of the United States of America*, 106(3), 912–917. <https://doi.org/10.1073/pnas.0807041106>
- Lively, S. E., Pisoni, D. B., Van Summers, W., & Bernacki, R. H.** (1993). Effects of cognitive workload on speech production: Acoustic analyses and perceptual consequences. *The Journal of the Acoustical Society of America*, 93(5), 2962–2973. <https://doi.org/10.1121/1.405815>
- Lovallo, W.** (1975). The cold pressor test and autonomic function: A review and integration. *Psychophysiology*, 12(3), 268–282. <https://doi.org/10.1111/j.1469-8986.1975.tb01289.x>
- Low, P. A.** (1993). Autonomic nervous system function. *Journal of Clinical Neurophysiology*, 10(1), 14–27. <https://doi.org/10.1097/00004691-199301000-00003>
- MacPherson, M. K., Abur, D., & Stepp, C. E.** (2017). Acoustic measures of voice and physiologic measures of autonomic arousal during speech as a function of cognitive load. *Journal of Voice*, 31(4), 504.e1–504.e9. <https://doi.org/10.1016/j.jvoice.2016.10.021>
- Mather, M., & Sutherland, M. R.** (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, 6(2), 114–133. <https://doi.org/10.1177/1745691611400234>
- Mendoza, E., & Carballo, G.** (1998). Acoustic analysis of induced vocal stress by means of cognitive workload tasks. *Journal of Voice*, 12(3), 263–273. [https://doi.org/10.1016/s0892-1997\(98\)80017-9](https://doi.org/10.1016/s0892-1997(98)80017-9)
- Middleton, F. A., & Strick, P. L.** (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive

- function. *Science*, 266(5184), 458–461. <https://doi.org/10.1126/science.7939688>
- Miller, H. E., Kearney, E., Nieto-Castañón, A., Falsini, R., Abur, D., Acosta, A., Chao, S.-C., Dahl, K. L., Franken, M., Heller Murray, E. S., Mollai, F., Niziolek, C. A., Parrell, B., Perrachione, T., Smith, D. J., Stepp, C. E., Tomassi, N., & Guenther, F. H. (2023). Do not cut off your tail: A mega-analysis of responses to auditory perturbation experiments. *Journal of Speech, Language, and Hearing Research*, 66(11), 4315–4331. [https://doi.org/10.1044/2023\\_JSLHR-23-00315](https://doi.org/10.1044/2023_JSLHR-23-00315)
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Patel, R., Niziolek, C., Reilly, K., & Guenther, F. H. (2011). Prosodic adaptations to pitch perturbation in running speech. *Journal of Speech, Language, and Hearing Research*, 54(4), 1051–1059. [https://doi.org/10.1044/1092-4388\(2010/10-0162\)](https://doi.org/10.1044/1092-4388(2010/10-0162))
- Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courcy, M., Deliyiski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology*, 27(3), 887–905. [https://doi.org/10.1044/2018\\_AJSLP-17-0009](https://doi.org/10.1044/2018_AJSLP-17-0009)
- Perrine, B. L., & Scherer, R. C. (2020). Aerodynamic and acoustic voice measures before and after an acute public speaking stressor. *Journal of Speech, Language, and Hearing Research*, 63(10), 3311–3325. [https://doi.org/10.1044/2020\\_JSLHR-19-00252](https://doi.org/10.1044/2020_JSLHR-19-00252)
- Ramaswamy, S., Shankar, S. K., Manjunath, K. Y., Devanathan, P. H., & Nityaseelan, N. (1994). Ultrastructure of the ganglion on human internal laryngeal nerve. *Neuroscience Research*, 18(4), 283–290. [https://doi.org/10.1016/0168-0102\(94\)90164-3](https://doi.org/10.1016/0168-0102(94)90164-3)
- Redding, G. M., & Wallace, B. (1985). Perceptual-motor coordination and adaptation during locomotion: Determinants of prism adaptation in hall exposure. *Perception & Psychophysics*, 38(4), 320–330. <https://doi.org/10.3758/bf03207161>
- Schäfer, C., Rosenblum, M. G., Abel, H. H., & Kurths, J. (1999). Synchronization in the human cardiorespiratory system. *Physical Review*, 60(1), 857–870. <https://doi.org/10.1103/physreve.60.857>
- Scheerer, N. E., Jacobson, D. S., & Jones, J. A. (2016). Sensorimotor learning in children and adults: Exposure to frequency-altered auditory feedback during speech production. *Neuroscience*, 314, 106–115. <https://doi.org/10.1016/j.neuroscience.2015.11.037>
- Scheerer, N. E., Tumber, A. K., & Jones, J. A. (2016). Attentional demands modulate sensorimotor learning induced by persistent exposure to changes in auditory feedback. *Journal of Neurophysiology*, 115(2), 826–832. <https://doi.org/10.1152/jn.00799.2015>
- Seidler, R. D., Kwak, Y., Fling, B. W., & Bernard, J. A. (2013). Neurocognitive mechanisms of error-based motor learning. In M. J. Richardson, M. A. Riley, & K. Shockley (Eds.), *Progress in motor control* (pp. 39–60). Springer. [https://doi.org/10.1007/978-1-4614-5465-6\\_3](https://doi.org/10.1007/978-1-4614-5465-6_3)
- Sletten, D. M., Suarez, G. A., Low, P. A., Mandrekar, J., & Singer, W. (2012). COMPASS 31: A refined and abbreviated composite autonomic symptom score. *Mayo Clinic Proceedings*, 87(12), 1196–1201. <https://doi.org/10.1016/j.mayocp.2012.10.013>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
- Taylor, E. W., Jordan, D., & Coote, J. H. (1999). Central control of the cardiovascular and respiratory systems and their interactions in vertebrates. *Physiological Reviews*, 79(3), 855–916. <https://doi.org/10.1152/physrev.1999.79.3.855>
- Taylor, J. A., & Thoroughman, K. A. (2007). Divided attention impairs human motor adaptation but not feedback control. *Journal of Neurophysiology*, 98(1), 317–326. <https://doi.org/10.1152/jn.01070.2006>
- Taylor, J. A., & Thoroughman, K. A. (2008). Motor adaptation scaled by the difficulty of a secondary cognitive task. *PLOS ONE*, 3(6), Article e2485. <https://doi.org/10.1371/journal.pone.0002485>
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981. <https://doi.org/10.1080/01690960903498424>
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3), 1429–1443. <https://doi.org/10.1016/j.neuroimage.2007.09.054>
- Van Lierde, K., Van Heule, S., De Ley, S., Mertens, E., & Claeys, S. (2009). Effect of psychological stress on female vocal quality. *Folia Phoniatrica et Logopaedica*, 61(2), 105–111. <https://doi.org/10.1159/000209273>
- Wass, S. V., Daubney, K., Golan, J., Logan, F., & Kushnerenko, E. (2019). Elevated physiological arousal is associated with larger but more variable neural responses to small acoustic change in children during a passive auditory attention task. *Developmental Cognitive Neuroscience*, 37, Article 100612. <https://doi.org/10.1016/j.dcn.2018.12.010>
- Watts, C. R., Awan, S. N., & Maryn, Y. (2017). A comparison of cepstral peak prominence measures from two acoustic analysis programs. *Journal of Voice*, 31(3), 387.e1–387.e10. <https://doi.org/10.1016/j.jvoice.2016.09.012>
- Weber, C. M., & Smith, A. (1990). Autonomic correlates of stuttering and speech assessed in a range of experimental tasks. *Journal of Speech, Language, and Hearing Research*, 33(4), 690–706. <https://doi.org/10.1044/jshr.3304.690>
- Weerathunge, H. R., Abur, D., Enos, N. M., Brown, K. M., & Stepp, C. E. (2020). Auditory-motor perturbations of voice fundamental frequency: Feedback delay and amplification. *Journal of Speech, Language, and Hearing Research*, 63(9), 2846–2860. [https://doi.org/10.1044/2020\\_JSLHR-19-00407](https://doi.org/10.1044/2020_JSLHR-19-00407)
- Witte, R. S., & Witte, J. S. (2017). *Statistics*. Wiley.
- Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, 60(3), 213–232. <https://doi.org/10.1037/h0044155>

## Appendix (p. 1 of 2)

## Sentence Task Stimuli

**Congruent:**

Lee flew red, yellow, and black kites at the spotted yard  
He tied blue, red and yellow knots on top of it.  
Oliver placed red, blue, and yellow keys to the right side

**Incongruent:**

Lee flew blue, black, and red kites at the spotted yard  
He tied yellow, black and blue knots on top of it.  
Oliver placed yellow, red, and black keys to the right side

**Motor Target:**

Lee flew red, yellow, and black kites at the spotted yard  
He tied blue, red, and yellow knots on top of it.  
Oliver placed red, blue, and yellow keys to the right side

**Assessing experimental fidelity**

To determine that the Stroop effect was increasing cognitive load, percent pause time (i.e., the amount of time spent pausing relative to the full amount of time spent on the task) was calculated and compared between conditions (using a paired  $t$  test) during Experiment 1. The amount of time spent pausing during speech has been used in the literature in various contexts to assess cognitive load (Khawaja et al., 2007). We found differences in the percent pause time (with a large effect size) in the two cognitive load conditions with a larger percent in the incongruent load. To ensure the cognitive load in both experiments was effectively arousing the ANS, multivariate ANOVAs (MANOVAs) were used to assess the effect of the cognitive load condition on the autonomic measures. The MANOVA model yielded a statistically significant effect of condition on the autonomic measures with a medium effect size in both experiments, confirming the cognitive load increased levels of ANS arousal at the group level. To the authors' knowledge, no prior studies have attempted to record physiological measures of the ANS during an altered auditory feedback experiment. The possibility of the altered auditory feedback itself eliciting additional ANS arousal was ruled out by comparing autonomic activity between the control and "shift-on" tasks in both cognitive load conditions (see below).

**Ruling out altered auditory feedback as a potential autonomic arouser**

Altered auditory feedback, when used properly, is typically meant to sound like one's own voice. Some participants are more sensitive to these perceived errors than others, hence, the variability in responses that is often observed with these paradigms. These alterations in auditory feedback, depending on the magnitude, may be surprising, unnatural, or even jarring to some participants and thus could activate the sympathetic nervous system. This potential increase in autonomic arousal, separate from the cognitive load conditions employed in the current study intended to modulate autonomic arousal, could pose as a confound for interpreting the results of this study. Thus, to ensure the autonomic arousal observed in our study was due to experimentally manipulating cognitive load, rather than altering auditory feedback, we compared autonomic measures between the control and "shift-on" tasks in both conditions.

We used paired  $t$  tests to assess the autonomic measures between the "shift-on" tasks and control tasks in both congruent and incongruent conditions. If altered auditory feedback was inducing autonomic arousal, we would expect a decrease in pulse volume amplitude and period, an increase in skin conductance response (SCR) amplitude and event responses, a decrease in SCR rise time, and increase in SCR half-recovery time in the shift tasks relative to the control tasks. We found no significant decreases in pulse volume amplitude, pulse period, SCR amplitude, SCR half-recovery time, and event-related SCRs between the control and shift task in either cognitive load condition. A significant decrease in SCR rise time was observed between the control and shift task in both conditions. In the congruent condition, there was a medium effect size ( $d = 0.61$ ), and in the incongruent condition, there was a small effect size ( $d = 0.4$ ). Note that we did not correct for multiple comparisons for this analysis and thus this result is interpreted with caution. Moreover, this measure is not used frequently in the literature as a proxy for autonomic arousal. Studies that have used SCR rise time have also included other more common metrics (i.e., SCR amplitude) and have used the rise time as more of a supplemental measure. One study has found that SCR rise time complimented amplitudes by differentiating between small and medium arousal levels (Jindrová et al., 2020). Changes in SCR rise time while the amplitude remains the same could indicate a small degree of arousal difference. We included cardiovascular measures for increased sensitivity and to account for small and quick changes in arousal. Although it is possible that the altered auditory feedback conditions elicited a small amount of arousal that was captured by the SCR rise time, it is unlikely due to the lack of differences in the cardiovascular measures as well as the rest of the SCR measures.

---

**Appendix** (p. 2 of 2)Sentence Task Stimuli

---

**Testing for mediation effects**

Autonomic arousal and cognition are bidirectionally related such that increases in cognitive load typically result in autonomic arousal, but also autonomic arousal can affect cognitive performance. Therefore, it is important when understanding the relationship between these factors and another variable of interest (such as voice outcomes) to account for the potential influence of the factors together. In the current study, we confirmed that cognitive load induced autonomic arousal through the MANOVA results described above. In some of our statistical models assessing the relationship between these factors and voice outcomes, we found a significant effect of autonomic arousal measures but not cognitive load and in another model, we found a significant effect of cognitive load but not autonomic arousal. To test whether one of these factors was mediating the relationship between the other factor and the outcome variable, we conducted mediation analyses in R software.

Mediation analyses were performed for all models in which either autonomic arousal (but not cognitive load) or cognitive load (but not autonomic arousal) was a significant factor on voice outcomes. Specifically, we tested two models from Experiment 1 (assessing CPPS and SPL) and one model from Experiment 2 (assessing  $f_0$  adaptation). In Experiment 1, autonomic arousal was related to CPPS and cognitive load was related to SPL, whereas in Experiment 2, autonomic arousal was related to  $f_0$  adaptation. For each model, the significant factor was treated as the mediator to assess whether an indirect mediation was occurring between the nonsignificant factor and voice outcome measure through the significant factor. For example, in the model assessing  $f_0$  adaptation in Experiment 2, we treated the significant factor of autonomic arousal (i.e., pulse volume amplitude) as the mediator to determine if a relationship between cognitive load and adaptation was being mediated by autonomic arousal.

We found no significant mediation effects in any of the tested models. One assumption that is typically met for a mediation effect is that both variables of interest relate to the outcome variable separately. To test this, we ran each model with and without the hypothesized “mediator” to determine if the other independent factor would be related to the outcome measure. However, in all of the models the nonsignificant factor remained nonsignificant when the hypothesized “mediator” was removed, indicating was unlikely to be a driving effect.

---