Attentional Demands Modulate Sensorimotor Learning Induced by Persistent Exposure to Changes in Auditory Feedback

Nichole E. Scheerer, Anupreet K. Tumber & Jeffery A. Jones

Psychology Department and Laurier Centre for Cognitive Neuroscience,
Wilfrid Laurier University
Waterloo, Ontario, Canada N2L 3C5

Running title: Attentional modulation of sensorimotor learning

*Corresponding author:
Jeffery A. Jones, Ph.D.
Psychology Department, Wilfrid Laurier University
Waterloo, Ontario, Canada N2L 3C5

Telephone: 1-519-884-0710 ext. 2992
E-mail: jjones@wlu.ca
Abstract

Hearing one’s own voice is important for regulating ongoing speech, and for mapping speech sounds onto articulator movements. However, it is currently unknown whether attention mediates changes in the relationship between motor commands and their acoustic output, which are necessary as growth and aging inevitably cause changes to the vocal tract. In this study, participants produced vocalizations while they heard their vocal pitch persistently shifted downward one semi-tone in both single- and dual-task conditions. During the single-task condition, participants vocalized while passively viewing a visual stream. During the dual-task condition, participants vocalized while also monitoring a visual stream for target letters, forcing participants to divide their attention. Participants’ vocal pitch was measured across each vocalization, to index the extent to which their ongoing vocalization was modified as a result of the deviant auditory feedback. Smaller compensatory responses were recorded during the dual-task condition, suggesting that divided attention interfered with the use of auditory feedback for the regulation of ongoing vocalizations. Participants’ vocal pitch was also measured at the beginning of each vocalization, before auditory feedback was available, to assess the extent to which the deviant auditory feedback was used to modify subsequent speech motor commands. Smaller changes in vocal pitch at vocalization onset were recorded during the dual-task condition, suggesting that divided attention diminished sensory-motor learning. Together the results of this study suggest that attention is required for the speech motor control system to make optimal use of auditory feedback for the regulation and planning of speech motor commands.

Keywords: Attention, Speech Motor Control, Sensorimotor Learning, Auditory Feedback
Introduction

Speech production is a motor skill that involves particularly complicated articulations. Auditory feedback plays a critical role in regulating ongoing speech and allows speech sounds to be mapped onto articulator movements during development (Guenther and Perkell 2004). The information provided by auditory feedback also allows the brain to maintain neural representations of sensory-motor relationships, despite inevitable physical changes that occur throughout one’s lifespan (e.g., growth, effects of aging). During speech, speakers must process auditory feedback while simultaneously processing information from other modalities. Since attention is a limited resource that must be divided amongst sensory input from different modalities (Posner and Snyder 1975), it is important to understand how attention modulates the integration of auditory and motor information to regulate ongoing and future speech motor commands.

The sensory-motor interactions that support fluent speech production are often investigated using a frequency-altered feedback (FAF) adaptation paradigm. For example, Houde and Jordan (1998) found that when participants heard predictable changes in the frequency of their vowel formants for an extended period of time they produced compensatory responses to the FAF. After-effects as a result of the exposure to the FAF also occurred, as the compensatory responses persisted following the removal of the feedback alteration. Compensatory responses to FAF highlight the importance of auditory feedback for maintaining fluent speech, while the persistence of the compensatory response following the removal of the FAF suggests that speech motor control is not solely regulated by auditory feedback. Rather, researchers posit that a feedforward controller plans and regulates speech motor commands, without relying on peripheral sensory feedback, which is delayed by neural processing (Houde et al. 2002; Guenther...
This feedforward controller utilizes sensory-motor representations that encode the relationship between speech motor commands and their sensory consequences. Once established, these sensory-motor representations drive speech production in a primarily feedforward manner. However, the after-effects witnessed after prolonged exposure to FAF demonstrate that sensory feedback is monitored and used to update the mapping of these sensory-motor representations, maintaining the precision of the feedforward controller. Although it is clear that fluent speech production relies on feedback and feedforward controllers, it is currently unclear how attention modulates these controllers, and their interactions.

Previous research has demonstrated that divided attention impairs performance on other goal-directed movements, such as arm-reaching (Taylor and Thoroughman 2007, 2008) and walking (Redding et al. 1985). Tumber and colleagues (2014) have also demonstrated that divided attention modulates feedback control of speech. However, it is unclear whether divided attention influences the sensorimotor interactions involved in the feedforward control of speech. For this reason, we exposed participants to FAF while they produced vowel sounds, with and without a secondary visual task. Since divided attention has been shown to impair sensory-motor integration during feedback control of speech (Tumber et al., 2014), as well as during the performance of other goal directed movements (Taylor and Thoroughman, 2007, 2008), we expected that asking participants to divide their attention between their auditory feedback and their visual input would diminish the compensatory responses and after-effects elicited by FAF.

Materials and Methods

Participants.
Thirty participants between the ages of 16 and 29 years ($M = 21.63$ years, $SD = 3.24$ years; 18 females and 12 males) participated. All participants were Canadian-English speakers who did not speak a tonal language, had normal or corrected to normal vision, had not been diagnosed with attention deficit (hyperactivity) disorder (ADD, ADHD), epilepsy (or had a family history of seizures), visual deficits that could not be amended by corrective lenses, and did not have a speech or language disorder. Vocal and behavioural responses were recorded from all 30 participants. However, since the primary objective of this study was to determine how attention modulates sensory-motor adaptation to FAF, 14 participants were excluded as they did not adapt to the FAF manipulation during the single-task condition, when no additional attentional load was applied. In total, 16 participants (9 females and 7 males) were retained for statistical analyses. All procedures were approved by the Wilfrid Laurier University Research Ethics Board and were in accordance with the World Medical Association 2013 Declaration of Helsinki.

**Apparatus.**

Participants were seated 76 cm from a 15-inch LCD monitor in a double-walled sound attenuated booth (Industrial Acoustics Company, Model 1601-01) and fitted with a headset-microphone (Sennheiser HMD 280-13 Pro, Sennheiser electronic GmbH & CO KG, Germany). One of two target notes was played to participants through headphones, and differed depending on their sex. These target notes were processed by combining three sine wave tones to create a harmonic tone using the program Praat (Boersma, 2001). The target notes were chosen to correspond to the normal vocal range of males and females. Thus, the target note heard by women was 261.4 Hz (C4) and the target note heard by men was 164.81 Hz (E3). Presentation of
the visual stimuli, notes, and shift onsets and offsets was controlled by Max/MSP 6 (Cycling '74, San Francisco, CA). Keyed behavioral responses to the questions posed to the participants during the dual task condition were also recorded using Max/MSP 6, using a standard keyboard with labeled keys. During the experiment, voice signals were sent to a mixer (Mackie Onyx 1200f, Loud Technologies, Woodinville, WA, USA), and then to a digital signal processor (VoiceOne, T.C. Helicon, Victoria, BC, Canada), which altered the fundamental frequency (F0) of the voice signal. This process introduced a small imperceptible (~10 ms) delay in the feedback signal, which was then presented back to the participant through headphones as auditory feedback. Loudness was not precisely monitored in this experiment, as previous research has shown that the relative loudness of auditory feedback has no influence on the amplitude of responses to FAF (Burnett et al., 1998). Despite this fact, the experimenter monitored the volume of the participants’ vocalizations and provided feedback to the participant if they were being too loud or too soft. The unaltered voice signal was digitally recorded (TASCAM HD-P2, Montebello, CA, USA) at a sampling rate of 44.1 kHz for later analysis.

Task. Participants completed a dual-task condition and a single-task condition. Each condition took place in a separate session, with the two sessions separated by approximately one week. For each session, participants produced 222 vocalizations of the vowel sound /a/ while exposed to FAF and a rapid serial visual presentation (RSVP) of letters.

Visual Stimuli.
The RSVP started with a 950 ms green fixation cross which was accompanied by the target note C4 for female participants, or the target note E3 for male participants. Following removal of the fixation cross and target note, a stream of successively presented letters appeared on the screen. Each stream of letters was approximately 5400 ms in duration, and consisted of 37 letters. Each letter appeared for 50 ms, followed by an interstimulus interval (ISI) of 100 ms. For both experimental sessions, each visual stream had two targets: 1. a randomly selected white letter from the alphabet (excluding the letter “X”); and 2. a black “X,” that occurred pseudorandomly before or after the white letter. The white letter appeared randomly between the 4th and the 37th letter, and the “X” randomly appeared between the 1st and 15th letter, before or after the white letter (minimum 2nd place, maximum 36th place in the visual stream). All of the letters in the stream were capitalized and in black font, with the exception of the white target letter. Each letter was displayed at the same location in the center of the grey screen, and during the ISI participants viewed a uniform grey field. For trials in the single-task condition, each visual stream ended with a blank grey field, whereas for trials in the dual-task condition, each visual stream ended with two questions. The first question appeared on the screen at the end of the visual stream and stated, “Identify the WHITE letter. If you are unsure, please guess.” The second question, which appeared immediately after the participant’s response to the first question stated, “Indicate when the “X” appeared with reference to the white letter.” The participant was required to press a key labelled “YES” if they believed that the “X” appeared before the white letter, and a key labelled “NO” if they believed that the “X” appeared after the white letter.

**Auditory Perturbations.**
For both the single- and dual-task conditions, the experimental session consisted of 19 unaltered (baseline) trials, followed by 180 FAF (shifted) trials, and 17 unaltered (test) trials. During the FAF trials, participants’ vocalizations were shifted downwards 100 cents from the start of the visual stream (and thus vocalization), and remained altered for the duration of their vocalization. Thus for the entire duration of the 180 FAF trials, participants only heard FAF. Participants auditory feedback was altered 100 cents downwards, as previous FAF adaptation studies have shown that persistent exposure to FAF of this magnitude and direction results in sensory-motor adaptation (Jones and Keough 2008, Keough et al. 2013). See Figure 1 for a depiction of the paradigm.

Experimental Procedure

The order of the single- and the dual-task conditions was counterbalanced across participants. During both conditions, participants were informed that at the start of each trial a fixation cross would appear on screen, and a target note would play in their headphones. Participants were instructed to start vocalizing an /a/ sound at a comfortable level once the target note was finished playing and the RSVP of letters began to flash on screen. Specifically the participants were instructed to start vocalizing when the first letter appeared on screen, and to stop vocalizing when the stream of letters stopped. In order to ensure the participants understood the task instructions, a series of practice vocalizations were performed at the start of each condition.

During the single-task condition, participants produced six practice vocalizations, followed by the experimental session of 198 vocalizations. During this condition, participants were required to passively view the RSVP of letters in order to know when to start and stop
vocalizing, however, their only task was to match the target note by producing a steady
vocalization. Participants progressed to subsequent trials with an inter-trial interval of 3000 ms.
The total duration of the single-task session was approximately 40 minutes.

During the dual-task condition, participants produced six practice vocalizations, followed
by the experimental session of 198 vocalizations. Participants were required to match the target
note by producing a steady vocalization, while they also attended to a RSVP of letters in order to
later answer questions about target letters in the visual stream. In this dual-task session, equal
emphasis was placed on matching the target note, and responding accurately to the two questions
at the end of each visual stream. Since participants could take as much time as they needed to
respond to the two questions, participants progressed to subsequent trials at their own pace. The
total duration of the experiment ranged from 40 to 60 minutes, depending on the amount of time
it took participants to respond to the questions during the dual-task.

_F0 analysis_

During each session, the participants’ unaltered voice recordings were segmented into
separate vocalizations. F0 values were calculated in Hertz for each vocalization using Praat
(Boersma 2001). A baseline F0 was then calculated for each participant by averaging the F0 of
the first 1500 ms of their last 5 unaltered vocalizations in the baseline period of the experimental
session. F0 values for each vocalization were then normalized to this baseline F0 by converting
Hertz values to cents using the formula:

$$\text{Cents} = 100 \times 12 \log_2 \left( \frac{F}{B} \right)$$

where F is the F0 value in Hertz, and B is the averaged baseline F0. Cents values were
calculated for the first 1500 ms of each vocalization.
In order to assess participants’ accuracy at matching the target note, F0 values were also normalized to the target note by substituting the frequency of the target note, 261.4 Hz (C4) for woman and 164.81 Hz (E3) for men, for B in this formula.

Results

Participants completed two experimental conditions in which auditory feedback was altered for an extended period of time, as participants vocalized the vowel sound /a/, while also simultaneously exposed to a RSVP of letters. In the single-task condition, participants passively viewed the RSVP of letters in order to know when to start and stop vocalizing. In the dual-task condition, participants actively monitored the RSVP of letters for cues to start and stop vocalizing, and also to identify two targets in the letter stream. Each condition took place in separate sessions where participants were exposed to 19 unaltered (baseline) trials, followed by 180 FAF (shifted) trials where feedback was held at -100 cents, and finally 17 unaltered (test) trials where feedback was returned to normal. These experimental trials were divided into four bins: baseline, shift-start, shift-end, and test, for subsequent analyses. The baseline phase consisted of the last 5 unaltered trials, the shift-start phase consisted of the first 5 FAF trials, the shift-end phase consisted of the last 5 FAF trials, and the test phase contained the first trial when feedback was returned to normal. The session-order was counterbalanced across participants, with an equal number of participants experiencing the single-task and dual-task conditions first.

Feedforward Control under Divided Attention

The median F0 value for the first 80 ms (the “median 80” value) of each vocalization was calculated to index participants’ F0 in the absence of auditory feedback. Previous research has
shown that compensatory responses to FAF have a latency of 100 to 150 ms (Burnett et al., 1998). This latency likely reflects the time required for the extraction of the F0 of the vocalization, detection of production errors, computation of the corrective command, and the delay between muscle activation and execution of the corrective movement(s) (Perkell et al., 1997). Thus, since corrective commands following changes to one’s auditory feedback are not generally observed for at least 100 ms, median 80 values can provide an index of feedforward control (Hawco and Jones, 2009; Keough and Jones, 2009; 2013). Averaged median 80 values were calculated for each experimental phase, for each participant, and then averaged across all participants for both the single- and the dual-task conditions.

In order to assess feedforward control across the experimental phases in both the single- and dual-task conditions, a repeated measures analysis of variance (RM-ANOVA) was conducted on the median 80 values with experimental phase (baseline, shift-start, shift-end, test), block order (single-task first, dual-task first), and condition (single-task, dual-task) as within subjects factors. A main effect of phase was observed \( (F(3,42) = 27.44, p < 0.001) \) as median 80 values were significantly larger in the test and shift-end phases, relative to the baseline and shift-start phases \( (p < 0.001) \). These results demonstrate that persistent exposure to the FAF caused participants to update their motor commands using information from the deviant auditory feedback. As a result, subsequent vocalizations were produced at a higher pitch, to compensate for the FAF alteration. The RM-ANOVA also revealed a significant condition by phase interaction \( (F(3,42) = 5.289, p = 0.021) \), as the single-task condition resulted in larger differences in median 80 values across the experimental phases, relative to the dual-task condition \( (p = 0.012; \text{see Figure 2}) \). This interaction demonstrates that during the single-task condition, the deviant auditory feedback had a larger influence on the way in which subsequent vocalizations
were produced. Since the deviant auditory feedback had less of an influence on subsequent vocalizations in the dual-task condition, these results suggest that the maintenance of feedforward control was impaired under divided attention. All other main effects and interactions were non-significant (\( p > .05 \)).

Feedback Control under Divided Attention

The median F0 value for the first 1500 ms (the “median 1500” value) of each vocalization was also calculated. Median 1500 values provide an index of changes in feedback control, since real-time sensory feedback information may be integrated into ongoing motor control as early as 100 ms. Averaged median 1500 values were calculated for each experimental phase, for each participant, and then averaged across all participants for the single- and the dual-task conditions.

A RM-ANOVA was conducted to assess the effect of condition (single-task, dual-task), block order (single-task first, dual-task first), and experimental phase (baseline, shift-start, shift-end, test) on the median 1500 cents values in order to understand how participants used their auditory feedback to alter their ongoing vocalizations. A significant main effect of experimental phase was found (\( F(3,42) = 15.944, p < 0.001 \)), as median 1500 values were larger during the test phase, relative to the baseline phase (\( p = 0.001 \)), and median 1500 values were also larger during the shift-end phase, relative to the test (\( p = 0.001 \)), shift-start (\( p = 0.007 \)), and baseline phases (\( p < 0.001 \)). The effect of phase on median 1500 cents values suggests that participants’ feedback driven compensatory responses were also affected by extended exposure to FAF.

Furthermore, the condition by phase interaction showed a trend towards significance (\( F(3,42) = 3.559, p = 0.074 \); see Figure 3), as median 1500 responses were significantly greater for the
single-task condition relative to the dual-task condition during shift-end ($p = 0.041$). All other main effects and interactions were non-significant ($p > .05$).

Responses to FAF under divided attention

The average median 1500 of the baseline phase, was subtracted from the average median 1500 of the shift-end phase, in order to index the magnitude of the participants’ compensatory responses after prolonged exposure to FAF. Similarly, the average median 80 of the baseline phase, was subtracted from the average median 80 of the shift-end phase, in order to index sensory-motor learning, or the extent to which the deviant auditory feedback was used to update the feedforward controller. After-effects were also calculated using median 80 values, by subtracting the mean of the baseline-phase from the mean of the test trials. After-effects were calculated in order to gauge the persistence of sensory-motor learning following the removal of the FAF manipulation.

To compare the effects of attentional load on compensatory responses, sensory-motor learning, and after-effects, multiple paired-samples t-tests were conducted. Significant differences were found between single- and dual-task condition for compensatory responses ($t(29) = 2.438, p = 0.028$), sensory-motor learning ($t(29) = 3.5, p = 0.003$), and after-effects ($t(29) = 2.169, p = 0.047$). For each comparison, median values were larger in the single-task condition, indicating that compensatory responses, sensory-motor learning, and after-effects were reduced when attention was divided.

The relationship between visual-task performance and compensations to FAF.
For each trial in the dual-task condition, records were kept of the white letter in the stream, whether the “X” appeared before or after the white letter, the participants’ keyed responses and reaction times to question one, as well as the participants’ keyed responses and reaction times to question two. Accuracy values for questions one and question two were averaged over the experimental block, and then across all participants for each question. Only accuracy data, not response time, was subjected to statistical analyses because response accuracy was emphasized in the instructions to the participants.

To assess the trade off between target identification accuracy, and the influence of deviant auditory feedback on vocal performance, Pearson-product moment correlations were conducted to assess the relationship between compensatory responses, sensory-motor learning, and after-effects for median 80 and median 1500 values, and the accuracy with which the white-letter and “X”-placement were identified during the dual-task session. All correlations yielded non-significant results ($p > 0.05$).

Vocal Accuracy under Divided Attention

Prior to each vocalization, participant’s heard a target note that they were asked to match. The F0 of participants’ vocalizations was normalized to the frequency of this target note in order to assess participants’ accuracy at matching the target note. Mean accuracy was then calculated by averaging the median 1500 of the last 5 unaltered vocalizations before the FAF manipulation was imposed. In order to investigate the influence of divided attention on accuracy, paired t-tests were conducted to compare accuracy in the single-task condition, to accuracy in the dual-task condition. The paired samples t-test revealed that accuracy did not differ across the single- and dual-task conditions, ($t(15) = 1.594, p > .05$).
Discussion

The aim of this experiment was to investigate whether divided attention modulates the integration of auditory and motor information during ongoing speech. Participants produced vocalizations while exposed to predictable changes in their auditory feedback. To manipulate attentional load, the vocalizations were produced while participants either passively viewed a RSVP of letters, or while they attended to a RSVP of letters in order to later identify target stimuli. Sensory-motor adaptation was assessed by measuring the participants’ F0 at the beginning of their vocalizations. Since auditory feedback is delayed by cortical processing, the F0 of the early portion of the participants’ vocalizations provided an index of feedforward control, or the extent to which the FAF was integrated into the planning of subsequent speech motor commands (Keough and Jones 2009, 2011, 2013; Hawco and Jones 2009). By contrast, participants’ median F0 after feedback became available was used to index the extent to which ongoing vocalizations were modified by FAF.

As expected, sensory-motor learning was demonstrated in both the single- and dual-task conditions, as median 80 values increased after persistent exposure to FAF. More specifically, median 80 values were larger in the shift-end and the test phases, relative to the baseline and the shift-start phases. These median 80 changes indicate that the way speech motor commands were executed by the feedforward controller changed as a result of exposure to FAF. An interaction between experimental phase and experimental condition was also found, which suggests that the influence of FAF on sensory-motor learning was modulated by the participants’ attention load. To elucidate the influence of divided attention on sensory-motor learning, averaged median 80 values extracted from the baseline phase were subtracted from averaged median 80 values
extracted from the shift-end phase, in both the single- and dual-task conditions. Comparing this
index of sensory-motor learning across the conditions revealed that more sensory-motor learning
occurred in the single-task condition. This result suggests that when participants were only
required to produce vocalizations, and thus their attention was undivided, participants used the
FAF to remap the sensory-motor representations utilized by their feedforward controller. As a
result of this remapping, participants’ subsequent vocalizations were produced at a higher F0,
which the participant would have perceived as closer to their baseline F0, as a result of the FAF.
On the other hand, less sensory-motor learning was observed when participants produced
vocalizations while also monitoring a visual stream for target letters. This indicates that under
divided attention, remapping of the sensory-motor representations that support the feedforward
controller was reduced, and the production of subsequent vocalizations was less affected by the
FAF. To further probe the influence of divided attention on sensory-motor integration, after-
effects, or the persistence of the compensatory response following the removal of the FAF
manipulation, were also calculated. These after-effects reflect the extent to which the motor
system was recalibrated as a result of the prolonged exposure to the FAF. Consistent with the
sensory-motor learning results, larger after-effects in the single-task condition were found, which
suggests the FAF had a larger effect on the feedforward controller when attention was undivided.
Together, these results suggest that sensory-motor learning decreased under divided attention,
which is consistent with the results of previous motor tasks performed under divided attention

By calculating the size of participants’ compensatory responses after feedback became
available, it was also possible to investigate the influence of divided attention on the feedback
controller. As expected, compensatory responses were elicited by FAF in both the single- and
dual-task conditions, as median 1500 values were larger during the shifted trials of the experiment. More specifically, the median 1500 of participants’ F0 was larger for vocalizations produced in the shift-end phase, relative to all other phases. This result indicates that the size of participants’ compensatory responses increased after repeated exposure to FAF, but then decreased toward baseline values when the FAF manipulation was removed and participants heard their unaltered voice. Note however, that the median 1500 of participants’ F0 was larger for vocalizations produced in the test phase, relative to the baseline phase, but smaller than their F0 in the shift-end phase. This result suggests that when participants’ auditory feedback was returned to normal their F0 decreased, but was still influenced by previous exposure to the FAF. This result demonstrates the interaction between the feedback and feedforward controllers. Remapping of the sensory-motor representations as a result of exposure to FAF caused the feedforward controller to produce vocalizations at a higher F0, however, once the no-longer-deviant auditory feedback was processed, the feedback controller reduced the F0 of the vocalization. To probe the influence of divided attention on these compensatory responses, the difference between the median 1500 of participants’ F0 at the end of the shifted phase, relative to the median 1500 of their F0 at the end of the baseline-phase, in both conditions was calculated. Larger compensatory responses were produced in the single-task condition, relative to the dual-task condition, which suggests that FAF had less of an influence on participants’ ongoing vocalizations while attention was divided. This finding is consistent with a previous study that demonstrated that reflexive responses to brief FAF perturbations are attenuated under divided attention (Tumber et al. 2013). Hu and colleagues recently showed that auditory event-related potentials elicited by FAF are likewise mediated by attention, though behavioural effects were
not observed in that study. Together these studies suggest that feedforward and feedback control of vocalizations are modulated by divided attention.

While the results of this study demonstrate that vocal responses to FAF can be influenced by divided attention, the results of previous arm-reaching (Taylor and Thoroughman 2007, 2008) and walking (Redding et al. 1985) studies have suggested that closed-loop monitoring of motor behaviours is generally unaffected by increases in attentional load. Although the modulation of motor responses to deviant sensory feedback by divided attention is a novel finding, divided attention has been shown to influence other types of auditory processing. Following increases in visual attentional load, the loudness of auditory tones has been shown to be attenuated by 7 dB (Dai 1991), which may be the result of reduced cochlear sensitivity (Delano et al. 2007), or decreased activation of the auditory cortex (c.f., Johnson and Zatorre 2005). Furthermore, focusing attention on an auditory channel has been shown to increase the loudness of the auditory input by 10 dB, relative to an unattended auditory channel (Choi et al. 2013). Together these studies suggest that divided attention may influence the saliency and the perceived loudness of auditory feedback. Based on the results of these studies, we suggest that performing the RSVP target identification task while vocalizing, reduced the saliency of the FAF, which resulted in smaller responses, and less sensory-motor learning.

Many theories have been proposed to explain how task performance is affected by divided attention. For example, the multiple resource theory of divided attention suggests that when simultaneously performing tasks, the degree to which performance on each task will decline is related to the similarity of the tasks, the attentional demands of each task, and how attentional resources are allocated between the tasks (Wickens 2007). Similarly, studies of cross-modal (e.g., visual and auditory) attention have suggested that separate, but linked attentional
systems exist (Spence and Driver 1996; Ferlazzo et al. 2002). According to this theory, simple
stimuli are processed by separate, modality specific resources, which eliminates any interference
that may occur as a result of simultaneously processing stimuli from different modalities
(Recanzone et al. 1993; Zenger-Landolt & Heeger 2003; Alais et al. 2006). However, when the
complexity of the stimuli increases, which in turn increases the demand for attentional resources,
attending to a stimulus in one modality may interfere with the processing of stimuli in different
modalities. To investigate how performance on the two tasks in this study may have been
influenced by simultaneously performing the two tasks, we investigated the relationship between
target letter identification accuracy, and vocal responses to FAF. The results of these correlations
indicated that changes in target identification accuracy were not related to the size of the vocal
responses to the FAF. Although we were unable to demonstrate a trade-off between sensory-
motor integration and RSVP target identification, target identification accuracy was high, thus
the RSVP task may not have been demanding enough to drive a target identification-auditory
processing response trade-off. Future studies probing the conditions under which divided
attention can impair sensory-motor integration may aid our understanding of how attention
modulates the integration of auditory and motor information to regulate ongoing and future
speech motor commands.

Speech motor control involves the precise coordination of vocal articulators, which are
controlled with the aid of auditory feedback (c.f., Waldstein 1990). The results of this study
suggest that dividing attention between the processing of auditory feedback and a secondary task
interferes with both online compensation to deviant auditory feedback, and sensory-motor
learning. Thus in order for the speech motor control system to make optimal use of auditory
feedback, to encode and transform auditory errors into corrective motor commands, attention is required.

Grants

This research was supported by a grant to Jeffery A. Jones from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Disclosure

The authors declare that they have no conflicts of interest.

References


Figure Legends

Figure 1. Experimental paradigm.

Participants observed a rapid serial visual presentation (RSVP) of letters with two targets (a white letter, and an “X”). Each RSVP trial started with a 9050 ms green fixation cross accompanied by the C4 target note for female participants, or the E3 target note for male participants. All trials were 5.4 s long where each letter in the stream appeared for 50 ms followed by an inter-stimulus interval of 100 ms. Participants completed a single-task condition and a dual-task condition. Each condition occurred in a separate session, taking place approximately one week apart. In the single-task condition, participants observed a blank screen for 3 s before the next trial. For the dual task condition, participants answered two questions at the end of each trial; they were asked to identify the white letter and whether the “X” appeared before or after the white letter. Participants vocalized the /a/ sound to match the target note during the letter stream in both conditions, while listening to their auditory feedback. For both conditions, the experimental block started with 19 unaltered baseline trials, followed by 180 shifted trials where feedback was held at -100 cents, then feedback was returned to normal for 17 test trials. Participants were expected to compensate to the shifted trials, and ultimately update their sensorimotor representations, showing aftereffects when feedback was returned to normal at test.
Figure 2. Compensation magnitudes early in utterances across experimental phases for the single-task condition and the dual-task condition.

The median F0 value for the first 80 ms of each utterance are normalized to baseline magnitudes and provide an index of feedforward control for each experimental condition. Participants produced larger compensation magnitudes at test during the single-task condition than the dual-task condition (see Results). Error bars represent standard error.

Figure 3. Compensation magnitudes late in utterances across experimental phases for the single-task condition and the dual-task condition.

The median F0 value for the first 1500 ms of each utterance are corrected to baseline magnitudes and provide an index of feedback control for each experimental condition. Participants produced larger compensation magnitudes overall at shift-end relative to test and shift-start, as well as larger compensation magnitudes at test relative to baseline (see Results). Error bars represent standard error.