EXPLORING THE INNER SPEECH PROCESS IN VERBAL WORKING MEMORY

by

Corrine Durisko

BS, Westminster College, 2002

Submitted to the Graduate Faculty of
Arts and Sciences in partial fulfillment
of the requirements for the degree of
Master of Science

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Corrine Durisko

It was defended on

September 8, 2006

and approved by

Carol Colby, Professor, Department of Neuroscience

Marc Sommer, Assistant Professor, Department of Neuroscience

Thesis Director: Julie Fiez, Professor, Departments of Psychology and Neuroscience

EXPLORING THE INNER SPEECH PROCESS IN VERBAL WORKING MEMORY

Corrine Durisko, M.S.

University of Pittsburgh, 2006

Verbal working memory (VWM) is the ability to dynamically preserve and manipulate verbal information for brief periods of time. VWM is maintained through a silent "inner speech" process (Baddeley, 1986; Baddeley & Hitch, 1974). It is well established in the behavioral and neuroimaging literature that VWM can be disrupted by the simultaneous (concurrent) performance of simple speech tasks (e.g. overt concurrent articulation of a word or digit) (Caplan *et al.*, 2000; Larsen & Baddeley, 2003). Our primary goal in these experiments is to test whether VWM and overt concurrent articulation will have one or more overlapping regions of activation in areas commonly associated with speech processing, and to determine whether such regions are active during simple tapping tasks. Due to concerns about overt movement artifacts, we also explore covert version of speech and tapping tasks. Experiment 1 was a behavioral study that examined the effects of overt and covert concurrent articulation and finger tapping on VWM. We found that overtly and covertly concurrently articulating "the" were the most detrimental to subjects' recall ability. These effects could be attributed to dual-task interference effects at the level of inner speech in VWM, thus, indicating a shared set of neural regions for all speech and VWM. At the same time, the effect sizes were different for the overt and covert versions of our tasks, raising questions about the common assumption of shared substrates. Experiment 2 was an imaging study designed to examine whether there were shared neural regions between simple speech tasks and VWM and to further explore differences between overt and covert tasks. The

results from this experiment provided only weak evidence implicating two candidate regions as the shared locus of activation: the left cerebellum and left superior temporal gyrus. We also found interesting evidence in support of distinct sets of regions for overt versus covert versions of the tasks.

TABLE OF CONTENTS

1.0 INTRODUCTION 1

2.0 EXPERIMENT 1 4

  2.1 INTRODUCTION 4

  2.2 MATERIALS AND METHODS 5

    2.2.1 SUBJECTS 5

    2.2.2 STIMULI AND PROCEDURE 5

  2.3 RESULTS 9

  2.4 DISCUSSION 11

3.0 EXPERIMENT 2 13

  3.1 INTRODUCTION 13

  3.2 MATERIALS AND METHODS 13

    3.2.1 SUBJECTS 13

    3.2.2 STIMULI AND PROCEDURE 14

  3.3 RESULTS 17

    3.3.1 BEHAVIORAL RESULTS 17

    3.3.2 NEUROIMAGING RESULTS 17

      3.3.2.1 TASK BY TIME INTERACTIONS 20

      3.3.2.2 GENERATION MODALITY BY TIME EFFECTS 21

3.3.2.3 TASK BY GENERATION MODALITY BY TIME EFFECTS.... 22

3.4 DISCUSSION... 26

4.0 CONCLUSIONS 32

APPENDIX A.. 34

APPENDIX B 38

BIBLIOGRAPHY 43

LIST OF TABLES

Table 1. Speech regions found from analysis of DSR data, mean signal change and SEM in each of the simple speech and motor tasks, and results of paired-sample t-tests between speech and tapping in both modalities. A trend toward a significant t-test result at $p \leq .05$ is listed in bold type... 19

Table 2. Name, coordinate, and Brodmann's area for the task x time interaction, percent signal change in the region for each task, paired sample t-test for differences between overt speech versus overt tap and covert speech versus covert tap, percent signal change in DSR task, and one sample t-test to determine significance in DSR task. A significant t-test result at $p \leq .05$ is represented in bold in the T-test Results column. Significant results on the one-sample t-tests for the DSR task are listed in bold in the last column. 22

Table 3. Regions found in the generation modality x time interaction and their representative patterns of activation: 1) covert positive and overt negative (Figure 3a), 2) overt positive and covert negative (Figure 3b), 3) covert negative and overt negative (Figure 3c), and 4) overt positive and covert no change (Figure 3d). This table also shows the percent signal change in the DSR task and if the region was significant in the DSR task from a one sample t-test. Significant one sample t-test results are marked by an asterisk. 24

LIST OF FIGURES

Figure 1. This is a schematic design of the basic DSR task that will be employed in these experiments. 7

Figure 2. Average VWM span of subjects with standard error bars on the DSR task with secondary conditions and the DSR task alone. 9

Figure 3. These four graphs show representatives of the patterns seen in the generation modality x time interaction. 3a. This is a representative graph of pattern 1 in the generation modality x time interaction. The covert task has a positive activation and the overt task has a negative activation. 3b. This is a representative graph of pattern 2 in the generation modality x time interaction. The overt task has a positive activation and the covert task has a negative activation. 3c. This is a representative graph of pattern 3 in the generation modality x time interaction. The covert task has a negative activation and the overt task has a negative activation. 3d. This is a representative graph of pattern 4 in the generation modality x time interaction. The overt task has a positive activation and the covert task has a no change in activation from baseline. 25

Figure 4.Two regions emerged throughout our analyses as possible candidates for the regions of significant overlap between VWM and overt simple speech. These two regions have only been weakly implicated. Regions are circled in yellow: 4a. This is the candidate region in the left cerebellum. 4b. This is the candidate region in the left superior temporal gyrus. 27

1.0 INTRODUCTION

Verbal working memory (VWM) is the ability to dynamically preserve and manipulate verbal information for brief periods of time. It has been argued that VWM is maintained through a silent "inner speech" process (Baddeley, 1986; Baddeley & Hitch, 1974). A commonly used example of VWM is when someone looks up a number in the phone book. After the individual finds the number, they repeat the seven digits over and over covertly until they can reach a phone to dial the number. This research focuses on the neural mechanisms that underlie inner speech, the maintenance process that supports VWM.

Evidence for the maintenance of verbal information via inner speech processes comes from two classic VWM effects demonstrated in the behavioral literature: the word length effect and the articulatory suppression effect. The word length effect occurs when subjects exhibit a decrease in VWM recall performance for words that take longer to articulate. For example, subjects typically have a lower VWM span for three-syllable as compared to one-syllable words (Baddeley *et al.*, 1984; Baddeley *et al.*, 1975). The word length effect is thought to occur because the inner speech process has a limited capacity (usually linked to decay of items in VWM) and items with longer articulation lengths pressure the boundaries of this system. The second effect, the articulatory suppression effect, occurs when the VWM span of subjects is decreased by the simultaneous (concurrent) performance of a simple speech task (e.g. overt concurrent articulation of a word, letter, or digit) (Murray, 1968). This phenomenon has been well studied in serial recall paradigms in which subjects are required to recall information in the order in which it was presented to them (Chein & Fiez, 2001).

While overt concurrent articulation causes decreases in the VWM span of subjects (Caplan et al., 2000; Larsen & Baddeley, 2003; Murray, 1968), overt finger tapping produces only slight decreases (Larsen & Baddeley, 2003; Saito, 1993, 1994, 1998). These results can be explained by the Multiple Resources Theory of dual-task interference, which argues that reductions in performance occur when two tasks tap into the same cognitive resource (Wickens, 1984). This theory suggests that performance deficits seen during concurrent articulation are the direct result of overt concurrent articulation "tying up" some of the same processes used for inner speech process, whereas overt tapping does not engage inner speech processes. To test this assertion we aim to replicate the well-known behavioral effects of overt concurrent articulation and overt finger tapping on a VWM tasks, and to extend the behavioral literature by identifying regions active during simple speech and motor tasks, and their overlap with regions active during VWM. We hypothesize that VWM and overt concurrent articulation will have one or more overlapping regions of activation in areas commonly associated with speech processing and these regions will not be active for overt finger tapping. Regions that demonstrate such a pattern could provide a neural account for the dual-task interference effects of concurrent articulation.

Historically, many neuroimaging researchers that were interested in overt tasks used covert versions instead because of concerns about image artifacts. Performance of overt speech tasks in the neuroimaging environment may cause artifacts due to physical movement and changes in magnetic susceptibility at boundaries between air and tissue (Barch *et al.*, 1999). Covert tasks have thus been employed to minimize these artifacts (Buckner *et al.*, 2000). Although using covert tasks in the neuroimaging environment is an intriguing alternative to overt tasks, there are also problems associated with covert task performance. For example, with covert tasks there is no way to monitor that subjects are performing the task as instructed, since there are no recorded measures such as performance rate and accuracy (Barch et al., 1999). Furthermore, the change from overt versions of the tasks to covert versions relies on the fundamental assumption that covert tasks are the same as overt tasks, without the end-stage motor processes (e.g., motor execution) (Barch et al., 1999). However, the cognitive psychological VWM literature highlights an important gap: to our knowledge, no behavioral data exists that directly contrasts overt and covert concurrent articulation effects. Instead, the behavioral literature focuses almost exclusively on overt tasks. In summary, there may be value in including covert speech and tapping tasks as a convergent measure of dual-task interference effects, but there is an absence of data to support assumptions of common mechanisms.

Our primary goal in these experiments is to identify the neural mechanisms that give rise to speech specific dual-task interference effects in VWM. Before conducting a neuroimaging study that contrasts speech and non-speech movement tasks, we will first conduct a behavioral study designed to test the assumption that overt and covert speech and motor tasks produce similar patterns of dual-task interference effects. If this is the case, the inclusion of both overt and covert versions of the simple speech and motor tasks in the subsequent neuroimaging study could provide a way to alleviate concerns about artifactual results associated with overt speech.

2.0 EXPERIMENT 1

2.1 INTRODUCTION

Our primary interest is to provide further evidence supporting a common neural substrate for VWM and overt concurrent articulation. Based on this goal there is a strong rationale to design an experiment looking at overt tasks. In spite of this, the trend in neuroimaging has been to utilize covert versions of tasks with the supposition that overt and covert are similar and differ only in the additional motor components involved in overt speech. Thus, as a secondary aim it is essential for us to compare both overt and covert versions of the speech tasks to ensure that the behavioral effects are the same. If the assumption of common mechanisms is correct, we should observe a decrease in VWM recall when both overt and covert concurrent articulation are performed in conjunction with a VWM, because both should overburden the same inner speech processes.

Additionally, we will look at subject performance on a VWM task with concurrent overt and covert finger tapping. To our knowledge, this contrast has not been examined in the behavioral literature. Behavioral studies have demonstrated modest effects of overt finger tapping on VWM, which have been attributed general attentional effects associated with almost any type of dual-task performance (Kemper *et al.*, 2003). However, these effects do not disrupt the memory ability of subjects at the same level as concurrent articulation. Thus, in this experiment we expect to see only slight decrements in VWM recall with overt finger tapping, and if covert and overt finger tapping engage the same mechanisms then covert tapping also should produce only slight performance decrements.

2.2 MATERIALS AND METHODS

2.2.1 SUBJECTS

Subjects were recruited from the University of Pittsburgh campus and surrounding metropolitan area through posted advertisements, e-mail notification of the studies, or through the Learning Research and Development Center (LRDC) online subject recruitment system (www.experimetrix.com/lrdc).

Nineteen subjects with no reported history of illiteracy, learning disabilities, psychiatric illness, or illicit drug use participated in the experiment. Prior to participation all subjects provided written informed consent on a University Of Pittsburgh Institutional Review Board approved form. All of the subjects were compensated $10.00 for their time and participation. The subjects ranged in age from 21–45 with a mean age of 27. There were 12 female and seven male subjects.

2.2.2 STIMULI AND PROCEDURE

The experiment employed a delayed serial recall task (DSR) (see Figure 1). Each trial of the DSR task consisted of three phases: 1) encoding, 2) maintenance, and 3) retrieval. During the six second encoding interval subjects were visually presented with a series of six letters at the center of a computer screen. Each letter appeared on the screen for one second until the next letter replaced it. Following the encoding interval there was a six second maintenance interval. During this interval all subjects were instructed to try and remember the letters presented during the encoding interval by repeating them covertly. Prior to the experiment all subjects read instructions and were again told not to use any mnemonic devices such as chunking or overt naming during the encoding or maintenance intervals to help them better remember the letters. After the maintenance interval, there was a six second retrieval interval. During this interval subjects were cued by red question marks that appeared at the center of the computer screen to recall the letters aloud in the order in which they were presented. This period, the retrieval interval, also lasted 6 seconds. The stimuli for this experiment consisted of eight non-rhyming letters (F, H, K, L, M, Q, R, S).

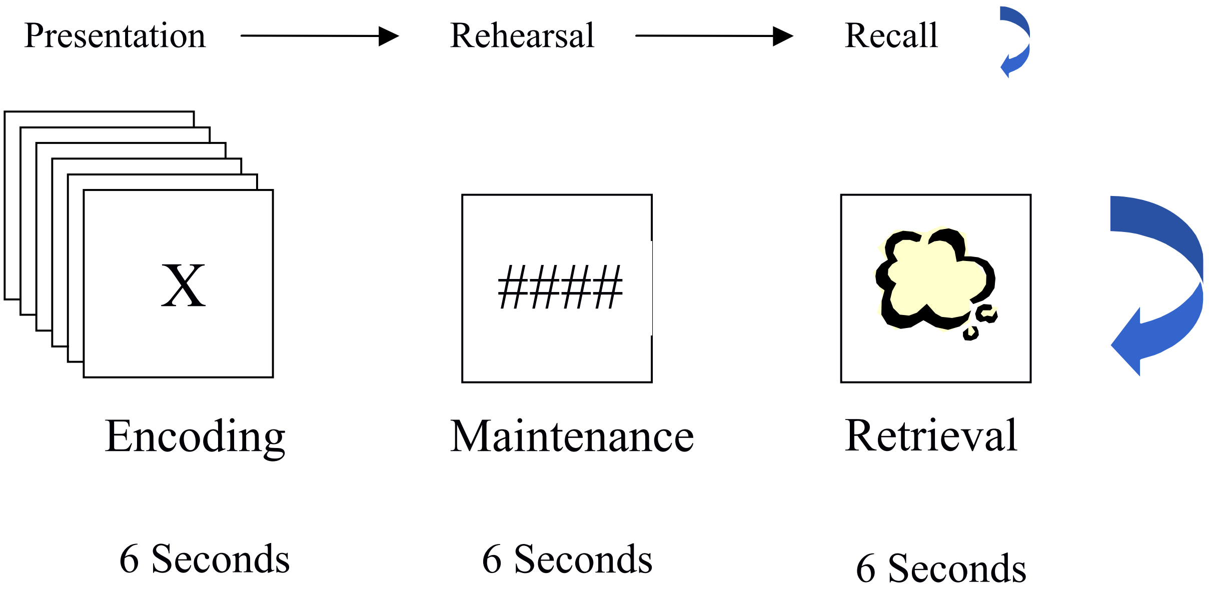


Figure 1. This is a schematic design of the basic DSR task that will be employed in these experiments.

In this study we manipulated the type of task performed concurrently with the DSR task. At the beginning of each trial subjects were presented with an instruction screen that provided them with information about whether the DSR would be performed alone or concurrently with a secondary task. Across trials, subjects were instructed to perform one of four types of secondary tasks: 1) <u>overt repetition</u> of “the” approximately three times per second; 2) <u>covert repetition</u> of “the” approximately 3 times per second; 3) <u>overt tapping</u> of the right index finger approximately three times per second; 4) <u>covert tapping</u> of the right index finger approximately three times per second. Subjects practiced the concurrent tasks with a metronome prior to the experiment so that they could learn the pace at which to perform. Their responses were recorded as they concurrently articulated and also while they provided their verbal responses during the retrieval interval using a microphone and Sony Sound Forge 8 audio software (Sony Corporation). The responses for the overt finger tapping were recorded using E-Prime (Psychology Software Tools, Inc.) software. Following the presentation of the instructions for each trial, subjects pressed a space bar to begin the encoding interval. In total, subjects performed 75 trials of the DSR task with 15 trials of each of the five conditions (DSR alone, DSR and overt concurrent articulation, DSR and overt finger tapping, DSR and covert concurrent articulation, and DSR and covert finger tapping) randomly intermixed. The total duration was approximately 45 minutes. Subjects were instructed to take breaks as they needed. A copy of the written instructions that were provided to the subjects for this experiment can be found in the Appendix.

Subject spans on the DSR task were determined by calculating the number of items recalled in the correct serial order and position on each trial, then averaging across trials to obtain a mean for each subject. To determine if the letter was correct the recorded letters were transcribed from the audio files and then compared to the letters presented to each subject from a trial log. Likewise, the rate of overt concurrent articulation was transcribed from the audio files. The total number of repetitions of the word “the” were tallied per trial and then divided by 12 (length of the encoding and maintenance intervals) to determine the average rate per second, and then these values were averaged across trials for each subject. For the overt tapping condition we obtained a text file that recorded the number of keypresses per trial. We divided the recorded number by 12 to determine the average rate per second, and then averaged over all of the trials for each subject.

2.3 RESULTS

First, we focused on subjects' rate of performance on the overt concurrent speech and overt finger tapping tasks. We found that concurrent articulation was performed at an average rate of 2.8 per second ± SEM = 0.1 and finger tapping was performed an average of 3.4 taps per second ± SEM = 0.2. Thus in terms of output rate subjects performed close to the rate that was practiced with the metronome prior to the experiment. We then performed a paired samples t-test to determine if the rate at which subjects overtly tapped differed from the rate at which they overtly articulated. In this analysis we had to exclude data from two of the subjects because there was a problem with data acquisition in the tapping condition. We found a significant difference between the rate at which these tasks were performed, t(1, 16) = 4.62, p<.001. Therefore, although the rates at which subjects' performed the overt versions of the secondary tasks differed only slightly from the practiced rate they were significantly different from one another.

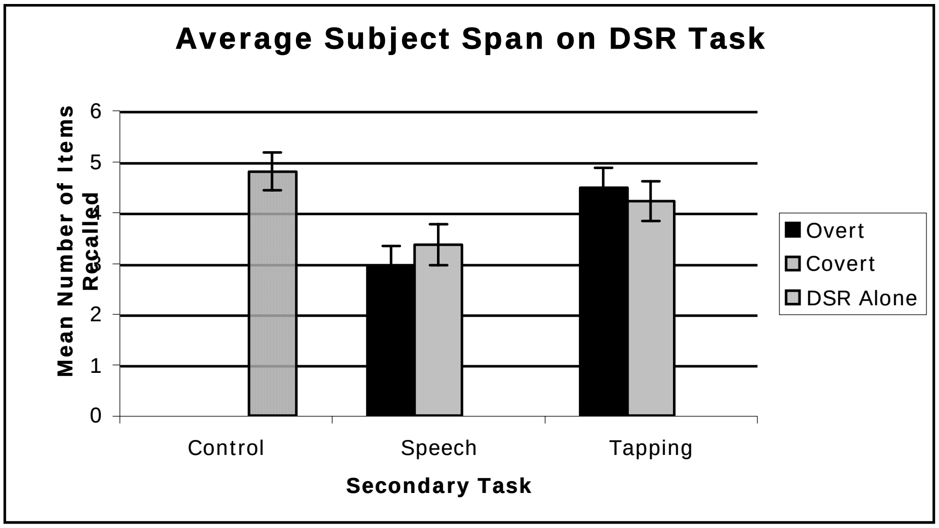


Figure 2. Average VWM span of subjects with standard error bars on the DSR task with secondary conditions and the DSR task alone.

In our initial analysis we examined the differences in subject performance for each secondary task condition versus the DSR task alone. Results of paired sample t-tests on the averaged subject span in each concurrent condition versus control indicated that all four secondary task conditions were significantly different from control (Figure 2), $t(1,18) \geq 3.082$, $p < 0.001$.

In order to further investigate the concurrent task conditions, a 2-way repeated measures ANOVA was conducted to contrast the effects of task type (speech or tap) and generation modality (overt or covert). There was a significant main effect of task type (speech < tap, $F(1,18) = 75.27$, $p < 0.001$) on memory performance. Post-hoc t-tests demonstrated that regardless of task generation modality (overt or covert) subjects performed significantly worse on the DSR task with concurrent performance of the speech tasks compared to concurrent performance of the tap tasks. That is, subjects' spans were significantly poorer with overt concurrent articulation than with overt finger tapping, ($t(18) = 8.04$, $p < .001$), and subjects' spans were significantly poorer with covert concurrent articulation than with covert finger tapping, ($t(18) = 5.30$, $p < .001$). There was no main effect of task generation (overt or covert), $F(1,18) = 0.38$, $p = .55$. However, there was a significant interaction between task type and task generation, $F(1,18) = 9.41$, $p = .007$. Post-hoc t-tests revealed that performance was relatively better for covert than overt speech ($t(1, 18) = 2.11$, $p=.05$) and relatively worse for covert tap than overt tap ($t(1,18) = 2.21$, $p=.04$).

2.4 DISCUSSION

The simple assumption that overt and covert versions of a task are the same, with additional motor processes, predicts that overt and covert concurrent articulation will have the same effects on subject VWM performance. In accord with this prediction both overt and covert concurrent articulation caused significant decreases in subjects' VWM recall ability that were greater than with concurrent finger tapping. This difference does not seem to reflect differences in motor output because subjects actually produced a faster rate of tapping than articulation. This finding provides a strong rationale for including both overt and covert tasks in the neuroimaging study planned for Experiment 2.

However, the data also suggest that some differences between overt and covert versions of the tasks may be found in Experiment 2. This is because we found that performing overt concurrent articulation in conjunction with the DSR was significantly more detrimental to subjects' overall span than performing a covert concurrent articulation task ($\underline{p} = .05$). Furthermore, based on the notion that overt and covert tasks are the same we anticipated that overt finger tapping and covert finger tapping would have similar effects on VWM. In contrast, the magnitude of these effects was different: subjects had a significantly poorer span in the covert finger tapping task compared to the overt version of the task ($\underline{p} = .04$).

Experiment 2 may be able to help elucidate the mechanisms that underlie the behavioral differences for the overt versus covert concurrent tasks. For instance, the greater decrease in VWM ability with covert finger tapping could be explained by subjects' resorting to verbal mediation in order to perform the finger tapping task covertly. If this is the case, we can expect there will be a set of regions active for the speech-mediated tasks (overt simple speech, covert simple speech, covert finger tapping) that overlap with VWM regions. Overt finger tapping should not show this overlap because the decreases in performance on this task are thought to be the result of the attentional demands associated with performing two tasks simultaneously (Kemper et al., 2003).

The fact that overt versus covert differences were obtained for both the speech and tapping tasks raises an alternative hypothesis: The overt and covert versions of these simple motor tasks may recruit different neural substrates. With this alternate view of the results we could predict a different set of results in Experiment 2. Particularly, there should be a set of convergent regions of activation associated with VWM that are engaged for both covert speech and tapping but not for overt speech and tapping. For instance, the difference in the dual-task interference between the overt and covert tasks could in part arise from differences between the regions involved in the motor production and perception aspects of generating movement in overt conditions versus the internal generation of movement in covert conditions (Mushiake *et al.*, 1991).

3.0 EXPERIMENT 2

3.1 INTRODUCTION

Our primary goal in this experiment is to test the hypothesis that VWM and overt concurrent articulation will have one or more overlapping regions of activation in areas commonly associated with speech processing and these regions will not be active for overt finger tapping. As a secondary goal we are interested in exploring the differences between overt and covert versions of the tasks in order to help identify differences that may account for the behavioral effects found in Experiment 1.

3.2 MATERIALS AND METHODS

3.2.1 SUBJECTS

Nineteen right-handed, native English speakers with no reported history of illiteracy, learning disabilities, psychiatric illness or illicit drug use participated in the experiment. Prior to participation all subjects provided written informed consent on a University of Pittsburgh Institutional Review Board approved form. After informed consent was obtained subjects completed an additional MR screening that excluded subjects who were pregnant, had metal or history of metal in any part of their bodies, medical or surgical complications, cardiac pacemaker, cochlear implant, aneurysm clip, IUD, shrapnel, weight over 300 pounds, or known problems of claustrophobia. Additionally, females of child-bearing potential were asked to complete a urine pregnancy test prior to participation. All subjects were compensated $60.00 for their time and participation in this research. The subjects ranged in age from 19–33 with a mean age of 23. Seven female and 12 male subjects were enrolled in this study.

3.2.2 STIMULI AND PROCEDURE

There were two parts to this experiment: a block of VWM trials using the DSR task and a block of simple motor tasks. The order of the blocks was counter-balanced across subjects. During the block of VWM trials subjects performed the 50 trials of the basic DSR task (Figure 1) described in Experiment 1. They did not perform any secondary tasks with the DSR task. We modified the parameters of the basic DSR task by adding a 14 second fixation cross at the center of the screen after the retrieval interval. This was added to allow the hemodynamic response signal to return to its baseline level. In addition, rather than being initiated by a subject keypress the encoding interval automatically started after the baseline fixation interval. The 50 trials of the DSR task were split up into five runs of 10 trials each. The total duration was 26 minutes and 40 seconds. A copy of the written instructions for this block of Experiment 2 can be found in the Appendix.

During the block of simple motor tasks subjects performed simple overt and covert speech and non-speech tasks. In particular, subjects executed the simple speech tasks by overtly or covertly repeated the word “the” approximately three times per second and the simple motor tasks by overtly or covertly tapping their right index finger approximately three times per second. Similar to Experiment 1, subjects practiced the rate of speech and tapping with a metronome prior to the experiment. These tasks were performed in six second epochs that were preceded by a two second instruction screen. Following each task epoch there was a 14 second baseline fixation with a fixation cross at the center of the screen. As with the VWM trials this served to allow the hemodynamic response to return to its baseline level. Each of the four types of task epochs occurred a 20 times total. The block was split up into five runs with each task epoch occurring randomly four times per run. The total duration was 29 minutes 20 seconds. A copy of the written instructions provided to each subject can be found in the Appendix.

Subjects' spoken responses on the DSR and overt speech task were recorded with a standard MR compatible microphone that was attached to a standard set of MR compatibles headphones (Avotec, Inc.). The microphone was affixed to the head coil to minimize the amount of cardiac and respiratory sounds recorded. The signal from the microphone ran to a PC and recorded with Sony Sound Forge 8 software (Sony Corporation). Prior to analysis, the audio was run through a noise cancellation program to eliminate a large portion of the background noise from the scanner (Jung *et al.*, 2005). Finally, an employee, blind to the experiment, transcribed the data in order to eliminate experimenter bias in the transcription.

Subjects' finger tapping responses were obtained from a MR compatible button glove placed on their right hand (Psychology Software Tools, Inc.). Each time the subject tapped their index finger the response was recorded in a text file using E-Prime software (Psychology Software Tools, Inc.). Subjects' span on the DSR task, average rate of speech, and average rate of tapping were calculated using the procedures described in Experiment 1.

All structural and functional scans were performed on a 3.0 Tesla (T) Siemens Magnetom Allegra head-only research scanner at the Brain Imaging Research Center (BIRC), a joint

University of Pittsburgh and Carnegie Mellon University facility. An FDA approved research head coil was used for all of the scans and certified MR technician was on hand at all times to monitor the subjects while they were in the scanner.

T2-weighted in-plane and high resolution structural scans were obtained using a standard EPI pulse sequence. Thirty-nine 3.5 mm thick oblique slices with no between slice gaps parallel to the ACPC were obtained during functional scans. Acquisition parameters used in this study to maximize the BOLD response were: TR=2000 ms, TE= 25ms, Flip Angle= 79, FOV=205 for 13.65 cm of coverage. These parameters gave us whole brain coverage in most of the subjects.

We performed analysis of the neuroimaging data using an integrative software package, Fiswidgets (Fissell *et al.*, 2003). After the raw data were reconstructed and checked for quality, an automated image registration procedure was used to correct for subject motion (Woods *et al.*, 1993). Any run where subjects moved more than four degrees or four mm was not included in the analysis. The images were then corrected for linear trends. A reference brain was chosen from among the subjects. After stripping the skull off of the subjects' brains, the structural scans were transformed into the reference space. The functional data were scaled to a global mean and a three dimensional Gaussian filter (8 mm FWHM) was used to smooth the data in order to account for between subject anatomical differences. Finally, the reference brain and functional maps were converted into Talairach space (Talairach & Tournoux, 1988) so that statistical analysis could be performed.

3.3 RESULTS

3.3.1 BEHAVIORAL RESULTS

In our initial analysis we examined the rate of performance on the overt simple speech and overt simple motor tasks. We found that overt speech was performed at an average rate of 2.6 per second ± SEM = 0.2 and finger tapping was performed an average of 3.7 taps per second ± SEM = 0.3. We performed a paired sample t-test to determine if the rate of overt speech differed from the rate of overt tapping. In this analysis we had to exclude data from three of the subjects because there were a problem with data acquisition in the speech and tapping conditions. There was a significant difference between the rates at which subjects performed the overt speech and overt tapping tasks, t(1, 15)=-4.29, p=.0001. The finger tapping task was performed significantly more quickly.

We also looked at subjects' performance on the DSR task and found an average span of 4.7 letters per trial +/- SEM = .2 letters.

3.3.2 NEUROIMAGING RESULTS

We had an *a priori* interest in brain regions associated with inner speech and clear set of predicted results from our primary question. Thus, as our first analysis we used a region-of-interest approach to compare activation across the DSR and the movement tasks. To begin, the DSR task was used as a localizer task to identify regions associated with inner speech. To do this we performed a voxel-wise ANOVA in which BOLD signal during the last four seconds of the maintenance interval was contrasted with the BOLD signal during the last four seconds of the baseline fixation interval. This analysis should reveal regions associated with the maintenance of information in working memory that that should persist even after the initial encoding period has ended. From the set of activated regions, we selected all of those that fell within brain areas previously implicated in speech processing and execution: 1) the medial frontal gyrus, 2) the precentral gyrus, 3) the superior temporal gyrus, 5) the anterior insula, and the cerebellum (for review see (Ackermann & Riecker, 2004; Ackermann *et al.*, 1998; Awh *et al.*, 1996; Chein *et al.*, 2002; Hickok, 2001; Hickok & Poeppel, 2000; Ravizza *et al.*, 2004)) A total of eight regions, largely in the left hemisphere, were thereby isolated for further study.

The set of eight regions were then applied to the data from the movement tasks to conduct a region of interest (ROI) based analysis of the movement data. In this next step, the average activation at each time point within each of the four different types of movement epochs was computed for each subject. From this "time series" we determined the average signal change for each condition in each subject. The average signal change was calculated by taking the average signal of the last four seconds of the maintenance interval minus the average signal of the last four seconds of the baseline interval, dividing the result from the average signal of the last four seconds of the maintenance interval, and then multiplying the resulting number by 100.

After we generated the percent signal change in each of the conditions we wanted to determine if there were significant differences between the speech and tapping tasks in each modality. We ran paired sample t-tests for overt speech versus overt tapping and covert speech versus overt tapping for each of the eight ROIs, using an alpha value of $p<.006$ to correct for multiple comparisons. The results are shown in Table 1. There were no significant differences between the speech and tap conditions. One region, in the left cerebellum, exhibited a tendency to produce a greater response in the speech than the tap condition.

Table 1. Speech regions found from analysis of DSR data, mean signal change and SEM in each of the simple speech and motor tasks, and results of paired-sample t-tests between speech and tapping in both modalities. A trend toward a significant t-test result at $p \leq .05$ is listed in bold type.

Name	Talairach Coordinate	Overt Speech	Overt Tap	T-test Results	Covert Speech	Covert Tap	T-test Results
Left Medial Frontal Gyrus (BA 6)	-1, 1, 50	.52 +/- .11	.47 +/- .09	p=.51	.23 +/- .04	.34 +/- .06	p=**.01**
Left Precentral Gyrus (BA 39)	-47, -6, 37	-.11 +/- .06	-11 +/- .04	p=.79	.14 +/- .04	.15 +/- .05	p=.87
Left Superior Temporal Gyrus (BA 22/39)	-54, -56, 19	-.27 +/- .13	-.23 +/- .10	p=.99	-.07 +/- .07	-.04 +/- .03	p=.63
Left Temporal Gyrus (BA 22/42)	-38, -31, 10	.26 +/- .08	.12 +/- .05	p=.26	-.13 +/- .05	-.08 +/- .03	p=.45
Right Insula	36, 14, 10	.20 +/- .06	.30 +/- .05	p=**.07**	.15 +/- .03	.20 +/- .05	p=.39
Right Cerebellum	38, -59, 30	.19 +/- .07	.25 +/- .06	p=.58	.10 +/- .05	.13 +/- .05	p=.46
Right Cerebellum	3, -71, 38	.20 +/- .09	.17 +/- .06	p=.57	.12 +/- .10	.03 +/- .05	p=.40
Left Cerebellum	-9, -72, -38	.23 +/- .09	.09 +/- .06	p=**.09**	.04 +/- .05	.01 +/- .08	p=.67

As a secondary analysis we used an exploratory voxel-wise approach to further probe for differences between speech and tapping tasks. For this analysis we conducted a 3-way, voxel-wise repeated measures ANOVA across the whole brain, with task type (speech vs. tongue), task generation modality (overt vs. covert), and time (four seconds at the beginning of baseline fixation vs. the last four seconds of baseline fixation) as factors. The factor of time (four seconds at the beginning of baseline fixation vs. the last four seconds of baseline fixation) should capture regions that exhibit task-related activity. The first four seconds of the fixation interval were used because we did not want to look at the actual movement epoch, since this could introduce movement artifacts into the data. The hemodynamic response is thought to peak six-nine seconds after a task ends. So, by using the first four seconds of the baseline interval following each task epoch we should still be capturing the peak hemodynamic response for each task condition. In the sections below, we report the results for the three contrasts of interest: task x time, generation modality x time, and task x generation modality x time. The data were thresholded using a p-value of $p \leq .0001$ and a voxel contiguity threshold of eight (cf (Tricomi *et al.*, 2004)).

3.3.2.1 TASK BY TIME INTERACTIONS

Regions that show a task x time interaction should show a speech versus tap difference between the task epoch and baseline fixation epoch. We found eight regions that were significantly active for this interaction (Table 2). After determining the regions, we computed the percent signal change for each condition using the procedures described in our primary analysis. We then probed for task effects using a paired samples t-tests on the percent signal change to separately examine the overt and covert conditions (overt speech vs. overt tap and covert speech vs. covert tap) to verify that one output modality (overt vs. covert) was not driving the task x time interaction. The results are listed in Table 2. We found a significant difference between overt speech and overt tap in all eight ROIs, with a correction for multiple comparisons, $\underline{p}<.006$. This is to be expected, since these regions were selected on the basis of a voxel-wise difference between the speech and tap tasks. Surprisingly, none of the differences between covert speech and covert tap survived the correction for multiple comparisons. However, the right postcentral gyrus and left cerebellum regions showed a trend towards significance.

Next, we applied the set of eight task x time regions to the data from the DSR task. The goal was to determine whether any of these regions identified in the movement condition were also active during the DSR task. For each region, the time series in the DSR condition was extracted and the average signal change was computed by comparing the last four seconds of the maintenance interval minus the last four seconds of the baseline interval. For each region, we conducted a one-sample t-test to determine if the activation significantly differed from zero. The results are listed in Table 2. We found three regions in which the DSR task was significantly different from zero: the right postcentral gyrus, left superior temporal gyrus, and left cerebellum. Thus, from these data three regions are the best candidate sites for dual-task interference effects between speech and VWM.

3.3.2.2 GENERATION MODALITY BY TIME EFFECTS

Our results from Experiment 1 showed overt and covert tasks disrupt performance differently. We hypothesized from these results that there may be differences between these two modalities that go beyond the additional motor output processes involved in motor execution. If there are in fact differences between these two modalities we expect to see them in the generation modality x time analysis. In general, this interaction should reveal regions that have a difference between overt and covert versions of the tasks. We identified 16 regions of significant activation, which had four distinct patterns of activation: 1) covert positive and overt negative (Figure 3a), 2) overt positive and covert negative (Figure 3b), 3) covert negative and overt negative (Figure 3c), and 4) overt positive and covert no change (Figure 3d). The regions and their pattern of activation can be seen in Table 3.

We also applied this set of 16 regions to the data from the DSR task to conduct a region of interest (ROI) based analysis of the DSR data, using the same procedures described for the analysis of the task x time regions. We found a significant difference from zero in the DSR task for eight of the 16 regions identified in the generation modality x time interaction. For two of these eight regions covert movement tasks produced a positive response, and overt produced a negative. These two regions would thus be the best candidate sites for dual-task interference effects for covert and not overt.

Table 2. Name, coordinate, and Brodmann's area for the task x time interaction, percent signal change in the region for each task, paired sample t-test for differences between overt speech versus overt tap and covert speech versus covert tap, percent signal change in DSR task, and one sample t-test to determine significance in DSR task. A significant t-test result at p ≤ .05 is represented in bold in the T-test Results column. Significant results on the one-sample t-tests for the DSR task are listed in bold in the last column.

Name	Talairach Coordinates	Overt Speech	Overt Tap	Paired Sample T-test	Covert Speech	Covert Tap	Paired Sample T-Test	DSR Task
Left Inferior Parietal (BA 2/40)	-41,-36,-53	-.21 +/- .05	.65 +/- .10	p≤.006	-.05+/.05	-.01 +/ - .03	p=.48	.07 +/- .05
Right Postcentral Gyrus (BA 3/1)	51,-12,43	1.02 +/-.12	-.09 +/-.03	p≤.006	.10+/-.05	-.05 +/- .04	**p=.04**	**.28 +/- .07**
Left Superior Temporal Gyrus (BA 22)	-63,-20,3	.10 +/- .10	-.13 +/- .04	p≤.006	-.03+/-.03	-.10 +/- .03	p=.14	**.13 +/- .05**
Right Superior Temporal Gyrus (BA 22)	52,-12,3	.81 +/- .07	-.15 +/- .02	p≤.006	.08 +/- .03	-.11 +/- .04	p=.46	-.02 +/- .03
Right Globus Pallidus	26,-14,-8	.29 +/- .05	-.01 +/- .03	p≤.006	.03 +/- .03	.01 +/- .03	p=.32	.01 +/- .03
Left Cerebellum	-25,-33,-26	-.43 +/- .07	-.16 +/- .03	p≤.006	.17 +/- .03	-.09 +/- .03	p=.99	-.04 +/- .04
Right Cerebellum	22,-46,-34	-.07 +/- .05	.50 +/- .06	p≤.006	-.05 +/-.02	.01 +/- .03	p=.15	.04 +/- .03
Left Cerebellum	-17,-59,-28	.47 +/- .06	.03 +/- .03	p≤.006	.01 +/- .02	-.09 +/- .03	**p=.01**	**.09 +/- .03**

3.3.2.3 TASK BY GENERATION MODALITY BY TIME EFFECTS

In Experiment 1 we found that overt and covert speech significantly decreased the VWM span of subjects. Interestingly, we also found that covert tapping causes a greater decrease in performance than overt tapping. From these results we posited that subjects may engage in verbal mediation during covert tapping, which could explain the greater decrease in performance with covert tapping. If such a speech region exists, we would expect it to show up in a three-way interaction with overt speech, covert speech and covert tapping producing a positive response, but overt tapping producing little change. We found a set of seven regions that exhibited a task x generation modality x time interaction: left inferior parietal (-34,-32,39), right precentral gyrus (51,-12,30), three regions in superior temporal gyrus (44,-23,-4, –60, -8, 4, and –47,-19,-1), and two regions in the cerebellum (-39,-56,50). None of these regions showed a pattern consistent with our hypothesis from Experiment 1 that there may be a region active for overt speech, covert speech, and covert tapping, but not for overt tapping.

Table 3. Regions found in the generation modality x time interaction and their representative patterns of activation: 1) covert positive and overt negative (Figure 3a), 2) overt positive and covert negative (Figure 3b), 3) covert negative and overt negative (Figure 3c), and 4) overt positive and covert no change (Figure 3d). This table also shows the percent signal change in the DSR task and if the region was significant in the DSR task from a one sample t-test. Significant one sample t-test results are marked by an asterisk.

Name	Talairach Coordinate	Pattern of Activation	DSR Task	One Sample T-test
Left Middle Frontal Gyrus (BA 6)	-25, -2, 57	1	.16 +/- .04	*
Right Precentral Gyrus (BA 4)	29, -14, 46	1	.08 +/- .05	
Left Middle Frontal Gyrus (BA 9)	-41, 15, 30	1	.20 +/- .08	*
Left Inferior Frontal Gyrus (BA 46)	-41, 43, 3	1	.00 +/- .04	
Right Insula	44, -18, -1	2	-.02 +/- .03	
Left Insula	-47, -17, 11	2	-.13 +/- .03	*
Left Precuneus (BA 7)	-9, -63, 43	3	-.05 +/- .02	
Right Anterior Cingulate (BA 10/47)	7, 46, -1	3	-.50 +/- .30	
Left Anterior Cingulate (BA 47)	-7, 36, -1	3	-.52 +/- .29	
Left Fusiform Gyrus	-33, -43, -16	3	-.10 +/- .03	*
Left Cerebellum (BA 36)	-25, -26, -26	3	-.21 +/- .10	
Right Parahippocampal Gyrus	26, -17, -26	3	-.23 +/- .10	*
Right Fusiform Gyrus	29, -39, -18	3	.09 +/- .04	*
Left Thalamus	-14, -19, -1	4	-.01 +/- .02	
Right Cerebellum	22, -56, -33	4	.13 +/- .03	*
Left Cerebellum	-16, -60, -27	4	.09 +/- .03	*

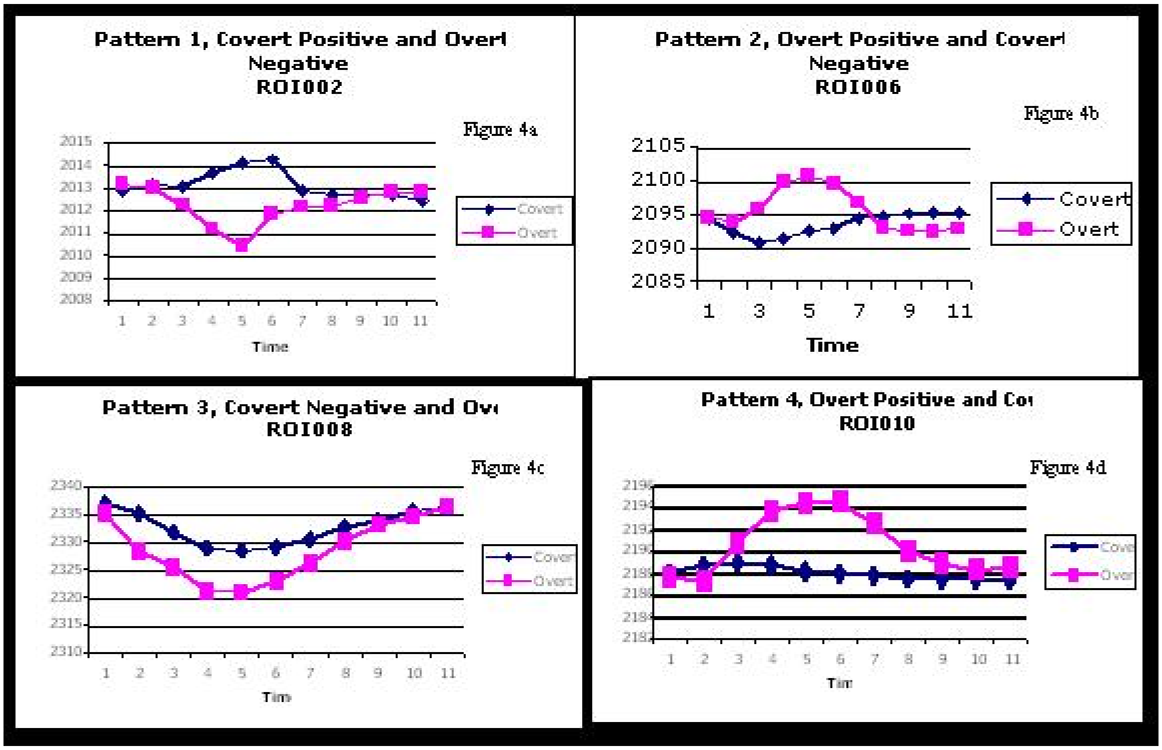


Figure 3. These four graphs show representatives of the patterns seen in the generation modality x time interaction. 3a. This is a representative graph of pattern 1 in the generation modality x time interaction. The covert task has a positive activation and the overt task has a negative activation. 3b. This is a representative graph of pattern 2 in the generation modality x time interaction. The overt task has a positive activation and the covert task has a negative activation. 3c. This is a representative graph of pattern 3 in the generation modality x time interaction. The covert task has a negative activation and the overt task has a negative activation. 3d. This is a representative graph of pattern 4 in the generation modality x time interaction. The overt task has a positive activation and the covert task has a no change in activation from baseline.

3.4 DISCUSSION

Our primary hypothesis was that VWM and overt concurrent articulation will have one or more overlapping regions of activation in areas commonly associated with speech processing, and these regions will not be active for overt finger tapping. Our analyses did not confirm this hypothesis. The lack of strong evidence in support of our primary hypothesis does not mean that it should be completely abandoned, for several reasons. First, we did find weak results in two regions, left cerebellum (Figure 4a) and left superior temporal gyrus (Figure 4b) that are consistent with our hypothesis. In the contrast between overt speech and overt tapping in our initial ROI analysis, the left cerebellum (-9, -72, -38) showed a trend towards significance (p=.09). This same region of the cerebellum (-17, -72, -38) was also active in the task x time analysis, where we found a significant difference between speech and tap in the overt condition ($\underline{p}<.006$), and a trend towards significance in the covert condition ($\underline{p}=.01$). When we performed an ROI analysis of the regions generated by the task x time interaction we found converging evidence that this region is involved in the DSR task. In particular, we saw that there was an increase in activation in the encoding, maintenance, and retrieval epochs. While, this region has been implicated in other studies of speech processing (Ackermann et al., 1998; Awh et al., 1996; Desmond *et al.*, 1997; Fiez *et al.*, 1996) the results from this study must be treated cautiously because we failed to find a significant effect in our *a priori* ROI-based analysis and this regions is primarily driven by overt speech.

The left superior temporal gyrus emerged as another weak candidate for the locus of the dual-task interference effects. The superior temporal gyrus was active in the task x time interaction in the overt, but not covert contrast between the speech and tap tasks. In our ROI-based analysis, the response in the overt speech condition was greater than the overt tap condition, but this difference failed to reach even trend levels of significance. This region has been associated with auditory processing, and its activation could be a result of auditory feedback that occurs during overt speech. If this were the case we could expect it to only be active for the retrieval epoch of the DSR task. In fact, this region is active for the encoding, maintenance, and retrieval epochs.

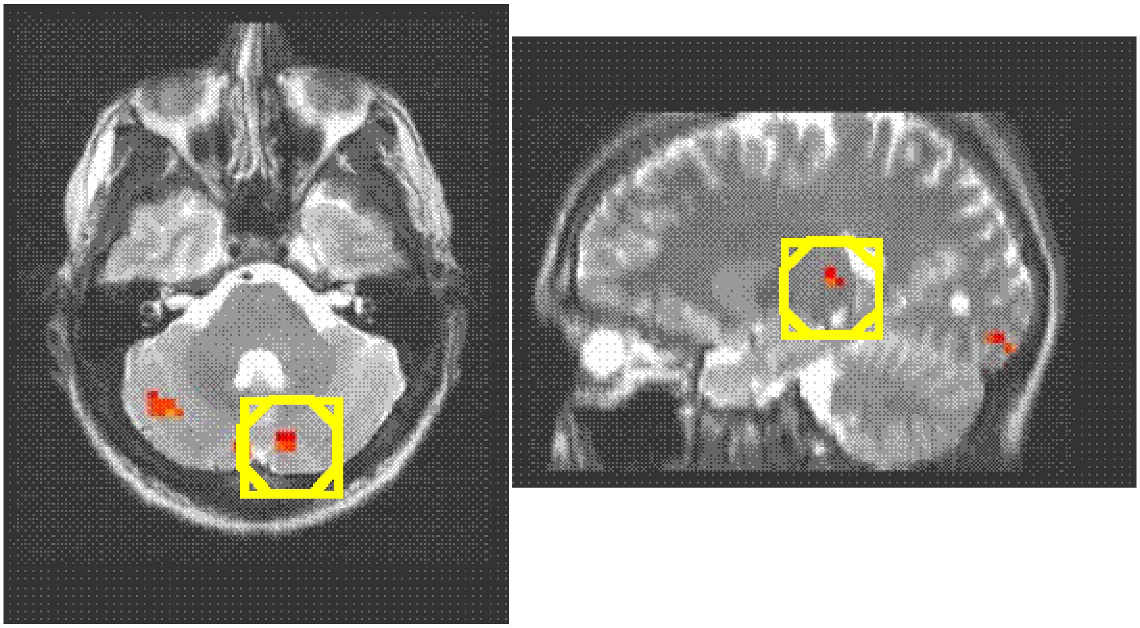


Figure 4.Two regions emerged throughout our analyses as possible candidates for the regions of significant overlap between VWM and overt simple speech. These two regions have only been weakly implicated. Regions are circled in yellow: 4a. This is the candidate region in the left cerebellum. 4b. This is the candidate region in the left superior temporal gyrus.

It is also possible that some of the underlying assumptions of this study may have impacted our ability to identify neural loci of dual-task interference effects. For instance, it is possible that when the simple speech tasks and the DSR task are performed in isolation they activate a different set of brain regions. If this is in fact the situation we could expect a different set of brain regions to be active when the tasks are performed concurrently than those identified in this study. It is also possible that we failed to confirm our hypothesis because neuroimaging does not permit the degree of spatial sensitivity to see differences between speech and tapping. The neural pools between the simple speech and motor tasks may be distinct, but they may also be contiguous or close in proximity to one another. If this is true, our neuroimaging acquisition parameters would not have provided the resolution to fractionate different neuronal subsets within a region in order to see separate areas of activation.

The strong form of our argument assumed that overt and covert versions of the tasks would show the same patterns of activation, with overt versions activating some additional motor regions. Experiment 1 demonstrated that performance of overt concurrent articulation and covert concurrent articulation are more detrimental to recall ability than covert and overt finger tapping. These results were taken to provide support for the idea that there is a common neural locus active for both of the simple speech tasks and VWM that would not overlap with simple motor tasks. At the same time, differences between the overt and covert tasks were noted.

As a secondary goal in these experiments we were interested in exploring the differences between overt and covert versions of the tasks in order to help account for the behavioral effects in Experiment 1. The results from Experiment 1 and behavioral findings showed that overt performance of tapping does not cause the same magnitude of disruption of VWM as concurrent articulation (Larsen & Baddeley, 2003; Saito, 1993, 1994, 1998). However, we did not know if there were differences in whether these tasks were performed overtly or covertly. Our results from both Experiment 1 and Experiment 2 are intriguing. Experiment 1 demonstrated different magnitudes of disruption between overt and covert versions of the concurrent speech and tapping tasks, and Experiment 2 showed a clear dissociation between the two overt and covert version of the simple speech and motor tasks. In particular, in Table 2 we note that overt speech and tapping are always significantly different from one another whereas covert speech and tapping are not significantly different from one another. These results not only demonstrate that the overt tasks are driving the activations that we found, but they also suggest that there is a difference between overt and covert that merits further investigation.

In Experiment 2 we found two patterns of overt versus covert activation that were especially interesting: 1) for some regions the covert response is blunted when compared to overt (e.g. Figure 3c) and 2) for some regions there is a "push-pull" relationship between overt and covert tasks (e.g. Figure 3a). The blunting effect that we found in some areas (left precuneus, right anterior cingulate, left anterior cingulate, left fusiform gyrus, left cerebellum, right parahippocampal gyrus, and right fusiform gyrus) has also been reported in previous neuroimaging studies contrasting overt and covert conditions. These studies revealed that regions that were significantly active for overt tasks have a greater degree of activation than during covert tasks (Barch et al., 1999; Palmer *et al.*, 2001; Shuster & Lemieux, 2005). The lower magnitude of activation in covert tasks could cause many regions not to show up in an analysis or to be overlooked. This blunting effect could have impacted our ability to see common speech vs. tap differences across the overt and covert tap conditions. For example, this effect could explain the differences between overt and covert versions of the tasks as seen in Table 1. In this table the overt tasks are much more active than the covert tasks overall. Based on the fact that there are no behavioral performance measures, and the fact that responses in covert tasks may be blunted, it may be in an experimenter's best interest to employ overt tasks in experiments because these profound differences between the two output modalities may cause problems in subsequent analyses.

A second pattern of activation that we observed in Experiment 2 was a “push-pull” relationship between overt and covert versions of the tasks (the modalities are inversely correlated). Areas where covert tasks produced a positive response and overt tasks produced a negative response were found in the left middle frontal gyrus, right precentral gyrus, and left inferior frontal gyrus. The bilateral insula showed the opposite pattern (positive response for overt tasks and a negative response for covert tasks). This inverse relationship between the two modalities provides strong evidence that overt and covert tasks do not simply differ at the point of motor execution. Instead, the push-pull relationship observed in some regions suggests that when one modality is utilized the other is suppressed. Because we collapsed across modalities on some of the analyses, this push-pull relationship between the two modalities may have eliminated any effect of speech versus tap that we hoped to see in regions associated with inner speech. Unlike the blunting effect that we saw in the covert conditions this is not commented about in the literature.

Finally, the use of overt responding in the neuroimaging environment has historically posed problems with subject movement and motion-related artifact (Yetkin *et al.*, 1995). To reduce subject movement in the scanner some studies resorted to using strict enforcement measures such as a bite plate (Small *et al.*, 1996). The use of a measure like the bite plate would not work in a study such as ours because we wanted to obtain subjects’ verbal responses as a performance measure, which would likely be muffled or unclear with a bite plate. However, more recent neuroimaging work had demonstrated that with the proper design the use of overt tasks in the scanner is plausible. One suggestion has been the use of event-related paradigms to reduce motion related artifact (Birn *et al.*, 1999; Palmer et al., 2001). Because motion-related artifact and activations have distinct time courses the use of this type of design is advantageous, it makes it easier to parse out speech related activations from artificial activation from subject movement. Also, (Barch et al., 1999) demonstrated that performing data analysis on group data rather than individual subjects decreases the number of artifactual activations. In this study we utilized a slow event-related design and only examined group data in an attempt to further minimize artifacts. The lack of motion related activations in this study can be used as further proof that with proper design and analysis overt responding in the scanner is a viable method to acquire subject data.

4.0 CONCLUSIONS

The purpose of these studies was to investigate dual-task interference effects on VWM. Historically, behavioral VWM tasks utilize overt concurrent articulation, whereas neuroimaging studies utilize covert version of the same tasks because of practical issues associated with movement and susceptibility artifact. To our knowledge there have been no studies that directly contrast overt and covert concurrent articulation effects. Experiment 1 aimed to look at the behavioral effects of simultaneous performance of overt and covert versions of concurrent articulation and finger tapping on DSR in order to further characterize the articulatory suppression effect. This experiment demonstrated that all four of the concurrent conditions caused decrements in the VWM ability of subjects. However, both overtly and covertly concurrently articulating “the” were more detrimental to subjects’ recall ability than overtly and covertly tapping. These effects could be attributed to dual-task interference effects at the level of inner speech in VWM, thus, indicating a shared set of neural regions for all speech and VWM.

However, the magnitude of the behavioral interference effects and convergent neuroimaging evidence indicates that the story may be more complex. Experiment 2 was designed to look at whether the decreases in subject performance on the DSR task in Experiment 1 arose from activation of shared neural regions. The results from this experiment provided only weak evidence in favor of our hypothesis that VWM and overt concurrent articulation will have one or more overlapping regions of activation in areas commonly associated with speech processing, and these regions will not be active for overt finger tapping. The best candidate for a shared locus of activation was the left cerebellum, but effects in this region did not reach significance in an *a priori* ROI-based analysis, and speech vs. tap differences were even weaker with covert response generation. There are a number of factors that may account for our failure to confirm our primary hypothesis. The tasks may activate different brain regions when they are performed in isolation versus when they are performed concurrently, neuroimaging may not give us the neural specificity necessary to view differences, and there may be an interaction between overt and covert versions of the simple speech and motor tasks that confound the results.

The results from the overt versus covert contrasts were interesting and unexpected. We anticipated that overt and covert versions of the tasks would behave similarly, however, they paint a picture closer to that seen in recent neuroimaging studies: overt tasks show a set of regions that are distinct from the set of regions that are active for both the covert tasks.

Thus, this study provides a basis for further studies into the neural basis of the articulatory suppression effect in VWM. Two regions emerged as weak candidates for the region of overlap between simple speech and VWM, the left cerebellum and the left superior temporal gyrus. This study lends strong support to the growing literature about differences between overt and covert versions of tasks, and because of these differences we argue in favor of comparing covert behavioral results to covert imaging results and overt behavioral results to overt imaging results because there may be significant differences between the two modalities.

APPENDIX A

EXPERIMENT 1 INSTRUCTIONS

Welcome and thank you for participating in our study. Your participation will help us to better understand the nature of the underlying mechanisms involved in maintaining items in short-term or working memory. Working memory refers to the ability to maintain information for a short period of time. For example, trying to remember a phone number just long enough to dial the number. We hope that with the data we collect from this experiment we will be able to expand upon current theories about how people store and process information in working memory.

In this experiment you will be presented with a series of six letters that will be displayed at the center of the screen.

You will be asked to study the letters as they are being presented and then to rehearse the letters over a short delay. We would like you to rehearse the letters by repeating them over and over in your head. Please, do not chunk the letters together; make sentences from the words; move your mouth; or any other deviation in order to better remember the words. It is very important that you remember them simply by repeating them silently.

After the short delay when you rehearse the letters you will be asked to verbally recall the letters in the order in which they were presented. At this time please clearly repeat out loud the letters in the order in which you remember them. If you do not remember a particular letter please say “skip”. All of your responses will be recorded.

The letters will be white and flash on a black background.

A… B… C… X... Y… Z…

After the letters flash on the screen your cue to begin rehearsing the letters will be green number signs. While these number signs remain on the screen please rehearse the letters by repeating them over and over in your head

# # # # # #

The number signs will change into red question marks, which will be the cue to repeat the letters out loud. When you see the question marks please loudly and clearly state the letters in the order in which you remember them. Again, if you do not remember a letter please say “skip”

? ? ? ? ? ?

While you are performing the short-term memory task you may be asked to perform a secondary task. There will be four secondary tasks that you are asked to perform 1) saying the word “the” out loud approximately three times per second; 2) imagining you are saying the word “the” out loud approximately three times per second; 3) tapping your right index finger approximately three times per second; or 4) imagining you are tapping your right index finger approximately three times per second.

If you are asked to perform a secondary task you will perform the task when the letters are presented (A… B… C…) and also when you are silently repeating them to yourself in order to remember them (when the green number signs are on the screen # # # # # #).

You should not be performing any secondary tasks when you are recalling the letters out loud in the order in which they were presented. In other words, you

? ? ? ? ? ?

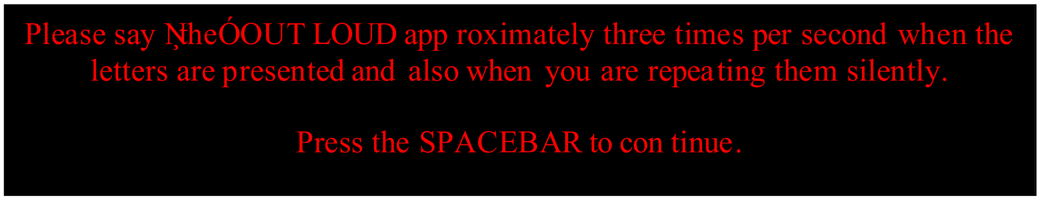
should stop any secondary tasks as soon as the red question marks that are your cue to recall the letters appears on the screen.

Before each trial you will see an instruction screen that will instruct you on how to perform the task. If you are not going to perform any secondary task you will see:

Please do not perform any secondary tasks.

Press the SPACEBAR to continue.

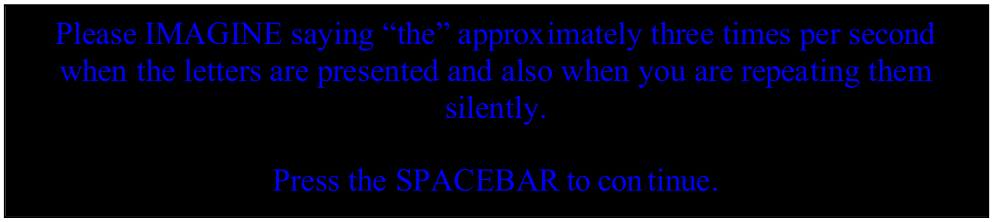
If you are going to perform the short-term memory task and say the word “the” out loud approximately three times per second you will see:



Please say ŅtheÓOUT LOUD approximately three times per second when the letters are presented and also when you are repeating them silently.

Press the SPACEBAR to continue.

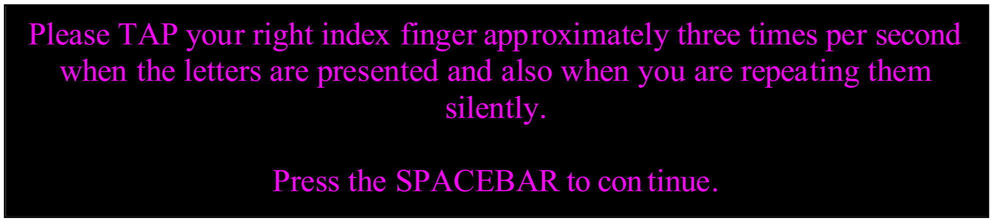
If you are going to perform the short-term memory task and imagine you are saying the word “the” out loud approximately three times per second you will see:



Please IMAGINE saying “the” approximately three times per second when the letters are presented and also when you are repeating them silently.

Press the SPACEBAR to continue.

If you are going to perform the short-term memory task and tap your right index finger approximately three times per second you will see:



Please TAP your right index finger approximately three times per second when the letters are presented and also when you are repeating them silently.

Press the SPACEBAR to continue.

If you are going to perform the short-term memory task and imagine you are tapping your right index finger approximately three times per second you will see:

Please IMAGINE tapping your right index finger approximately three times per second when the letters are presented and also when you are repeating them silently.

Press the SPACEBAR to continue.

If you do not have any questions at this time we will allow you to practice the secondary tasks with a metronome so that you perform the secondary tasks at the correct rate. Then, we will start a practice session of the task so that you are familiar with it for the experiment.

Please ask any questions now.

APPENDIX B

EXPERIMENT 2 INSTRUCTIONS

B.1 VERBALWORKING MEMORY BLOCK

Welcome and thank you for participating in our study. Your participation will help us to better understand the nature of the underlying mechanisms involved in maintaining items in short-term or working memory. Working memory refers to the ability to maintain information for a short period of time. For example, trying to remember a phone number just long enough to dial the number. We hope that with the data we collect from this experiment we will be bale to expand upon current theories about how people store and process information in working memory and about how these processes are mediated by the brain.

In this experiment you will be presented with a series of six letters that will be displayed at the center of the screen.

You will be asked to study the letters as they are being presented and then to rehearse the letters over a short delay. We would like you to rehearse the letters by repeating them over and over in your head. Please, do not chunk the letters together; make sentences form the words; move your mouth; or any other deviation in order to better remember the words. It is very important that you remember them simply by repeating them silently.

After the short delay when you rehearse the letters you will be asked to verbally recall the letters in the order in which they were presented. At this time please clearly repeat out loud the letters in the order in which you remember them. If you do not remember a particular letter please say “skip”.

The letters will be white and flash on a black background.

A… B… C… X... Y… Z…

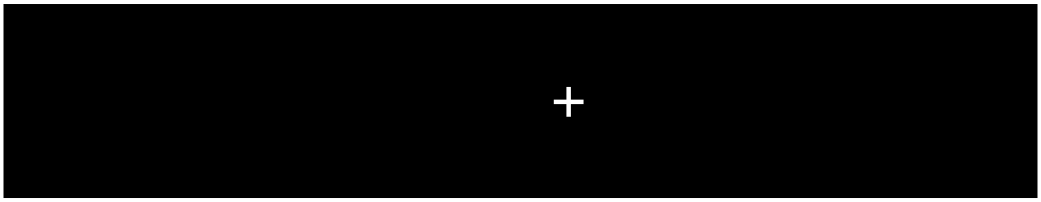
After the letters flash on the screen your cu to begin rehearsing the letters will be green number signs. While these number signs remain on the screen please rehearse the letters by repeating them over and over in your head

# # # # # #

The number signs will change into red question marks which will be the cue to repeat the letters out loud. When you see the question marks please loudly and clearly state the letters in the order in which you remember them. Again, if you do not remember a letter please say “skip”

? ? ? ? ? ?

Following the red question marks you will see a cross at the center of the screen. This cross will remain on the screen for 14 seconds. When you see the cross please just clear your mind and relax.



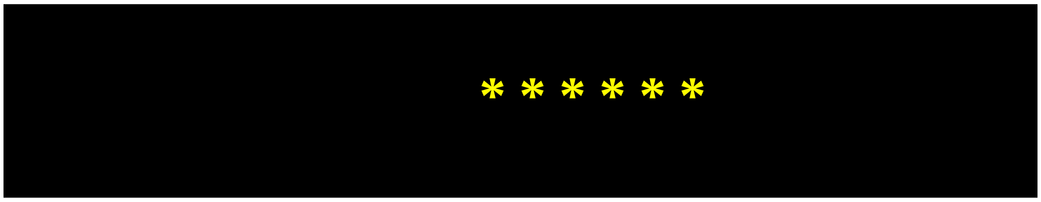
Remember to remain still the entire time you are in the scanner and to speak clearly.

If you do not have any questions we will start a practice session of the task so that you are familiar with it when you get into the scanner.

B.2 SPEECH AND MOTOR BLOCK INSTRUCTIONS

Welcome and thank you for participating in our study. Your participation will help us to better understand the nature of the underlying mechanisms involved in speaking and simple motor tasks. We hope that the data we collect from this experiment will help us to understand how the brain mediates these processes.

In this experiment you will be asked to perform four different tasks. These four tasks will be presented randomly. You will see a two-word instruction for each of the four tasks task followed by yellow stars at the center of the screen. The stars will be your cue to perform the task that the preceding instructions specified.



In the first task you will see the words “Actual Speech” on the computer screen in red. After you see this prompt you will be asked to repeat the word “the” outloud at the rate of approximately 3 times per second.

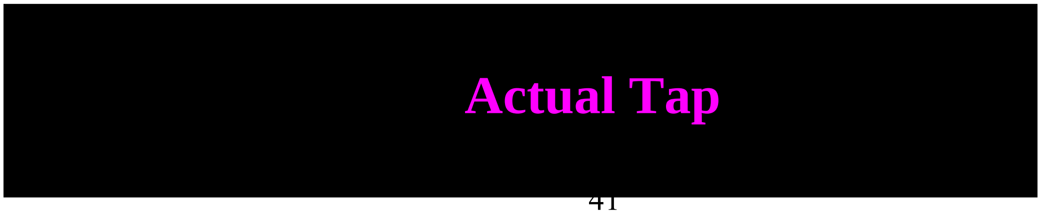

Actual Speech

Again, your cue to begin speaking will be yellow stars at the center of the screen. When these yellow stars appear after the instruction “Actual Speech”, please repeat the word “the” outloud approximately three times per second. Continue to repeat the word “the” the entire time the stars remain on the screen.

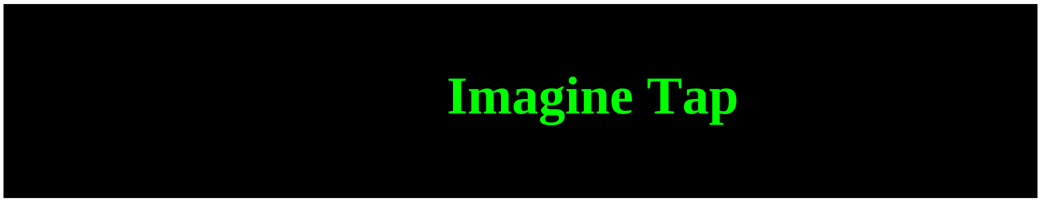
In the second task you will see the words “Imagine Speech” on the computer screen in blue. After you see this prompt you will be asked to imagine you are repeating the word “the” approximately three times per second. In other words, without moving your lips, tongue, or mouth muscles and without producing a sound imagine that you are producing the word “the”. Your cue to begin the task will be the yellow stars at the center of the screen following the instruction “Imagine Speech”.


Imagine Speech

In the third task you will see the words “Actual Tap” on the computer screen in pink. After you see this prompt you will be asked to tap your right index finger at the rate of approximately 3 times per second. You will have a glove on in the scanner so you will be given the opportunity to feel how hard you have to tap in order for the glove to record your response. Your cue to begin tapping will be the yellow stars at the center of the screen following the instruction “Actual Tap”.

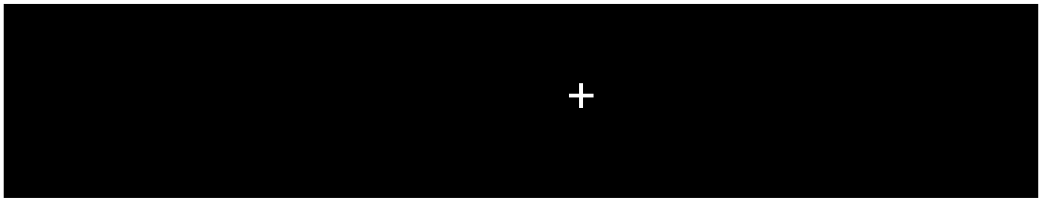

Actual Tap

In the fourth task you will see the words "Imagine Tap" on the computer screen in green. After you see this prompt you will be asked to imagine you are tapping the index finger on your right hand approximately 3 times per second. In other words, without moving any of your hand muscles imagine that your are producing the movements necessary to tap your right index finger approximately 3 times per second and hearing the sound associated with the tapping. As with the other three tasks your cue to imagine you are tapping your right index finger will be the yellow stars at the center of the screen following the instruction "Imagine Tap"



Imagine Tap

Following each task you will see a white cross at the center of the screen. This cross will remain there for 14 seconds. When you see this cross you do not need to do anything. Just relax and clear your mind and stare at the cross.



Remember to remain still the entire time you are in the scanner and to speak clearly.

If you do not have any questions we will start a practice session of the task so that you are familiar with it when you get into the scanner.

BIBLIOGRAPHY

Ackermann, H., & Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language, 89*, 320-328.

Ackermann, H., Wildgruber, D., Daum, I., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fmri) study in humans. *Neuroscience Letters, 247*, 187-190.

Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory. *Psychological Science, 7*(1), 25-31.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A. D., & Hitch, G. J. (1974). *Working memory*. New York: New York Academic Press.

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Experimental Psychology Society, 36A*, 233-252.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., & Cohen, J. D. (1999). Overt verbal responding during fmri scanning: Empirical investigations of problems and solutions. *NeuroImage, 10*, 642-657.

Birn, R. M., Bandettini, P. A., Cox, R. W., & Shaker, R. (1999). Event-related fmri of tasks involving brief motion. *Human Brain Mapping, 7*, 106-114.

Buckner, R. L., Koutstaal, W., Schacter, D. L., & Rosen, B. R. (2000). Functional mri evidence fo a role of frontal and inferior temporal cortex in amodal components of priming. *Brain, 123*, 620-640.

Caplan, D., Alpert, N., Waters, G., & Olivieri, A. (2000). Activation of broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping, 9*, 65-71.

Chein, J. M., & Fiez, J. A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. *Cerebral Cortex, 11*, 1003-1014.

Chein, J. M., Fissell, K., Jacobs, S., & Fiez, J. A. (2002). Functional heterogeneity within broca's area during verbal working memory. *Physiology & Behavior, 77*, 635-6339.

Desmond, J. E., Gabrieli, J. D. E., Wagner, A. D., Ginier, B. L., & Glover, G. H. (1997). Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional fmri. *The Journal of Neuroscience, 17*(24), 9675-9685.

Fiez, J. A., Raife, E. A., Balota, D. A., Schwarz, J. P., Raichle, M. E., & Petersen, S. E. (1996). A poitron emission tomography study of the short-term maintenance of verbal information. *The Journal of Neuroscience, 16*(2), 808-822.

Fissell, K., Tseytlin, E., Cunninham, D., Carter, C., Schneider, W., & Cohen, J. D. (2003). Fiswidgets: A graphical computing environment for neuroimaging analysis. *Neuroimformatics, 1*(1), 11-125.

Hickok, G. (2001). Functional anatomy of speech perception and speech production: Psycholinguistic implications. *Journal of Psycholinguistic Research, 30*(3), 225-235.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, 4*(4), 131-138.

Jung, K.-J., Prasad, P., Qin, Y., & Anderson, J. R. (2005). Extraction of overt verbal response from the acoustic noise in a functional magnetic resonance imaging scn by use of segmented active noise cancellation. *Magnetic Resonance in Medicine, 53*, 739-744.

Kemper, S., Herman, R. E., & Lian, C. H. T. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech or noise. *Psychology and Aging, 18*, 181-192.

Larsen, J., & Baddeley, A. D. (2003). Disruption of verbal short term memory by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source. *Quarterly Hournal of Experimental Psychology A, 56*(8), 1249-1268.

Murray, D. J. (1968). Articulation and acoustic confusability in short term memory. *Journal of Experimental Psychology, 78*(4, Pt 1), 679-684.

Mushiake, H., Inase, M., & Tanji, J. (1991). Neuronal activity in primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. *Journal of Neurophysiology, 66*, 705-718.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. S., Kelley, W. M., & Petersen, S. P. (2001). An event-related fmri study of overt and covert word stem completion. *NeuroImage, 14*, 182-193.

Ravizza, S. M., Delgado, M. R., Chein, J. M., Becker, J. T., & Fiez, J. A. (2004). Functional dissociations within the inferior parietal cortex in verbal working memory. *NeuroImage, 22*, 562-573.

Saito, S. (1993). The disappearance of teh phonological similarity effect by complex rhythmic finger tapping. *Psychologia: An International Journal of Psychology in the Orient, 36*, 27-33.

Saito, S. (1994). What effect can rhythmic finger tapping have on the phonological similarity effect. *Memory and Cognition, 22*(2), 181-187.

Saito, S. (1998). Phonological loop and intermittent activity: A whistle task as articulatory suppression. *Canadian Journal of Experimental Psychology, 52*, 18-23.

Shuster, L. I., & Lemieux, S. K. (2005). An fmri investigation of covertly and overtly produced mono- and multisyllabic words. *Brain Lang, 93*(1), 20-31.

Small, S. L., Noll, D. C., Perfetti, C. A., Hlustik, P., Wellington, R., & Schneider, W. (1996). Localizing the lexicon for reading aloud: Replication of a pet study using fmri. *Neuroreport, 7*, 961-965.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity for action contingency. *Neuron, 41*, 281-292.

Wickens, C. D. (1984). *Processing resources in attention*. London: Academic Press.

Woods, Cherry, & Mazziotta. (1993). Mri-pet registration with automated algorithm. *Journal of Computer Assisted Tomography, 17*, 536-546.

Yetkin, Z. F., Hammeke, T. A., Swanson, S. J., Morris, G. L., Mueller, W. M., McAuliffe, T. L., et al. (1995). A comparison of functional mr activation patterns during silent and audible language tasks. *American Journal of Neuroradiology, 16*, 1087-1092.