2017

Nonword Repetition Task to Evaluate Syllable Stress as a Motor Class

Emily R. Cobun

Nonword Repetition Task to Evaluate Syllable Stress as a Motor Class

Emily R. Cobun

Thesis submitted
to the College of Education and Human Services
at West Virginia University

in partial fulfillment of the requirements for the degree of

Master of Science in
Speech-Language Pathology

Kimberly M. Meigh, Ph.D., Chair
Dennis Ruscello, Ph.D.
Michelle Moore, Ph.D.

Department of Communication Sciences and Disorders

Morgantown, West Virginia

2016

Keywords: motor learning, generalized motor program theory, speech motor control

ABSTRACT

Nonword Repetition Task to Evaluate Syllable Stress as a Motor Class

Emily R. Cobun

Current speech therapy methods and theories are based on generalized motor program (GMP) theory (Schmidt, 1975). GMP theory states a single GMP, or motor program, directs multiple movements of speech (Maas et al., 2008). Additionally, GMP theory asserts these similar muscle movements are part of the same motor class, which allows a GMP to direct performance on novel, untrained patterns of movements (i.e., what is termed "transfer performance"; Chamberlin & Magill, 1992; Schmidt, 1975). Alternatively, movements outside of a learned motor class will be more difficult to perform because a different GMP is controlling these movements. Currently, syllable stress patterns are theorized as the GMP when planning motor speech tasks. This study aims to help clarify the method through which motor speech movements are learned.

Meigh et al. (in press) conducted a study to learn more about speech motor planning. This study found that syllable stress, which was the expected GMP for speech production, did not direct transfer performance on untrained stimuli following training on a speech-like task. Instead, participants encoded speech sound (i.e., phoneme) information during training that influenced transfer results. In Meigh's study, participants were trained using a speech production task but the testing procedure was not speech-based. Meigh's results and interpretation may have been impacted by the study design because of the "mismatch" between modes of training and testing in this study. Therefore, the current study replicated and extended Meigh's experiment using a speech-based training and transfer task.

Twenty-four participants (16 females and 3 males) produced nonsense words (i.e., nonwords) using a motor learning design, which included mass amounts of training followed by an evaluation of performance on untrained stimuli. During training, participants produced different syllable stress patterns while repeating a training list of nonwords. Following training, participants repeated a list of both trained and untrained nonwords that varied in similarity to the trained stimuli. All untrained stimuli varied by motor class (i.e., syllable stress pattern), as well as the phonemes (or sounds). Accuracy of nonword productions were evaluated across transfer stimuli sets, and results revealed participants had learned syllable stress *and* phoneme information during training. These results align with a GMP theory and Meigh (in press) suggesting that more than one GMP memory representation may be encoded during motor learning.

Acknowledgements

I would like to thank the College of Education and Human Services for supporting this research with the Wirt C. and Mae S. Belcher Graduate Education Award. Additionally, thank you to my thesis adviser Dr. Kimberly Meigh, as well as my committee members Dr. Dennis Ruscello and Dr. Michelle Moore. I appreciate your guidance and support throughout this process. I would also like to thank the students in Dr. Meigh's lab who put a considerable amount of time into assisting with this study.

Table of Contents

Introduction

Generalized motor program (GMP) theory is a predominant theory of motor planning and execution in the limb and speech production literatures (e.g., Maas et al., 2008; Schmidt, 2003). GMP theory does not require individuals to store a motor memory for each movement. Instead, this theory allows individuals to encode vital information about the components of movement as a set of rules known as a GMP (Schmidt, 1975). According to GMP theory, only the most important components of the motor behavior must be stored in memory; therefore, memory storage is not taxed by storing all of the details necessary to complete a movement (Schmidt, 1975; Schmidt & Wrisberg, 2004). Moreover, GMP theory provides a straightforward prediction of transfer between motor tasks that are well-learned and novel motor behaviors. Positive transfer from trained to untrained behaviors will occur when both behaviors share the same underlying GMP (Chamberlin & Magill, 1992a, 1992b; Schmidt, 1975). Defining GMPs for a given set of motor tasks, however, has proven difficult for this motor theory.

Researchers have hypothesized numerous possible GMPs for speech, including syllables (e.g., Aichert & Ziegler, 2004; Cholin & Levelt, 2009). Meigh (in press) evaluated syllable stress as a GMP; however, results indicated transfer performance was due to phoneme similarity and not a syllable-sized GMP. Meigh's (in press) methods may not have been sensitive to detecting speech transfer performance as the transfer task did not involve a speech-like production task. The current study aims to replicate and extend Meigh (in press) by using a transfer task that uses the same modality as the training task (a speech-like production task). A general discussion of GMP theory and motor learning will preface a description of Meigh's (in press) study design.

Generalized Motor Programs

The prevailing motor theory used to explain the programming of muscle movements for speech production is generalized motor program (GMP) theory (Schmidt, 1975, 2003). GMP schemas provide an explanation for how novel, untrained movements may be learned without training (Schmidt & Wrisberg, 2004). Pre-stored GMPs may direct several similar movements as long as the movements all share the same underlying set of rules (Schmidt, 1975, 2003). Thus, GMP theory provides a parsimonious account of motor learning.

According to GMP theory, motor programs comprise sets of rules, called schemas, which control the invariant features of movements (Schmidt, 1975, 2003; Schmidt & Wrisberg, 2004). These invariant features provide the core structure necessary to complete the desired movement by programming the relative timing, force, and sequence of the movement (Schmidt, 1975, 1983; Schmidt & Wrisberg, 2004). In this context, relative refers to the abstract set of rules of the motor program. For example, there is not one prescribed value for each invariant feature that must be achieved to execute a movement. Instead, the schema directs a range of values for each invariant feature so that the total amount of time or force for a given movement is detailed (Schmidt, 1975; Schmidt & Wrisberg, 2004). Additionally, the sequences of steps that are required to complete a successful movement are detailed within the GMP, but the specific effectors required to carry out the movement are not described.

Using a basketball jump shot as an example, the invariant features of the throw would include the overall force and time required to make a jump shot from anywhere on the court (e.g., velocity between 6-8 meters per second (m/s)), as well as the sequence of movements to complete the shot (e.g., squat, extend, jump, release the ball). GMPs require specification for a movement to be executed. These variant features, termed parameters in GMP theory, provide *absolute* information about a movement, including the specific timing of the movement, force of the movement, and the specific muscles engaged during the movement (Schmidt, 1975, 2003). Parameters are selected based on the needs of a particular situation or environment, and represent the absolute values required to perform each movement (Cummings & Caprarola, 1986; Schmidt, 1975). Using the previous example, a GMP velocity of 6-8 m/s is required to make a jump shot from anywhere on the basketball court. However, standing at the free throw line requires a parameter of 6.5 m/s using both hands to push the ball up towards the basket. A different parameter may be selected if only the right hand were to be used to throw the ball. To summarize, parameters add specificity to the generalized invariant features. Combining the GMP and parameters allows a basketball player to vary jump shots in different positions around the court under a variety of circumstances (e.g., games, practice). This flexibility allows a basketball player to adapt to numerous contexts even if they were not practiced previously or if the environment is not the same every time. In summary, GMP theory states the GMP schemas, or a set of rules, are encoded into memory, not a series of individualized movements or degrees of freedom (Schmidt, 1975, 2003). This provides efficiency in producing a movement, as fewer components need to be loaded and programmed into memory prior to execution. Prior to initiating the movement, parameters are selected that provide specification, depending on the available environmental and contextual cues, to the abstract rules directed by the GMP (Cummings & Caprarola, 1986; Schmidt, 1975).

GMPs and Motor Learning

Motor programs can be learned through practice. Individuals refine the GMP by trialing different parameters until the desired motor outcome is achieved (e.g., the basketball moves through the net; Cummings & Caprarola, 1986). With each trial, the difference between the desired outcome and the actual outcome are minimized, and the GMP schemas are updated and encoded into memory (Kantak & Winstein, 2012). As the GMP becomes more efficient, the selection of parameters to meet the task goals also becomes more efficient (Cummings & Caprarola, 1986; Schmidt, 1975). In summary, during motor learning the individual encodes the fundamental GMP rules for a given motor behavior, as well as the parameters required for given variations associated with that behavior.

After learning takes place, an individual can retrieve and use an encoded GMP to aid performance on a new untrained movement that is similar to the movement directed by the learned GMP. GMP theory predicts that positive transfer will occur between similar movements that share the same GMP (Schmidt, 1975, 2003). These behaviors are considered to be in the same motor class, and have similar transfer characteristics as both behaviors share the same underlying rules (or GMP; Chamberlin & Magill, 1992a; Schmidt, 1975). Indeed, transfer for within-class movements should be uniform, as the rules of the GMP should not apply to different behaviors within a class with greater or lesser amounts (Chamberlin & Magill, 1992a, 1992b; Crump & Logan, 2010). However, when movements are part of a different motor class, negative transfer, or a decline in performance quality and accuracy, will occur because the movements no longer share the same GMP or set of rules (Chamberlin & Magill, 1992a; Schmidt, 1975, 2003; Wulf & Schmidt, 1988).

Thus, understanding transfer in GMP theory relies on examining motor class boundaries as the point where positive transfer stops occurring. Research has historically fallen short of confirming the size and makeup of a motor class (e.g., Schmidt, 2003; Shea & Wulf, 2005), but research trends show that, although the exact makeup of a GMP is unknown, there has been success in demarcating motor class boundaries, i.e., learning what movements do not belong to the same GMP. For example, learning swimming strokes is unlikely to influence transfer performance for basketball shots. Motor learning studies have aided researchers in determining motor class boundaries by examining where positive transfer (an improvement in performance for untrained movements) and negative transfer (a decrement in performance for untrained movements) takes place (Adams, 1987; Cummings & Caprarola, 1986; Schmidt & Young, 1987). GMP theory infers transfer performance as an indicator of when a well-learned GMP influences novel behavior. As we are able to better define motor class boundaries across a variety of behaviors, it may provide insight into how the motor class develops and the underlying GMP is refined through learning.

GMPs and Speech Production

GMP theory has been adapted to speech motor control theory, and there are many different hypotheses about what GMPs direct speech movements. Researchers have hypothesized GMPs for speech production as phonemes (e.g., Austermann-Hula, Robin, Maas, Ballard, & Schmidt, 2008; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000; Wambaugh, Martinez, McNeil, & Rogers, 1999), syllables (e.g., Aichert & Ziegler, 2004; Cholin & Levelt, 2009; Cholin, Levelt, & Schiller, 2006; Levelt, Roelofs, & Meyers, 1999), and/or words (e.g., Klapp, 2003). However, there is not yet consensus on whether one or more of these levels constitute a speech GMP.

Meigh (in press) conducted a study that investigated syllable stress as a speech GMP by examining motor class boundaries and transfer performance. This study hypothesized syllable stress as a GMP, where first and second syllable stress position constituted a motor class. Other stress positions, e.g., third syllable stress position, were considered to be under the direction of a separate GMP and within a different motor class. Meigh (in press) trained participants to produce first and second syllable stressed nonwords. During production training, participants were given intermittent feedback on their syllable stress production accuracy. Following training, an old-new judgment task was administered in which participants were presented with a nonword and asked to judge if it was “old” (trained nonword) or “new” (novel, untrained nonword). All untrained nonwords were systematically manipulated to vary along two gradients of similarity compared to the training stimuli: 1) syllable stress and 2) phonetic similarity. Untrained stimuli either shared the same syllable stress as the trained stimuli (within motor class variable) or had third syllable stress (outside motor class variable). Additionally, untrained stimuli also varied in the number of shared phonemes, as well as the order of phonemes in a consonant-vowel (CV) unit. During the old-new judgment task, participants made the determination of “old” or “new” by pushing a button on a response box with their index finger. Reaction times from this task were used in Meigh’s (in press) analyses.

Using GMP theory as a framework, Meigh (in press) hypothesized syllable stress as the GMP influencing nonword production. Therefore, positive transfer would be observed for nonwords with first and second syllable stress (the trained motor class); specifically, these untrained nonwords would have relatively the same uniform reaction time pattern as the training stimuli. However, negative transfer for nonwords with third syllable stress (an untrained motor class) was predicted to result in slower reaction times for these transfer stimuli compared to other untrained transfer stimuli with a trained stress pattern.

Meigh’s (in presss) results indicated that phoneme similarity *and* syllable stress position impacted reaction times. Specifically, no significant differences in reaction time were noted between untrained stimuli with trained syllable stress versus untrained syllable stress (as originally predicted). However, there was a significant decrease in reaction time for within-class untrained stimuli with the same phonemes and phoneme order as the trained stimuli. Stimuli with similar phonemes, phoneme order, *and* syllable stress position were the slowest reacted to stimuli of all untrained stimuli. Meigh (in press) interpreted these results to mean the trained syllable stress pattern may not be a speech GMP and that several factors (phoneme novelty and syllable stress position) were learned during training to influence transfer performance.

In Meigh's (in press) study, participants were trained with a nonword repetition (NWR) task (a speech "production" measure), and the transfer task was conducted using audition as the modality (a perceptual measure). The transfer task was an old-new judgment task during which the participant pushed a button on a response key to indicate if the presented nonword was old or new. Meigh's (in press) results may have been secondary to the difference in modalities that were evaluated and not differences in the underlying GMP. Specifically, this mismatch in modalities makes it unclear if a motor representation was being evaluated during the transfer task. The mismatch between modalities (i.e., a production task during training and a perceptual task during testing) during the experiment may have influenced the results of Meigh's (in press) study.

The current study aims to replicate and extend Meigh's (in press) study, but training and testing modalities will be matched in the present study. The current study will both train *and* evaluate transfer using speech-like production tasks. This evaluation of methods will help to determine if Meigh's (in press) interpretation of the study was influenced by the study design. In the present study, participants will be trained in first and second syllable stress with a list of nonwords using the same training procedure as Meigh (in press); however, learning and motor memory will be evaluated by a NWR task. It is expected that, in this study, positive transfer will be seen for nonwords that are within the same motor class (first and second syllable stressed words) as the trained words while negative transfer will be seen for words outside of the trained motor class (third syllable stressed words). By matching training and testing modalities, this study avoided the questions raised by the mismatch of modalities in Meigh's (in press) study when training was conducted with a production task and evaluation of transfer with a perceptual task.

This study aimed to evaluate syllable-stress as a potential speech GMP. Specifically, the study determined the extent to which performance transfers when the mode of training and transfer task are the same following extensive motor learning training. The expected findings of this study are detailed in the hypotheses below and align with motor class predictions from Schmidt's (1975, 2003) GMP theory.

Hypothesis 1

Reaction times for stimuli outside of the trained motor class will be significantly slower than stimuli that belong to the same motor class as the trained nonwords. Transfer Sets 3 and 4 will be compared as both stimuli are equally dissimilar from the trained stimuli, i.e., these sets share the same phonemes (not present in the trained stimuli) in the same CV order. The only difference between these two stimuli sets is the proposed motor class.

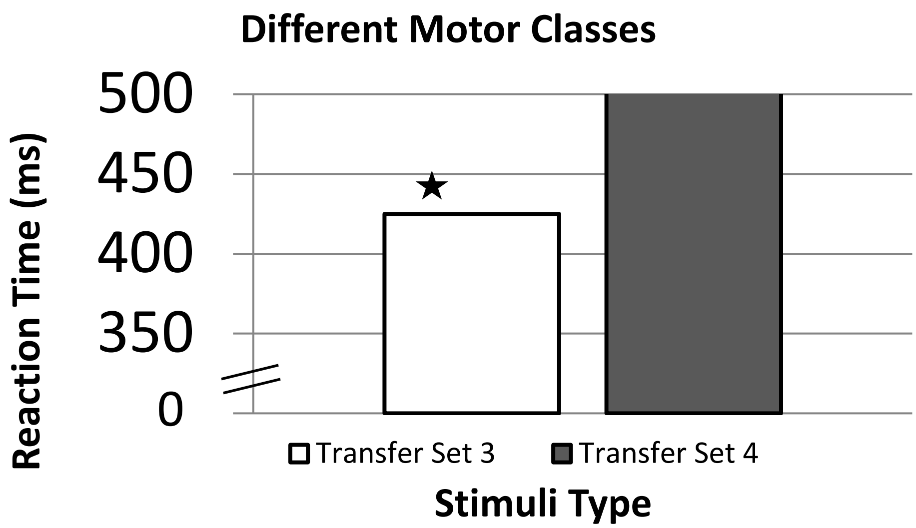

Figure 1: An example of reaction time differences across different motor classes.

Note: The star symbol in this figure indicates a significant effect.

Hypothesis 2

Reaction times for within-class stimuli will be uniform in their distribution (i.e., without significant difference), as these stimuli share the same motor class as the trained nonwords.

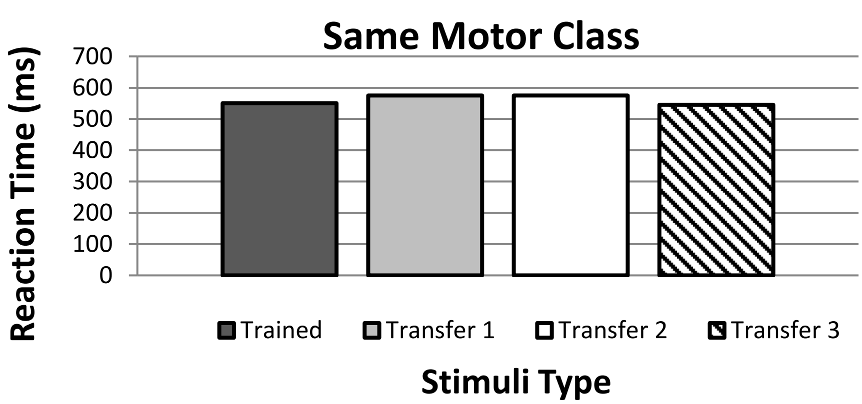

Figure 2: An example of uniform reaction times for stimuli within the same proposed motor class.

Methods

Participants

Nineteen participants (16 females, 3 males) between the ages of 18 and 35 were recruited to participate in this study. All participants were prescreened for language and education prior to coming into the lab. Specifically, participants self-reported being monolingual with English as their dominant language as defined by a custom-designed language questionnaire (Appendix A) and holding a high school diploma (or equivalent). When participants met this initial criteria, they were scheduled to come into the lab and were screened for normal speech and hearing skills.

To be considered as having normal speech production, participants were required to be within normal limits for all categories on an oral mechanism exam: lingual protrusion and retraction; facial symmetry; labial protrusion, retraction, and closure; elevation and depression of the mandible; and velopharyngeal movement and symmetry. Also, no evidence of tremor in the face, lips, or neck could be observed. Additionally, vowel prolongation and diadochokinetic rates were required to be within one standard deviation of the minimum normative values (Duffy, 1995). Finally, participants could have no articulation errors or disfluent speech when producing any of the stimuli on the Test of Minimal Articulation Competence (T-MAC; Secord, 1981) or during conversational speech. For this study, normal hearing was defined as passing a pure tone hearing screening at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz in at least one ear at 40 dB (American Speech-Language-Hearing Association, 1990). Also, participants were required to correctly repeat the Northwestern University Test #6 words (NU-6; Tillman & Carhart, 1966) with no more than one mistake (45/46 on list 2A male speaker recordings). Participants who failed the screening were dismissed from the study and referred to the West Virginia University Speech and Hearing Center for a complete speech and language and/or audiological evaluation.

Participant recruitment took place using IRB-approved fliers posted in public spaces in West Virginia University buildings, IRB-approved ads posted to the Speech Motor Control Lab Facebook page, and IRB-approved ads sent via email listservs. Additional recruitment occurred through the West Virginia University psychology pool (SONA). All study procedures were conducted in the WVU Speech Motor Control Laboratory (807-G Allen Hall) by the primary investigator or a trained IRB-approved investigator.

Participants interested in the study contacted Dr. Meigh or the PI via email or phone regarding their interest in the study. At that time, a pre-screening language questionnaire (see Appendix A) was administered. Participants who passed the pre-screening were scheduled for an experimental session and those who did not were thanked for their time. During the scheduled experimental session, written consent was obtained according to procedures outlined by West Virginia University's Institutional Review Board. Following consent, participants completed the screening procedure to determine eligibility for the study as outlined above. Screening procedures took 30 minutes to complete. Participants who passed the screening protocol were reimbursed with a $20 gift card at the end of the study.

Experimental Setup

The experimental procedure consisted of a training task where participants learned to produce novel nonwords with a first or second syllable stress pattern and a transfer task where participants applied their training knowledge to repeat trained and untrained nonwords. After the screening, participants were seated in a comfortable chair, and a dynamic headset unidirectional microphone (Shure WH20XLR) was placed approximately one-inch mouth-to-microphone distance. The microphone was connected to a 64-bit Dell Latitude 3340 laptop utilizing Windows 7 operating system, which ran custom software "Stimulate" (described below). A stereo speaker (Bose Companion 2 Series 3) was centered on the table approximately 15 inches in front of the participant, and a second monitor was placed directly in front of the participant. During the transfer task, the headset microphone was connected to a serial response box (Psychology Software Tools; Model 200a), which captured participants' vocal reaction time for each stimulus. A digital voice recorder (Olympus DM-901) was also centered on the table approximately 6 inches from the participant to record the NWR task.

Stimuli

All stimuli used in this study were also used in Meigh (in press); however, Meigh derived these stimuli from Kendall et al. (2005). These stimuli were used due to their complex and novel construction. Specifically, the stimuli were controlled on the following parameters: consonant and vowel structure, lexical familiarity, frequency of consonants and vowels, tenseness or laxness of vowels, and interphonemic transitional gestural frequency (a measure of articulatory complexity and frequency). As with Meigh (in press), participants encountered trained and untrained stimuli during different phases of this experiment. Participants encountered trained stimuli during syllable stress training, and untrained stimuli were added during the NWR task. Each type of stimuli and the study phases in which they were encountered are explained in more detail below.

Stimuli Features

Prior to data collection, Meigh (in press) re-recorded the stimuli to control for differences in duration that were present in Kendall et al. (2005). After re-recording the stimuli, all of the nine phonemes in each nonword were verified for phonetic accuracy by three blinded individuals who were familiar with and used phonetic transcription regularly. Three blinded listeners also verified syllable stress placement in each recording. For complete lists of all stimuli, see Appendix A.

Training Stimuli

Thirty nonwords (three-syllables in length) with first or second syllable stress were used as the training stimuli for this study. These syllable stress patterns were based on frequency of stress in English. First and second syllable stress in three syllable words is very common in English (mean frequencies: 9.38 and 10.12 respectively) while third syllable stress is much less common (mean frequency: 3.79; Clopper, 2002). Of the thirty stimuli, ten were experimental (taken from Kendall et al., 2005) and the rest were filler stimuli taken from Roy and Chiat's (2004) three-syllable stimuli, Kendall et al.'s (2005) high frequency stimuli, and Dollaghan and Campbell's (1998) three syllable stimuli. Meigh (in press) used filler stimuli to ensure an equal number of "old" and "new" responses in the old-new judgment task. In this study, the filler stimuli will be used during training to maintain continuity with Meigh (in press)'s methods, to maximize distance between the presentations of similar stimuli, and to decrease potential repetition effects.

Untrained Stimuli

Untrained stimuli were only encountered during the NWR task. There were four sets of untrained stimuli, which varied systematically across two gradients of similarity: phoneme similarity (as quantified by phonemes and phoneme order in a CV unit) and syllable stress (first, second, or third syllable stress).

Transfer Set 1

The first stimuli set was the same as the trained experimental stimuli set except the first and second syllables were reversed. For example, the training item "/zæ**nɔ**dʒəθ/" became Transfer Set 1 "/**nɔ**zædʒəθ/;" note the stressed syllable is bolded. These stimuli have first and second syllable stress and are considered within the same motor class as the experimental training stimuli.

Transfer Set 2

Transfer Set 2 stimuli shared the same phonemes that were found in the trained stimuli, but each syllable had a different phoneme order than the trained stimuli. For instance, the training item "/zæ**nɔ**dʒəθ/" and Transfer Set 2 item "/næ**θo**dæp/" (stressed syllable is bolded) share several phonemes but not in any given CV unit. These stimuli share the same syllable stress pattern and motor class as the trained stimuli.

Transfer Set 3

Transfer Set 3 was created by Meigh (in press) to be maximally different in phonetic composition compared to the trained stimuli and is composed of phonemes not found in the training stimuli (e.g., training stimuli "/zæ**nɔ**dʒəθ/" and Transfer Set 3 stimuli "/**gi**gʊðib/"). However, these stimuli share the same syllable stress patterns (bolded in the example) and proposed motor class as the trained experimental stimuli.

Transfer Set 4

These are the same stimuli as used in Transfer Set 3, except the syllable stress pattern is on the final, third syllable (e.g., Transfer Set 3 stimuli **"/gi**gʊðib/" and Transfer Set 4 stimuli "/gigʊ**ðib/**"). Thus, this transfer set represents an untrained motor class compared to the trained stimuli.

Motor Class Training

Motor class training included three tasks: perception-production training, syllable stress training, and recognition probes. For this experiment, the motor class was represented by first and second syllable stress. Perception-production training was used so participants could familiarize themselves with the training program, “Stimulate,” and the experimental stimuli. The purpose of syllable stress training was to provide participants frequent practice opportunities to encode the trained syllable stress pattern as a GMP with first and second syllable stress positions constituting a trained motor class. Recognition probes were used to assess phonetic accuracy of the nonwords being encoded into memory during training by evaluating participants’ ability to recognize familiar and unfamiliar auditory presentations of the trained stimuli.

Perception-Production Training

Perception-production training was used to acclimate participants to the training stimuli and the Stimulate software. Stimulate is a custom-built software platform that presents the trained stimuli, records the participant’s production, and analyzes the participant’s production for differences in intensity (used to mark syllable stress). For example, on a given trial the participant heard a recorded nonword and repeated the nonword and its associated stress pattern. Stimulate recorded the participant’s response, analyzed this response for changes in intensity, and displayed visual feedback on the participant’s stress production. This visual feedback included three horizontal blue lines that represented the expected stress pattern for each syllable of the nonword (Figure 1). Stress was represented by the vertical placement of the blue lines on the screen, where only one of the three blue lines was elevated (either first or second blue line) to indicate the targeted syllable stress pattern. The participant’s production of stress in the nonword was represented by yellow dots superimposed over the modeled stress profile (the blue lines).

Therefore, participants simultaneously viewed the modeled stress profile and their resultant stress as feedback. During this portion of training, participants were given two trials of every nonword with 100% feedback on each trial to ensure the participant was familiar with the stimuli and Stimulate.

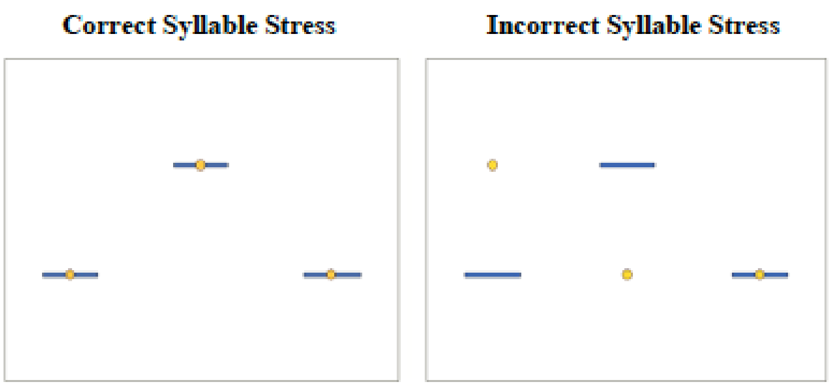

Figure 1: Stimulate visual feedback

Syllable Stress Training

The procedure for syllable stress training was exactly the same as perception-production training except for the number of trials practiced and the amount of feedback the participant received. The participant completed up to two separate training blocks of stimuli each consisting of 120 pseduo-randomized training nonwords. Accuracy in producing the targeted syllable stress pattern was monitored by Stimulate, as well as perceptually by the examiner. To aid learning of the nonwords, participants only received feedback 65% of the time (e.g., Almelaifi, 2013; Meigh & Shaiman, 2010). Participants' accuracy was variable across training blocks with an average training accuracy of 84%. All participants required both training blocks, and training was discontinued after two training blocks regardless of accuracy. Accuracy variability was attributed to participant fatigue and inattention despite providing multiple breaks and maximum motivational cues. Following training, it was hypothesized that participants had encoded syllable stress as a GMP with first and second syllable stress positions constituting a trained motor class.

Recognition Probes

After each training block, the participant completed a recognition probe consisting of 10 trials. Recognition probes provide an indirect measure of the phonetic accuracy being encoded into memory during syllable stress training. During each recognition probe trial, the participant heard recordings of three nonwords, including one nonword practiced during syllable stress training. The other two words were foils. One foil changed the vowel in the stressed syllable while the other changed difficult to perceive phonemes, such as /ð/ and /v/ in the final position of the nonword. The participant was asked to identify which nonword they practiced during training. Participants' accuracy was variable across recognition probes with an average recognition probe accuracy of 76%. All participants required both recognition probes, and training was discontinued after two probes regardless of accuracy. As noted previously, accuracy variability was attributed to participant fatigue and inattention despite providing multiple breaks and maximum motivational cues. Fewer number of trials within a probe may have attributed to these lower accuracy scores as each probe had only ten trials.

Nonword Repetition Task

A NWR task was used to evaluate transfer of performance between the trained nonwords and the untrained nonwords that varied in syllable stress and phonetic similarity. Although the training methods used for this study were replicated from Meigh (in press), this transfer task differed from Meigh's methods. In Meigh (in press), an old-new judgment task was used to determine what variables affected transfer following extensive syllable stress training. Results revealed transfer was affected by phoneme similarity *and* syllable stress. Therefore, Meigh (in press) inferred that syllable stress might not be the GMP directing transfer performance. However, Meigh's interpretation of the results may have been impacted by the study design, as the old-new judgment task involved perceptual judgments of auditory presentations of the nonwords and did not incorporate *production* of the nonwords. The transfer task in the present study is critical to re-evaluating Meigh's (in press) results, as it aligns the training and transfer task to the same modality: movement of the articulators.

For this task, the participants listened to a nonword and then repeated it twice in succession. The experiment was controlled by Eprime (Schneider, Eschman, & Zuccolotto, 2002) using the following protocol: an auditory presentation of the word "ready," an auditory presentation of the nonword stimulus, a period of silence randomized between 200 and 400 milliseconds (ms), a 500 Hz alert tone, a response period of four seconds to record the participant's production, a second period of silence randomized between 800 and 1200ms, a second 500 Hz tone, and another four second response period to record the participant's second production. The period of silence was varied during the trial to increase participants' attention to the stimuli, as well as decrease participants' over-anticipation of when they should respond (Magill, 2001). Only a single auditory presentation of each nonword occurred per trial, which is similar to Meigh's (in press) old-new judgment task. However, two repetitions of the nonword were elicited to compensate for potential false starts, participant production errors, errors in recording the signal, and unwanted increases in reaction time due to repetition effects (Magill, 2001). Similar research using NWR tasks have included a range of successive repetitions, including two repetitions (Warker, Xu, Dell, & Fisher, 2009) to greater than four (Edwards & Lahey, 1998; Lisman, 2011). Meigh (in press) noted that during perception-production training, most participants were able to remediate articulation errors based on feedback from one repetition.

During the NWR task, participant's responses (spoken into the headset microphone) activated a voice response key, which collected a vocal reaction time. In addition to the vocal reaction time, trained study personnel made a perceptual judgment of correct pronunciation during the NWR task. A digital recorder was used to record the task in case the voice key failed on the response box.

Data Analysis

Trained transcribers reviewed the recordings from the NWR task for accuracy in producing the nonwords correctly, as well as producing the correct syllable stress pattern. Transcribers were instructed to listen to each experimental trial as many times as necessary to evaluate phoneme and syllable stress production. To evaluate phoneme production accuracy, transcribers coded each phoneme within a nonword trial as correct, a substitution (where the substituted phoneme was also transcribed), an addition, omission, or distortion. To evaluate syllable stress accuracy, transcribers rated their perception of syllable stress occurring in the first, second, or third position of the nonword. Each participant was rated by at least one transcriber.

Syllable stress production accuracy was required for inclusion in the analysis. However, a less restrictive accuracy criterion was used to assess phoneme production based on the rules in Table 1. These rules included execution errors (e.g., distortions) as accurate productions. It was anticipated that with more repetitions participants would have eliminated these motor execution errors; however, the complexity of the stimuli paired with the NWR task demands (a single presentation of the stimuli) limited overall production accuracy. The exclusion of all other errors (including phonetic substitutions and additions) was based on the programmatic nature of the errors indicating the wrong memory representation was selected for programming and execution. Only a single phoneme error was allowed for any given stimulus to be included in the analysis. If multiple errors were present, even if these errors were present in Table 1, they were excluded from the analysis.

Table 1: Types of phoneme errors permitted using a lax scoring system for including trials in the analysis

Rule	**Justification**
1. Omitted final consonant	The majority of these omissions were difficult to perceive phonetic contrasts (e.g., /θ/; Jongman, Wayland, & Wong, 2000).
2. Change in manner of final consonant from fricative to affricate (specific substitution of /ʒ/ → /dʒ/ in the final position)	Participants may have had a difficult time perceiving the difference due to frication (Jongman et al., 2000).
3. Substitutions within the same manner of articulation with an adjacent place of articulation (e.g., producing a /ʃ/ for /s/)	Manner of articulation has been postulated as a segment GMP (e.g., Austermann-Hula, Robin, Maas, Ballard, & Schmidt, 2008; Ballard, Robin, McCabe, & McDonald, 2010). Using a GMP framework, these substitutions maintained the same motor class (e.g., fricatives) across different motor class members (e.g., alveolar vs. palatal).

Rule	Justification
4. Devoicing of any consonant	Devoicing is estimated to have occurred secondary to production complexity, where participants were unable to accurately produce all features of the stimuli when attempting to produce the correct manner and place of articulation.
5. A single vowel substitution for any other vowel	Vowel substitutions were adjacent to the targeted sound (e.g., /ɔ/ and /ɑ/) or centralized (e.g., /ə/ or /ʌ/) indicating a potential motor execution error.
6. Any consonant or vowel distortion	This error was a considered a motor execution error.

The resulting number of included stimuli for the reaction time analysis was unequal between stimuli types: Trained (168), Transfer Set 1 (145), Transfer Set 2 (128), Transfer Set 3 (89), and Transfer Set 4 (61). Therefore, the original proposed analysis was not performed. A dependent variable of accuracy was used in the statistical analyses below. All data was coded as either "correct" or "incorrect" for phonetic production (using the rules above), as well as syllable stress production. The hypotheses for this experiment were maintained with the new dependent variable of accuracy. It was hypothesized that participants would be significantly more accurate when producing first and second syllable stress stimuli (i.e. trained motor class) than stimuli with third syllable stress (i.e. untrained motor class). Also, it was expected that phonetic similarity would not impact participants' accuracy on the NWR task. Therefore, it was hypothesized that participants would produce all nonwords within the trained motor class (Trained and Transfer Sets 1-3 stimuli) with uniform accuracy.

Results

Cochran's Q Test Syllable Stress Accuracy

Cochran's Q Test (Cochran, 1950) was run to determine if the percentage of correctly produced syllable stress patterns varied across stimuli type. Sample size was adequate to use the χ^2 distribution approximation (Tate & Brown, 1970). Results from 189 nonwords were used for this analysis. Participants were 95.8% accurate in producing correct syllable stress in Trained nonwords, 97.4% accurate for Transfer Set 1 stimuli, 95.8% accurate for Transfer Set 2 stimuli, 70.9% accurate for Transfer Set 3 stimuli, and 47.6% accurate for Transfer Set 4 stimuli. The percentage of correctly produced syllable stress was statistically significant across stimuli, $\chi^2(4)=230.774$, $p<0.001$. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons (adjusted p values are presented). Participants produced stimuli from Transfer Sets 3 and 4 with significantly less accuracy than nonwords from Trained, Transfer Set 1, or Transfer Set 2 stimuli sets ($p<0.001$). Participants produced words from Transfer Set 4 with significantly less syllable stress accuracy than words from Transfer Set 3 ($p<0.001$). There was no statistically significant difference in syllable stress accuracy between Trained stimuli and Transfer Sets 1 or 2 stimuli ($p=1.00$).

Cochran's Q Test Phoneme Accuracy

Cochran's Q test (Cochran, 1950) was run to determine if the percentage of accurately produced phonemes in nonwords varied across stimuli type. Sample size was adequate to use the χ^2 distribution approximation (Tate & Brown, 1970). Results from 190 nonwords were used for this analysis. Participants were 93.2% accurate at producing phonemes correctly in Trained nonwords, and overall accuracy declined across untrained stimuli: Transfer Set 1 - 77.4%, Transfer Set 2 - 72.1%, Transfer Set 3 stimuli - 62.6%, and Transfer Set 4 - 65.3%. The percentage of phonemes that were correctly produced was statistically significant across different stimuli, $\chi(4)=68.413$, $p<0.001$. Pairwise comparisons were performed using Dunn's (1964) procedure with Bonferroni correction for multiple comparisons (adjusted p values presented). Phoneme production accuracy in Trained stimuli was significantly greater than Transfer Set 1 ($p=0.001$), Transfer Set 2 ($p<0.001$), Transfer Set 3 ($p<0.001$), and Transfer Set 4 ($p<0.001$) stimuli. Transfer Set 1 stimuli were produced with significantly greater accuracy than both Transfer Sets 3 ($p=0.004$) and 4 ($p=0.035$). No significant difference was found between Transfer Set 2 and Transfer Sets 1 ($p=1.00$), 3 ($p=0.223$), or 4 ($p=0.987$). Transfer Sets 3 and 4 were not significantly different ($p=1.000$).

Discussion

This experiment investigated syllable stress as a GMP with first and second syllable stress positions as a motor class. Participants were trained extensively to produce first and second syllable stress in complex nonwords. After training, they completed a NWR task to evaluate repetition accuracy across two variables of interest: syllable stress and phonetic similarity. It was hypothesized participants would repeat nonwords with first and second syllable stress (i.e., trained motor class) with significantly greater accuracy than nonwords with third syllable stress (i.e., untrained motor class). Additionally, it was hypothesized that phonetic similarity would not influence accuracy during the NWR task. Thus, it was expected that participants would not vary in their repetition accuracy when producing nonwords within the trained motor class (Trained and Transfer Sets 1-3 stimuli). The results of this study only partially aligned with these predictions.

Based on the above hypotheses, a motor class boundary between Transfer Set 4 and the other stimuli (Trained, Transfer Sets 1-3) was predicted for the first analysis. Transfer Set 4 stimuli contained an untrained syllable stress pattern, and it was predicted that participants would not be able to rely on the trained syllable GMP to help them repeat Transfer Set 4 stimuli. The results of the first analysis align with these findings as participants had significantly more incorrect productions with Transfer Set 4 stimuli than any other stimuli set. However, the results of the syllable stress analysis for this hypothesis also revealed Transfer Set 3 was significantly less accurate from Trained and Transfer Sets 1 and 2 stimuli. This was an unexpected result that does not align with the current proposed speech GMP. Transfer Set 3 stimuli was within the trained motor class, therefore participants were expected to perform with a uniform level of accuracy compared with other within-class stimuli (Trained, Transfer Set 1, and Transfer Set 2 stimuli). The main difference between the within-class stimuli is the phoneme order and phonemes used in each set. Thus, this finding may indicate that phonemes may have also influenced accuracy during the NWR task.

Similar results are noted in the second analysis where participant accuracy decreased as phonemes within the untrained stimuli became less similar to the trained stimuli. As noted earlier, the trained syllable stress pattern should have aided repetition of untrained stimuli with the same stress pattern. However, this analysis suggests that phonetic differences in the stimuli were influencing repetition accuracy despite the presence of the trained stress pattern.

The results of both analyses do not support a GMP model based exclusively on syllable stress as the GMP. The results suggest that *both* features of syllable stress and phonemes may have been encoded into memory during syllable stress training. In the current GMP model, participants should not have encoded any features of the phonemes that were present in the trained nonwords. Yet, both analyses suggest phonemes influenced repetition accuracy despite the presence of a motor class boundary based on trained versus untrained stress patterns. These findings were similar to those of Meigh (in press), which found that features of phonemes and syllable stress were encoded during training.

There were several limitations of this study that may have influenced the above results. First, only nineteen participants successfully completed this study, which is a smaller sample size than originally projected. The proposed sample size was n=24. Twenty-three participants were screened for this study; however, only data from 19 participants were used in the analysis. Of the four individuals who did not complete the experimental portion of the study, three participants failed the screening criterion, and one participant did not adequately learn to produce first and second syllable stress. The proposed sample size of n=24 was based on Meigh's previous sample sizes (Meigh, in press); however, participants found the NWR task more difficult to perform than Meigh's old-new judgment task. The literature shows a wide range of sample sizes for studies that included NWR tasks performed by neurologically intact individuals (some of the studies included both children and adults within the sample size). The sample sizes ranged from 18 to 58; however, most of the sample sizes were greater than the analyzed sample of 19 from this study (Gupta, 2010; Sadagopana & Smith, 2013; Sasisekaran, 2013; Vitevitch & Luce, 2005).

Therefore, the literature supports adding additional participants to this study. With additional participants, it is possible that more statistical power can be achieved and a parametric test could be performed.

A second limitation of this study was the use of rules to determine phonetic accuracy. The use of these rules may have biased the outcome of this study by including stimuli that favored certain stimuli sets. As noted previously, participants were much less accurate at repeating nonwords than investigators anticipated during the NWR task. As noted in the methods, the use of rules was implemented when the reaction time analysis was still being attempted. The use of rules allowed for inclusion of more data points, and all rules were justified as being indicative of motor execution or perceptual difficulties with the stimuli. Specifically, it was assumed that motor execution issues (e.g., distortions) did not reflect difficulty in selecting the wrong motor program (i.e., different processing levels of speech motor control; Van der Merwe, 1997) but were used by the participant to simplify motor execution. In the future, a less confounded analysis would be to perform a percent consonant correct analysis (PCC; Dollaghan, 1998), where each individual consonant is determined to be correct or incorrect and then divided by the total number of consonants. Like the rules implemented in this study, the PCC analysis does not consider distortions or additions as incorrect (Dollaghan, 1998). The PCC analysis does code substitutions and deletions as incorrect, which indicate a different phoneme than the intended phoneme was selected from memory and produced. This analysis has been used widely in healthy, neurologically intact individuals (Dillion & Pisoni, 2006; Wren, McLeod, White, Miller, & Roulstone, 2013).

A third limitation of this study is the limited number of transcribers coding the NWR data for accuracy. At least one trained transcriber listened to each participant's NWR recording and evaluated syllable stress and phonetic accuracy. However, multiple raters should have been used to ensure judgments were accurately made and were replicable across raters. In the future, at least two transcribers should analyze each recording, and a third transcriber should listen to resolve any discrepancies between the first two transcribers.

A fourth limitation of this study was the number of nonword repetitions produced by participants during the NWR task. Participants listened to a single presentation of a nonword and then repeated it twice. However, the number of repetitions that the participant produced may have impacted their accuracy. Although two repetitions of the nonwords were selected for this study, the literature does not suggest a specific number of repetitions for optimal performance and accuracy during a NWR task. In fact, research shows that repetitions of nonwords in previous studies range from two to over four (Edwards & Lahey, 1998; Lisman, 2011; Warker, Xu, Dell, & Fisher, 2009). Meigh et al. (in press) noted that participants were typically able to remediate articulation errors after receiving feedback from just one trial. However, many participants who made errors when speaking made errors on both the first and second repetition. Also, participants spoke the nonword twice in quick succession with only an 800 to 1200 m/s pause between the stimuli. Thus, there was no time between stimuli to deliver feedback during the NWR task. Adding additional audio presentations of the nonword and/or allowing the participant to have more opportunities to repeat the nonword might allow the participant to self-correct syllable stress or phoneme errors.

Although the above limitations may have contributed to the results of this study, an alternative interpretation of GMP theory should also be considered. GMP theory states that programs for motor movements are made up of schemas or rules, which direct invariant features that are necessary to complete the movement (Schmidt, 1975). Invariant features provide the basic information that is necessary to perform a movement. They include the relative time, force, and sequences required to complete a desired movement (Schmidt, 1975, 2003; Schmidt & Wrisberg, 2004). Syllable stress fits nicely into this model because the time, force, and sequence required to produce a certain type of syllable stress can be controlled (Aichert & Ziegler, 2004; Cholin & Levelt, 2009). However, other aspects of speech production may also have invariant characteristics. As noted in the introduction, a variety of speech GMPs have been proposed. The results of this study suggest phonemes influenced nonword repetitions on the NWR task, which might fit a GMP model with a phoneme motor program.

It is unclear what features of phonemes might be considered GMPs. Manner of articulation has been put forth as one alternative (Austermann-Hula et al., 2008; Ballard et al., 2007, 1999; Knock et al., 2000). Different constrictions of airflow based on manner may align with invariant force and timing characteristics required for a GMP. Additionally, phonemes may be considered a programming unit of speech production (e.g., Davis, Farias, Baynes, 2009; Ziegler, Thelen, Staiger, & Liepold, 2008). Moreover, other features of phonemes (e.g., place of articulation) align with a phoneme-based motor class. For instance, regardless of place of articulation all stops are produced with the same timing and force characteristics.

Although a phoneme-level GMP sounds promising, this study was unable to investigate this type of GMP because stimuli were not controlled for manner of articulation or other features associated with a phoneme GMP. In a future study to investigate this type of GMP, stimuli should be developed that vary along two motor classes (e.g. fricatives and stops). These stimuli sets should be controlled and standardized for other phoneme-level features, such as place of articulation and voicing. A similar procedure to this study could be implemented, where participants learn nonwords containing a single manner of articulation (e.g., fricatives).

Following training, a NWR task with trained and untrained nonwords that varied in trained manners (e.g., fricatives) and untrained manners (e.g., stops) could be implemented. Trained motor class variables could also be evaluated by including untrained stimuli within the same manner of articulation (e.g., fricatives) that only varied by place of articulation. This type of study would allow further examination of manner of articulation as a speech GMP, and it could clarify a portion of the current results that indicate that features of phonemes were encoded during learning.

The results of this study also suggested a motor class boundary had formed following syllable stress training. Participants were less accurate in producing an untrained syllable stress pattern (third syllable stress) compared to producing the trained stress pattern. These results differ from those of Meigh (in press), which did not find a significant difference (or motor class boundary) between Transfer Set 3 and Transfer Set 4 stimuli. This difference in results may be attributed to the different dependent measures evaluated; however, the overall conclusions between Meigh (in press) are similar. These results suggest syllable information was encoded, as well as phoneme information during training. Speech is a complex hierarchical process, and there are many levels within the process that need to be controlled: speech processing, programming, and sequencing, as well as execution including vocal tract and articulator movements (Meyer & Gordon, 1985; Munhall, 1993; Van der Merwe, 1997). Therefore, it is not difficult to imagine that multiple GMPs may be operating simultaneously in speech production. Using a Chunking Theory framework (Miller, 1994), individuals encode and group smaller memory representations into larger, more manageable memories that are easier to remember and use. In speech, the smaller units may be phonemes while the larger units could be syllables or words. Chunking Theory states that we learn the smaller units, such as phoneme GMPs, and as we become more practiced with these GMPs we group them into larger units, like syllable GMPs (Gilbert, Boucher, & Jemel, 2015; Miller, 1994). For example, the training stimuli /zaʃɔdʒəz/ is difficult to learn as an entire unit. In Chunking Theory, the phonemes (e.g. /z/, /a/, /ʃ/, /ɔ/, /dʒ/, /ə/) would be encoded initially during training; however, as training progressed, this information would be condensed and encoded into larger syllable units (e.g. /za/, /ʃɔ/, /dʒəz/; Gilbert, Boucher, & Jemel, 2015; Miller, 1994). Thus, during this study, participants may have simultaneously learned GMPs for both phonemes and syllable stress during syllable stress training and encoded these GMPs into memory. As training progressed, these GMPs would have blended so that participants could use them concurrently.

Chunking Theory may be used to explain the results of this study. Transfer Sets 2 through 4 were too dissimilar from the trained stimuli for both GMPs to be used together. In Transfer Set 2 the phonemes were the same as those that were trained, but they were in a different order within the syllables. Therefore, the chunked syllable GMP may have broken apart causing the participants to rely only on the phoneme GMP. Transfer Set 3 had completely different phonemes than those that were trained but was still within the trained motor class. This could have forced the participants to abandon the learned phoneme GMP and rely only on the syllable stress GMP. Lastly, Transfer Set 4 had the lowest accuracy in both the syllable stress and phoneme accuracy analysis. This may be because Transfer Set 4 had different phonemes than those that were trained and it was outside of the trained motor class. Therefore, the participants could no longer use either the syllables stress or the phoneme GMPs that they had previously learned.

In summary, this study investigated syllable stress as a GMP for speech motor movements. After extensive training in producing first and second syllable stress, participants engaged in a NWR task. The results of this study did not support a GMP for speech based solely on syllable stress. Instead, results indicate both phonemes and syllable stress impacted the participants' accuracy in producing the nonwords. These results align with those of Meigh (in press), which found that both phoneme and syllable stress information were encoded into memory and used by participants during the NWR task. This may indicate GMPs for both syllable stress and phonemes may be learned during training and influence performance on untrained nonwords. Future studies should further investigate GMPs for speech. These studies could help clarify the type of GMP(s) that govern speech motor movements.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74. http://dx.doi.org/10.1037/0033-2909.101.1.41

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88(*1), 148–159. http://doi.org/10.1016/S0093-934X(03)00296-7

Almelaifi, R. (2013). *The role of "focus of attention" on the learning of non-native speech sounds: English speakers learning of mandarin Chinese tones* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh. PA.

American Speech-Language-Hearing Association. (1990). Guidelines for screening hearing impairment and middle-ear disorders. *ASHA*, *32*(2), 17–24.

Austermann-Hula, S. N., Robin, D. A., Maas, E., Ballard, K. J., & Schmidt, R. A. (2008). Effects of feedback frequency and timing on acquisition and retention of speech skills in acquired apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1088–1113. http://dx.doi.org/10.1044/1092-4388(2008/06-0042)

Ballard, K. J., Maas, E., & Robin, D. A. (2007). Treating control of voicing in apraxia of speech with variable practice. *Aphasiology*, *21*(12), 1195–1217. http://doi.org/10.1080/02687030601047858

Ballard, K. J., Robin, D. A., Knock, T. L., & Schmidt, R. A. (1999). Influence of frequency of feedback and order of stimulus presentation on treatment for apraxia of speech. Presented at the Clinical Aphasiology Conference. Key West, FL.

Ballard, K. J., Robin, D. A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *53*(5), 1227-1245. http://dx.doi.org/10.1044/1092-4388(2010/09-0130)

Chamberlin, C. J., & Magill, R. A. (1992a). A note on schema and exemplar approaches to motor skill representation in memory. *Journal of Motor Behavior*, *24*(2), 221–224. http://dx.doi.org/10.1080/00222895.1992.9941617

Chamberlin, C. J., & Magill, R. A. (1992b). The memory representation of motor skills: A test of schema theory. *Journal of Motor Behavior*, *24*(4), 309–319. http://dx.doi.org/10.1080/00222895.1992.9941627

Cholin, J., & Levelt, W. J. (2009). Effects of syllable preparation and syllable frequency in speech production: Further evidence for syllabic units at a post-lexical level. *Language and Cognitive Processes*, *24*(5), 662–684. http://dx.doi.org/10.1080/01690960802348852

Cholin, J., Levelt, W. J., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*(2), 205–235. http://doi.org/10.1016/j.cognition.2005.01.009

Clopper, C. G. (2002). Frequency of stress patterns in English: A computational analysis. *Indiana University Linguistics Community Working Papers*, *2*(1), 1–9.

Cochran, W.G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256-266. http://dx.doi.org/10.2307/2332378

Crump, M. J. C., & Logan, G. D. (2010). Episodic contributions to sequential control: Learning from a typist's touch. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 662–672. http://dx.doi.org/10.1037/a0018390

Cummings, J. F., & Caprarola, M. A. (1986). Schmidt's schema theory: Variability of practice and transfer. *Journal of Human Movement Studies*, *12*, 51–57.

Davis, C., Farias, D., & Baynes, K. (2009). Implicit phoneme manipulation for the treatment of apraxia of speech and co-occurring aphasia. *Aphasiology*, *23*(4), 503-528. http://dx.doi.org/10.1080/02687030802368913

Dillon, C. M., & Pisoni, D. B. (2006). Nonword repetition and reading skills in children who are deaf and have cochlear implants. *The Volta Review*, *106*(2), 121-145.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136–1146. http://dx.doi.org/10.1044/jslhr.4105.1136

Duffy, J. R. (1995). *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, MO: Mosby.

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252. http://dx.doi.org/10.1080/00401706.1964.10490181

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*(2), 279–309. http://doi.org/10.1017/S0142716400010079

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. http://dx.doi.org/10.3758/BRM.41.4.1149

Gilbert, A. C., Boucher, V. J., & Jemel, B. (2015). The perceptual chunking of speech: A demonstration using ERPs. *Brain Research*, *1603*, 101-113. http://dx.doi.org/10.1016/j.brainres.2015.01.032

Gupta, P. (2010). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology: Section A*, *56*(7), 1213-1236. http://dx.doi.org/10.1080/02724980343000071

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3 Pt 1), 1252-1263. http://dx.doi.org/10.1121/1.1288413

Kantak, S. S., & Winstein, C. J. (2012). Learning–performance distinction and memory processes for motor skills: A focused review and perspective. *Behavioural Brain Research*, *228*(1), 219–231. http://dx.doi.org/10.1016/j.bbr.2011.11.028

Kendall, D., McNeil, M. R., Shaiman, S., & Pratt, S. (2005). Phonetic encoding of infrequent articulatory phonetic transitions. *Aphasiology*, *19*(1), 39–52.

Klapp, S. T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior*, *35*(2), 135–50. http://dx.doi.org/10.1080/00222890309602129

Knock, T. L., Ballard, K. J., Robin, D. A., & Schmidt, R. A. (2000). Influence of order of stimulus presentation on speech motor learning: A principled approach to treatment for apraxia of speech. *Aphasiology*, *14*(5/6), 653–668. http://dx.doi.org/10.1080/026870300401379

Levelt, W. J., Roelofs, A., & Meyers, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. http://dx.doi.org/10.1017/S0140525X99001776

Lisman, A. L. (2011). *Attentional focus and motor speech performance* (Unpublished master's thesis). University of Colorado at Boulder, Bouldar, CO.

Maas, E., Robin, D. A., Austermann-Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, *17*, 277–298. http://dx.doi.org/10.1044/1058-0360(2008/025)

Magill, R. A. (2001). *Motor Learning: Concepts and Applications* (6th ed.). Boston, MA: McGraw-Hill Higher Education.

Meigh (in press). A novel investigation of GMP Theory: Syllable stress as a motor class variable. *Journal of Speech, Language, and Hearing.*

Meigh, K. M., & Shaiman, S. (2010). *Evaluation of a comparable control parameter for speech and nonspeech: The effect of amount of practice on intraoral pressure accuracy*. Poster presented at the Motor Speech Conference, Savannah, GA.

Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. *Journal of Memory and Language*, *24*(1), 3–26. http://doi.org/10.1016/0749-596X(85)90013-0

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *101(*2), 343-352. http://dx.doi.org/10.1037/0033-295X.101.2.343

Munhall, K. G. (1993). The skill of speech production. In J. L. Starkes & F. Allard (Eds.), *Cognitive Issues in Motor Expertise* (pp. 201–221). Amsterdam, NL: Elsevier Science Publishers B.V.

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2-to 4-year-olds: Evidence from typically developing children. *Journal of Speech, Language and Hearing Research*, *47*(1), 223-234. http://dx.doi.org/10.1044/1092-4388(2004/019)

Sadagopana, N., & Smith, A. (2013). Age differences in speech motor performance on a novel speech task. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1552-1566. http://dx.doi.org/10.1044/1092-4388(2013/12-0293)

Sasisekaran, J. (2013). Nonword repetition and nonword reading abilities in adults who do and do not stutter. *Journal of Fluency Disorders, 38*(3), 275-289. http://dx.doi.org/10.1016/j.jfludis.2013.06.001

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*(4), 225-260. http://dx.doi.org/10.1037/h0076770

Schmidt, R. A. (1983). On the underlying structure of well-learned motor responses a discussion of Namikas and Schneider and Fisk. *Advances in Psychology*, *12*, 145–165. https://doi.org/10.1016/S0166-4115(08)61990-1

Schmidt, R. A. (2003). Motor schema theory after 27 years: Reflections and implications for a new theory. *Research Quarterly for Exercise and Sport*, *74*(4), 366–375. http://doi.org/10.1080/02701367.2003.10609106

Schmidt, R. A., & Wrisberg, C. (2004). *Motor Learning and Performance: A problem-based learning approach.* (3rd ed.). Champagne, IL: Human Kinetics.

Schmidt, R. A., & Young, D. E. (1987). *Transfer of movement control in motor skill learning*. San Diego, CA: Academic Press, Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). Eprime User's Guide (Version 2.0). Pittsburgh: Psychology Software Tools, Inc.

Secord, W. (1981). Test of minimal articulation competence. San Antonio, TX: Psychological Corporation.

Shea, C. H., & Wulf, G. (2005). Schema theory: A critical appraisal and reevaluation. *Journal of Motor Behavior*, *37*(2), 85–102. http://dx.doi.org/10.3200/JMBR.37.2.85-102

Tate, M., & Brown, S. (1970). Note on the Cochran *Q* test. *Journal of the American Statistical Association*, 65(329), 155–160. http://dx.doi.org/10.1080/01621459.1970.10481069

Tillman, T. W., & Carhart, R. (1966). An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6. DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0639638

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (Vol. 2, pp. 3–18). New York, NY: Thieme.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*(2), 193-204. http://doi.org/10.1016/j.jml.2004.10.003

Wambaugh, J. L., Martinez, A. L., McNeil, M. R., & Rogers, M. A. (1999). Sound production treatment for apraxia of speech: Overgeneralization and maintenance effects. *Aphasiology*, *13*(9–11), 821–837. http://doi.org/10.1080/026870399401902

Warker, J. A., Xu, Y., Dell, G. S., & Fisher, C. (2009). Speech errors reflect the phonotactic constraints in recently spoken syllables, but not in recently heard syllables. *Cognition*, *112*(1), 81–96. http://doi.org/10.1016/j.cognition.2009.03.009

Wren, Y., McLeod, S., White, P., Miller, L.L., & Roulstone, S. (2013). Speech characteristics of 8-year old children: Findings from a prospective population study. *Journal of Communication Disorders*, *46*(1), 53-69. http://dx.doi.org/10.1016/j.jcomdis.2012.08.008

Wulf, G., & Schmidt, R. A. (1988). Variability in practice: Facilitation in retention and transfer through schema formation or context effects? *Journal of Motor Behavior*, *20*(2), 133–149. http://dx.doi.org/10.1080/00222895.1988.10735438

Ziegler, W., Thelen, A., Staiger, A., Liepold, M. (2008). The domain of phonetic encoding in apraxia of speech: Which sub-lexical units count? *Aphasiology, 22*(11), 1230-1247. http://dx.doi.org/10.1080/02687030701820402

Appendix A

Pre-Screening Language Questionnaire

This prescreening was administered over the phone or via email once a subject had contacted the WVU Speech Motor Control Lab for further information about participating in the study. If the subject passed the prescreening, he/she was scheduled for an appointment to participate in the experiment. If the subject did NOT pass the prescreening, he/she was thanked for his/her time. The subject was informed of his/her ineligibility to participate in the study, and all information obtained during the prescreening was destroyed.

1. When you were learning to speak as a child, did you learn any language other than English?

_____**YES**: a.) Did you speak more than a few phrases at home?

_____**YES**: Not eligible for the study

_____**NO**: Still Eligible, continue with question b.)

b.) Did you understand more than a few phrases at home?

_____**YES**: Not eligible for the study

_____**NO**: Eligible for the study, continue to question 2.)

______**NO**: Did anyone in your family, like your parents or grandparents, speak a language other than English?

_____**YES**: a.) Did you ever speak more than a few phrases to them in that language?

______**YES**: Not eligible for the study

_______**NO**: Still Eligible, continue with question b.)

b.) Did you understand more than a few phrases when they were speaking that language?

______**YES**: Not eligible for the study

_______**NO**: Eligible for the study, continue to question 2.)

_____**NO**: Eligible for the study, continue to question 2.)

2.) Have you taken more than 2 semesters of a foreign language in high school or college?

_____**YES**: **Are you able to speak more than a few phrases fluently?**

____**YES**: Not eligible for the study

_____**NO**: Still Eligible, continue with question b.)

b) Are you able to comprehend more than a few phrases fluently?

____**YES**: Not eligible for the study

______**NO**: Eligible

______**NO**: Eligible for scheduling

Appendix B: Stimuli

Training Stimuli

Training stimuli were composed of experimental stimuli taken from Kendall et al. (2005), as well as filler stimuli adapted from Kendall et al. (2005), Roy and Chiat (2004), and Dollaghan and Campbell (1998). These stimuli were used during motor class training and the NWR task.

Training Stimuli: Experimental

Each nonword had first or second syllable stress (bolded and underlined) and was considered to be part of the trained motor class.

Table 2: Experimental training stimuli

Training Stimuli
/te**næ**rok/
/**kæ**θotæs/
/**sæ**θodæk/
/**zo**tenav/
/za**ʃɔ**dʒəz/
/zæ**ʃɔ**dʒəθ/
/dʒə**zɔ**zæk/
/zæ**nɔ**dʒəθ/
/dʒʌ**nɔ**zæk/
/**θʌ**rasæθ/

Filler Stimuli:

Each nonword had first or second syllable stress (bolded and underlined) and was considered to be part of the trained motor class. These stimuli were included in this study to maintain continuity with Meigh's (in press) original stimuli and to create adequate space between similar stimuli during the NWR task.

Table 3: Filler experimental stimuli

Filler Stimuli
/**<u>ze</u>**səfis/
/ze**<u>sə</u>**fin/
/**<u>vɔ</u>**səfis/
/zi**<u>rə</u>**fin/
/**<u>di</u>**səfis/
/nʌ**<u>zi</u>**rəz/
/**<u>nʌ</u>**zirəv/
/rʌ**<u>zi</u>**rəs/
/**<u>sʌ</u>**zirəʤ/
/pʌ**<u>zi</u>**rəʤ/
/**<u>sɪ</u>**nədɔb/
/la**<u>dɪ</u>**heb/
/**<u>nə</u>**nɑbəp/
/tɑ**<u>fu</u>**pəl/
/**<u>gæ</u>**zəmin/
/rɪ**<u>gə</u>**sɛt/

Untrained Stimuli

Untrained stimuli varied along two experimental parameters: phoneme similarity and syllable stress. These stimuli were only encountered during the NWR task.

Transfer Set 1

Each nonword had first or second syllable stress (bolded and underlined) and was considered to be part of the trained motor class. These stimuli were the same as the trained nonwords except the first and second syllables were reversed.

Table 4: Transfer Set 1 stimuli

Transfer Set 1 Stimuli
/**<u>næ</u>**terok/
/θo**<u>kæ</u>**tæs/
/θo**<u>sæ</u>**dæk/
/te**<u>zo</u>**nav/
/**<u>ʃɔ</u>**zadʒəz/
/**<u>ʃɔ</u>**zædʒəθ/
/**<u>zɔ</u>**dʒəzæk/
/**<u>nɔ</u>**zædʒəθ/
/**<u>nɔ</u>**dʒʌzæk/
/ra**<u>θʌ</u>**sæθ/

Transfer Set 2

Each nonword had first or second syllable stress (bolded and underlined) and was considered to be part of the trained motor class. These stimuli used the same phonemes as the training nonwords and Transfer Set 1, but the phonemes were in a different order in each syllable.

Table 5: Transfer Set 2 stimuli

Transfer Set 2 Stimuli
/**<u>ʃɔ</u>**ʤəzɔd/
/**<u>vu</u>**zæʃɔm/
/**<u>fo</u>**zæʃɔd/
/**<u>ko</u>**zæʃɔm/
/**<u>dɔ</u>**ʤəzɔd/
/**<u>næ</u>**θodæp/
/ra**<u>sæ</u>**θon/
/sʌ**<u>ve</u>**næθ/
/na**<u>sæ</u>**θoʃ/
/vi**<u>ʃə</u>**dæk/

Transfer Set 3

Each nonword had first or second syllable stress (bolded and underlined) and was considered to be part of the trained motor class. These stimuli were created by Meigh (in press) and contained different phonemes than the nonwords in Training Stimuli, Transfer Set 1, and Transfer Set 2.

Table 6: Transfer Set 3 stimuli

Transfer Set 3 Stimuli
/gi**bɪ**ðɪb/
/ʒi**bʊ**tʃeð/
/**tʃe**ðugʊʒ/
/**ʒʊ**gijub/
/gʊ**gi**ðʊtʃ/
/**bɪ**ðetʃug/
/**gi**gʊðib/
/**tʃe**jiwɪʒ/
/bʊ**tʃi**tʃeʒ/
/tʃʊ**tʃu**bɪʒ/

Transfer Set 4:

Each nonword had third syllable stress (bolded and underlined) and was considered to be *outside* of the trained motor class. These stimuli were identical to Transfer Set 3 except stress was placed on the third syllable.

Table 7: Transfer Set 4 stimuli

Transfer Set 4 Stimuli
/gibɪ**ðɪb/**
/ʒibʊ**tʃeð/**
/tʃeðu**gʊʒ/**
/ʒʊgi**jub**/
/gʊgi**ðʊtʃ**/
/bɪðe**tʃug**/
/gigʊ**ðib**/
/tʃeji**wɪʒ/**
/bʊtʃi**tʃeʒ**/
/tʃʊtʃu**bɪʒ/**