The Effect of Blocked Versus Random Practice on Nonword Acquisition and Retention

Lauren Rose Scheiner

University of Colorado at Boulder, lauren.scheiner@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/slhs_gradetds

Part of the Physiology Commons, and the Speech and Hearing Science Commons

Recommended Citation


https://scholar.colorado.edu/slhs_gradetds/8

This Thesis is brought to you for free and open access by Speech, Language and Hearing Sciences at CU Scholar. It has been accepted for inclusion in Speech, Language, and Hearing Sciences Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
The Effect of Blocked Versus Random Practice on Nonword Acquisition and Retention

by

Lauren R. Scheiner

B.A., Reed College, 2005

A thesis submitted to the
Faculty of the Graduate School
of the
University of Colorado-Boulder
in partial fulfillment of the requirement for the degree of
Master of Arts
Department of Speech, Language and Hearing Sciences
July 2011
This thesis entitled:
The Effect of Blocked Versus Random Practice on Nonword Acquisition and Retention

Written by Lauren R. Scheiner

Approved for the Department of Speech, Language and Hearing Sciences

__________________________________________
Dr. Neeraja Sadagopan, Committee Chair

__________________________________________
Dr. Lori Ramig

__________________________________________
Dr. David Sherwood

__________________________________________
Dr. Phillip Gilley

Date _______________________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB Protocol #10-0253
Abstract

Scheiner, Lauren R. (M.A., Speech, Language and Hearing Sciences)

The Effect of Blocked Versus Random Practice on Nonword Acquisition and Retention

Thesis directed by Assistant Professor Neeraja Sadagopan

The purpose of the current research is to elucidate the effects of two practice schedules (blocked vs. random) on motor learning in the context of speech production. Specifically, this study examined the effect of blocked vs. random practice on the acquisition and retention of novel speech motor sequences (nonwords) in healthy young adults. Participants underwent a comparable amount of practice in either blocked or random order on Day 1 and were tested for retention on the following day.

Kinematic measures of timing (duration, duration variability and relative duration variability) and measures of behavioral accuracy were obtained for four time points during acquisition on Day 1 and during the retention test on the following day in order to test two primary hypotheses: 1) the blocked practice group would outperform the random practice group on measures of accuracy, duration, duration variability and relative variability during acquisition; and 2) the random practice group would outperform the blocked practice group at retention testing along the same outcome measures.

Consistent with patterns established in the limb motor learning literature, participants in the blocked practice demonstrated better overall performance during acquisition, though not necessarily greater gains in performance. While blocked practice
resulted in poorer performance than random practice on Day 2, the difference between the
groups was significant only for accuracy.
Acknowledgments

First and foremost, I’d like to thank my advisor Dr. Sadagopan for all of her help and support. She was the first professor at the University of Colorado who sincerely encouraged my research interests and has consistently made my experience worthwhile. She has helped me to think critically and methodically. She inspires me to work harder and to do better.

My committee members, Dr. Ramig, Dr. Sherwood and Dr. Gilley deserve many thanks for helping to shape my thesis into a robust and meaningful study. Their collective wisdom is immense and I’m incredibly appreciative of their advice.

While I was in California Dr. Kan was an invaluable resource and support. Her passion for research is an inspiration. Her respect for me as a researcher gave me confidence that my project would be a success.

I’d also like to thank Dr. Gilley and Nicholas Walker for their help with the statistics for this experiment. I would have been utterly lost in a statistical quagmire without their help.

Many thanks to everyone in the CU Speech Lab who helped me recruit and run participants, and for keeping me company during long days in the lab.

Thanks to my Colorado friends who kept me sane (Shani O’Brien) or drove me insane (Jesse Capecelatro). Special thanks to friends who saw me through my undergrad thesis and are miraculously still around for this one: Ian Schreiner, Andrew Maser, Beth Ungerecht, Mike Lubing and Ethan Nowak. I’d like to thank my tumblr readers for their encouragement and support through my #thesis.

Finally, I’d like to thank my family back on Long Island, especially my Mom, Dad, sister and Victoria Maccarone. Your love and support means so much to me: thank you, thank you, thank you.
Table of Contents

List of Tables ........................................................................................................................................... vii
List of Figures .............................................................................................................................................. viii
1 Introduction ................................................................................................................................................ 1
  1.1. Learning in the context of speech motor control ................................................................. 1
  1.2 Neurophysiological mediation of motor learning ................................................................................. 2
  1.3 Practice schedules and motor learning ........................................................................................... 4
      1.3.1 Limb motor learning .............................................................................................................. 4
      1.3.2 Speech motor learning .......................................................................................................... 8
  1.4 Rationale for the current study and hypotheses .................................................................................. 10
      1.4.1 Hypothesis ............................................................................................................................ 11
2 Methods .................................................................................................................................................... 12
  2.1 Participants ......................................................................................................................................... 12
  2.2 Testing ............................................................................................................................................... 13
  2.3 Apparatus .......................................................................................................................................... 14
  2.4 Stimuli ................................................................................................................................................. 15
  2.5 Protocol and Data Collection ........................................................................................................... 18
  2.4 Data Analysis ..................................................................................................................................... 22
      2.4.1 Behavioral Accuracy (Percent phonemes correct) .............................................................. 23
      2.4.2 Kinematic Analysis (Mean nonword duration, duration variability and relative duration
                    variability) ............................................................................................................................... 23
  2.5 Statistical Analysis ............................................................................................................................. 25
      2.5.1 Acquisition ............................................................................................................................. 25
      2.5.2 Retention ............................................................................................................................... 26
3 Results ....................................................................................................................................................... 27
  3.1 Acquisition ......................................................................................................................................... 27
  3.2 Retention .......................................................................................................................................... 43
Discussion ..................................................................................................................................................... 50
  Acquisition ............................................................................................................................................. 50
  Retention ............................................................................................................................................... 55
  Effects of nonword complexity ............................................................................................................. 56
Conclusion .................................................................................................................................................... 58
Bibliography ............................................................................................................................................... 60
List of Tables

Table 2.1 Experimental nonwords and accompanying visual stimuli .......................... 16
Table 2.2 Nonword stimuli length, durations and phonotactics ................................. 17
Table 2.3 Schematic of experimental protocol ............................................................. 22
Table 3.1 Planned comparisons results for accuracy over time for the blocked and random practice groups ................................................. 31
Table 3.2 Planned comparisons for duration for the blocked and random practice groups ................................................................. 35
Table 3.3 Planned comparisons results for duration variability for both practice groups 39
Table 3.4 Planned comparisons results for relative variability for both practice groups 43
List of Figures

Figure 2.1 Example of a presentation slide .................................................. 19
Figure 3.1 Nonword phoneme accuracy during acquisition ......................... 28
Figure 3.2 Trajectory of phoneme accuracy over time for both practice groups .... 29
Figure 3.3 Phoneme accuracy at acquisition time points for both practice groups .... 30
Figure 3.4 Nonword durations during acquisition ........................................ 32
Figure 3.5 Trajectory of duration over time for both practice groups ............... 33
Figure 3.6 Duration at acquisition time points for both practice groups .......... 34
Figure 3.7 Nonword duration variability at acquisition for both practice groups .... 36
Figure 3.8 Trajectory of duration variability over time for both practice groups .... 37
Figure 3.9 Duration variability at acquisition time points for both practice groups .... 38
Figure 3.10 Nonword relative variability during acquisition .......................... 40
Figure 3.11 Trajectory of relative variability over time for both practice groups .... 41
Figure 3.12 Relative variability at acquisition time points for both practice groups .... 42
Figure 3.13 Phoneme accuracy by group at retention testing ......................... 44
Figure 3.14 Nonword phoneme accuracy at retention testing ....................... 45
Figure 3.15 Duration by group at retention testing ....................................... 46
Figure 3.16 Nonword duration at retention testing ..................................... 46
Figure 3.17 Duration variability by group at retention testing ....................... 47
Figure 3.18 Nonword duration variability at retention testing ....................... 48
Figure 3.19 Relative variability by group at retention testing ....................... 49
Figure 3.20 Nonword relative variability at retention testing ....................... 49
1 Introduction

1.1. Learning in the context of speech motor control

It is clear that practice and learning are inextricably linked. Schmidt & Lee (2005) define motor learning as a set of internal processes associated with practice of experience leading to relatively permanent changes in the capability for responding. In studies of learning involving motor tasks, learning is inferred from behavioral changes in skilled motor performance. The above definition implies that practice is associated with change in performance over time. The initial phase of motor learning is acquisition, during which a skill is practiced. Acquisition is measured by assessing improvement in motor skill performance. This is also commonly known as the practice effect. To assess whether there has been a change in the capability for responding, performance must be tested after practice is complete. Performance can be measured by assessing retention of the motor skill or generalization to a novel, untrained skill (Schmidt & Lee, 2005; Maas et al., 2008).

Speech is a fine motor skill, and it is reasonable to suggest that parallels may exist between the extensively researched principles of limb motor learning (see Schmidt & Lee, 2005) and speech motor learning (Maas et al., 2008). Developmental trends in speech motor control, including findings that young children demonstrate slower, more variable speech movement sequences than older children and young adults (e.g., Sadagopan & Smith, 2008), indicate the potential role of practice (among other factors) in the development of a highly coordinated, mature speech motor system. Behavioral and kinematic changes in speech production associated with practice have been documented for healthy adults. Schulz, Stein, & Micallef (2001) found that both healthy younger and
older adults demonstrated speech motor learning during a novel utterance production task. With practice, participants produced the utterances (e.g. “This is a thraimpoframodis”) more accurately and with reductions in duration and duration variability.

Walsh, Smith & Weber-Fox (2006) found short-term practice effects in movement coordination and duration for young children during practice of novel nonwords during a single experimental session. Sasisiekaran, Smith, Sadagopan, & Weber-Fox (2010) found similar within-session practice effects on movement coordination during nonword repetition in children and adults. They also found improvement in movement coordination and behavioral accuracy at next-day retention. These data suggest that practice has measureable effects on speech motor control both in the short-term and across longer periods of time.

1.2 Neurophysiological mediation of motor learning

Motor learning is neurophysiologically mediated: short-term and long-term changes in coordination are thought to reflect transient and persistent changes in cortical and subcortical motor maps (Kleim, Barbay & Nudo, 1998). Experience, including training, practice, environmental cues, disease and injury (Robbins et al., 2008) can alter the functioning of the brain.

Conversely, changes in the brain can affect motor learning. Schulz, Sulc, Leon & Gilligan (2000) compared speech motor learning in three healthy adults and three individuals with idiopathic Parkinson’s disease (PD). The study used a novel nonword as the practice target. Healthy elderly adults showed evidence of motor learning; there was a
negative correlation between the number of correct syllables produced and utterance duration. Adults with PD, however, did not show evidence of motor learning. For that group, correlations between accuracy and duration, and accuracy and variability, and duration and variability were neither strong nor significant. Further, participants with PD increased accuracy at the expense of duration and variability. Schulz posits that the basal ganglia— the site of brain deterioration in PD— is involved in the acquisition of skilled speech motor patterns.

Sasisekaran et al. (2010) describe a possible mechanism underlying motor learning. According to the neuromotor noise hypothesis, changes in motor learning are the consequences of reducing variability in neural command signals. The acquisition of a skill is associated with increased neural noise. With maturation and practice comes an increase in neuronal synchronization and reduced noise levels, which in turn leads to greater motor coordination.

Several general principles governing neural plasticity have been enumerated by the research community (Kleim & Jones, 2008; Ludlow et al, 2008; Robbins et al., 2008). Research has shown that use of a behavior is imperative in maintaining the neural connections necessary for its continued function. This is known as “use it or lose it.” Furthermore, use leads to an improvement in that function, in other words, “use it and improve it.” Other factors like age, salience of the learning, and the repetition and intensity of training are also known to affect learning.

The general principles of neural plasticity can be applied to motor learning. Practice can be thought of a physical embodiment of the principle of “use it and improve it”. It is widely accepted in the literature motor learning occurs as a result of practice and
experience. The purpose of the current research is to elucidate what types of practice schedules are most beneficial to acquisition and retention in the context of speech motor skill learning.

1.3 Practice schedules and motor learning

Research on motor learning describes two main types of practice: blocked and random. In blocked practice the learner practices a discrete skill multiple times before a new skill is targeted, and that skill in turn is practiced repeatedly. In random practice the order of targeted skills is random and is therefore unknown to the learner. The majority of research on the role of practice schedules on acquisition and retention of motor skills has been conducted on limb movements. This literature is reviewed first in the following section, followed by an overview of research on speech motor learning.

1.3.1 Limb motor learning

Most research on the effect of practice schedules on learning comes from limb motor research. Shea and Kohl (1990) explain two theories of how practice impacts motor learning. The first is the specificity of learning principle, which states that motor skills are specific and only superficially resemble other similar skills. Any change in a motor task (e.g., amplitude, force) results in a new motor task, for which a new motor pattern must be developed. A second theory, the variability of practice hypothesis, is based on Schmidt's (1975) schema theory. The variability of practice hypothesis states that a learner abstracts sensory information about task variations to form a rule or schema that relates to the
sensory consequences of the outcome of the movement. The specificity hypothesis predicts that blocked practice would lead to the best retention of a skill as long as the outcome measure is identical to the practiced task. In contrast, the variability of practice hypothesis predicts that varied or random practice would result in the best retention and transfer because random practice creates a stronger schema for a given motor pattern.

Another factor that may modulate the strength of a motor schema is the level of contextual interference inherent in the practice (Battig & Shea, 1980). Lee, Wulf and Schmidt (1992) define contextual interference as the influence on performance when more than one task is practiced during an experimental session. By this definition, contextual interference arises as a result of practice schedule, with blocked practice having low interference and random practice having high interference. Interference has been found to degrade performance during acquisition practice, but to enhance learning when performance is measured on retention and transfer tests (see Magill & Hall, 1990, for a review).

Shea and Morgan (1979) examined the effect of contextual interference in a motor learning paradigm by comparing groups of random and blocked practice at acquisition, retention and transfer. The groups practiced three tasks, either in blocked or random order. Participants were asked to perform the tasks as quickly as possible. The authors found that the blocked group performed more quickly during acquisition trials than the random group. The two groups differed even during the first trials of practice. Since they were repeating the same movements over and over, the blocked group was able to plan the sequence of motor skills needed to accomplish the task. They were able to perform the tasks more quickly than the random group. In contrast, the random group was not able to
plan their movements ahead of time because they did not know which of the three tasks they would be asked to complete. Although the groups differed significantly overall, their performance converged over the acquisition trials. The random group started at a higher total time than the blocked group but had a steeper slope of improvement over the first few trials.

Shea and Morgan (1979) had four counterbalanced retention trials. Half of the participants completed retention using the same practice schedule as acquisition and the other half switched schedules. The authors found that regardless of retention testing (blocked or random), participants performed better if they had practiced under random conditions. They conclude that although random practice was not beneficial during acquisition, it facilitated retention.

There are at least two plausible explanations why high interference (random practice) results in improved retention. In their paper, Shea and Morgan (1979) suggest that for the random group, high levels of interference caused a decreased dependence on the memory of the contextual factors present during the acquisition trials. The random group had practiced under myriad contextual conditions and therefore had more elaborate and flexible mental representation of the task than the blocked group. The action plans responsible for initiating and controlling the movements for successful completion of the task were accessed more efficiently in the random practice group.

Lee, Wishart, Cunningham and Carnahan (1997) put forward an alternative “forgetting and reconstruction hypothesis”. During practice, an action plan for a motor sequence is held in working memory long enough to be carried out. In random practice this action plan must be abandoned to perform a different task on the next trial. A different
action plan is then called into working memory for that task. After each trial the action plan is forgotten and is not brought back to working memory (reconstructed) until it is needed. The forgetting and reconstructing of action plans makes random practice difficult during acquisition, but results in stronger memory traces for the action plans later at retention.

Lee et al. (1997) suggest that anything that keeps the action plan in working memory (e.g., a model of the target skill) obviates the need to generate an action plan. In blocked practice, the action plan may only be created once because successive trials do not require new plans. Providing a model of the target task can have the same effect, because it places the action plan in the participant's mind without any effort (reconstruction) on their part. To test whether the presence of a model before trials would result in similar patterns of performance as blocked practice, Lee et al. (1997) compared three groups (blocked, random and random +) on a computerized timing task. Participants practiced three patterns of keystrokes. Each pattern had an associated goal time, which could be under- or over-shot. In other words, the goal was to complete the keystroke pattern in a given time, not to complete it as quickly as possible. The blocked group performed practice in blocks of 30 trials of each pattern before practicing the next pattern. The random groups practiced the patterns in a random order. Participants in the random + group received an auditory-visual model of the target pattern with the appropriate timing three times before each trial. All participants performed identical retention tests, at three minutes post- and then again at one day post-acquisition. Results demonstrated that the random + and blocked practice groups were more accurate than the random practice group at acquisition. The random group outperformed the other two groups at retention. The random + group essentially behaved the same as the blocked group. Consistent with their hypothesis, the authors
found that the presence of the model was sufficient to eliminate the effect of contextual interference due to random practice.

It remains to be seen whether contextual interference results in more elaborate action plans, or causes the action plans to be forgotten and reconstructed (or, perhaps both). Either way, it seems that random practice requires additional processing over and above that required by blocked practice. Multiple studies of limb motor learning have replicated the Shea and Morgan (1979) results (Magill & Hall, 1990) both in laboratory and real-life sports tasks. Schmidt and Lee (2005) conclude that in general, random practice has almost always been found to be as good as, if not better than, blocked practice for learning, as measured by retention and transfer.

1.3.2 Speech motor learning

In comparison to the body of research on limb motor learning, relatively little is known about the impact of practice schedules on speech motor learning. A literature search revealed only two relevant published articles, described below. What little evidence exists seems to support the notion that there are similar underlying processes for limb motor learning and speech motor learning. If true, practice schedules that benefit limb motor learning could be applied to speech.

One study by Knock, Ballard, Robin and Schmidt (2000) assessed the effects of random and blocked practice on relearning in adults with acquired apraxia of speech. Two participants were enrolled in this repeated baseline study. Practice consisted of eliciting targeted speech behaviors (consonant/vowel patterns and words) either under random or blocked order stimulus presentation, a total of 50 times each. Phonemic accuracy was
assessed during acquisition and later during retention testing. An interesting dissociation was revealed in their analysis. Targets practiced in blocked order were acquired more accurately than those practiced in random order. However, during retention testing at 1- and 4- weeks post-acquisition, both participants showed more improvement in speech production of targets that had been practiced in random order. In fact, targets practiced in blocked order were not retained at all. The authors did not find that random practice led to slower acquisition; both behavior sets were acquired at the same rate.

A study by Adams and Page (2000) examined random and blocked practice of a slowed rate of speech in unimpaired female speakers. The authors measured the absolute error (the difference between the target speech duration and the subjects’ attempted speech duration). It was hypothesized that the blocked practice group would outperform the random practice group during acquisition. While this pattern did emerge as a trend, the two groups did not differ significantly in terms of absolute error. However, during Day 2 of the experiment (retention testing), participants who had undergone random practice had significantly lower absolute error than the blocked practice group on the speech production task.

Knock et al. (2000) suggest that random practice is more difficult than blocked practice because it forces the learner to retrieve and organize a different response on each trial. The authors propose that the increased difficulty of random practice results in greater retention than blocked practice. This is similar to Lee et al.’s (1997) forgetting and reconstructing hypothesis.

The results of these studies are consistent with results from the limb motor literature. This suggests that despite different enervation patterns (see Chu & Barlow, 2000, for a
review) there may be common mechanisms underlying both speech and limb motor learning.

1.4 Rationale for the current study and hypotheses

There is a need for more normative data (i.e., regarding healthy adults) in the area of speech motor learning. Thus far, no studies have compared blocked and random practice schedules during the acquisition and retention of novel speech sequences (nonwords). The current study seeks to fill this gap in the literature. Understanding the differential effects of blocked and random practice on acquisition and retention will expand what is known of the effect of practice schedules on behavioral plasticity.

As mentioned earlier, motor learning can be assessed during acquisition and at retention. Retention performance is considered a more reliable measure of motor learning than acquisition (Knock, et al., 2000; Schmidt, 1975), in that it represents a more permanent change in the capability of skilled response. However, it is generally of interest to characterize both short- and longer-term changes. The current study was focused on examining the effects of blocked vs. random practice schedules on both the acquisition and retention of novel speech sequences (nonwords) during a production task. Nonwords were chosen as the experimental stimuli because in order to assess acquisition and retention, it is necessary to present stimuli with which the participants have no familiarity. Accuracy, duration and duration variability of nonword production were chosen as dependent variables based on previous findings that acquisition and retention of novel motor sequences are associated with improved accuracy, shorter durations movement and
decreased variability (in terms of duration variability and relative duration variability) with increased practice (Newell, 1976; Schulz, Stein, & Micallef, 2001; Walsh, Smith, & Weber-Fox, 2006; Sasisekaran, Smith, Sadagopan, & Weber-Fox, 2010), and longer-term maintenance of improvements, evidencing retention. These changes are expected within (acquisition) and across (retention) experimental sessions.

1.4.1 Hypothesis

If practice schedule does not impact acquisition and retention there should be no differences between the random and blocked practice groups on accuracy, duration, duration variability and relative duration variability at acquisition or retention. However, it is expected that at acquisition the blocked practice group will produce the nonwords more accurately, with shorter durations and with less duration variability and relative duration variability compared to the random practice group. At retention testing, it is expected that the random group will demonstrate more accurate productions, with shorter durations and less variability compared to the blocked practice group. These hypotheses are consistent with the findings of previous studies manipulating practice schedules (Shea & Morgan, 1979; Lee, Wishart, Cunningham & Carnahan, 1997; Lee, Wulf & Schmidt, 1992; Adams & Page, 2000; Knock & Ballard, 2000).
2 Methods

2.1 Participants

A total of 20 adults (10 males and 10 females), who were monolingual English speakers, ages 18-25 years (M = 21.5; SD= 2.4) were enrolled in the study. Initial eligibility for participation in the study was determined through a brief screening form. Participants with a history of speech, language or neurological impairments were excluded from participation because these impairments can alter brain function. Bilingual speakers were excluded from this study because knowledge of the phonological systems of other languages could alter the perception and production of the nonword stimuli used in the experiment. All participants were required to have at least a high school degree because the experimental task involved listening to and repeating fairly complex nonwords. Participants taking medication that could impact performance (e.g., for ADHD, depression, anxiety) were excluded from the study. Participants with orofacial piercings (lip and tongue) were eligible to participate if piercings were removed during the experiment.

Participants who passed the initial screening were asked to: a) abstain from drugs and alcohol for 24 hours before the study, as these substances can disrupt normal speech and cognitive function (McKinney & Coyle, 2004); b) avoid exposure to loud noise in the 24-hour period before the experiment, because extremely loud noise (greater than 90 dB) can cause temporary shifts in hearing levels and may cause changes in the perception of the novel word stimuli (Danenberg, Loos-Cosgrove, & LoVerde, 1987); and c) get what they considered a good night’s rest on the nights before the study since sleep deprivation has been shown to negatively affect cognitive variables such as attention and memory (Babkoff,
Zukerman, Fostick, & Ben-Artzi, 2005; Harrison & Horne, 1999), as well as motor memory consolidation (e.g., Walker et al., 2002). In addition, because the experiment required reading words and recognizing images from a distance of about 10 feet, participants were asked to wear eyeglasses or contacts that they would normally wear for distance vision.

On Day 1 of the experiment, participants completed a checklist to verify that they had followed all instructions. Participants then completed a series of brief screenings: a) a hearing screening at 25 dB at 500, 1000, 2000, 4000 and 6000Hz in each ear; b) an oral mechanism screening; and c) a vision screening using a standard eye chart. Participants who had followed all instructions and passed the screenings were determined eligible to participate in the experiment.

2.2 Testing

Participants completed tests of cognition and language: the vocabulary, digit-span and digit-symbol coding subtests of the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-3; Wechsler, 1997), the Wisconsin Card Sorting Task-4 (WCST-4; Psychological Assessment Resources, 2003), over the course of the experimental sessions. Additionally, a brief nonword-repetition test (Nonword Repetition Test; Dollaghan & Campbell, 1998) was completed on the first day of the experiment. Pre-recorded stimuli were auditorily presented and participants were instructed to repeat each nonword as accurately as they could. This test was used to examine how accurately participants could produce novel nonwords and is believed to reflect a participant’s phonological working memory function. Mean total accuracy for the nonword repetition test was 95.5% (SD = 3.29). Nonword repetition scores did not differ between participants in the blocked and random groups (p
2.3 Apparatus

Participants were seated in a chair about eight feet in front of the Optotrak camera system (Northern Digital Inc., Waterloo, Ontario, Canada). The Optotrak is a commercially-available motion tracking instrument for the study of human movement. Situated on a shelf above the Optotrak, was a 24-inch flat panel computer screen (Dell G2410) used to present visual cues for the elicitation of the novel nonword stimuli used in the experiment. All pre-recorded audio stimuli were presented though computer speakers (Dell AX210) attached to the wall behind the participants.

Participants wore Plexiglas goggles with plastic extensions down each side of the goggles. Eight infra-red light emitting diodes (IREDs) covered in protected plastic casing were attached to the participants’ face as follows: 1st on the midline of the forehead; 2nd-5th on the goggle splints (two at the level of the corners of the eye and two at the level of the corners of the lips); 6th on the midline of vermilion border of the upper lip; 7th on the midline of the lower lip; 8th on a splint attached to the jaw. The IREDs were attached by wires to a unit that was electronically interfaced to a computer. Together, the 1st through 5th IREDs served as a reference system for tracking the movements of the upper lip, lower lip and jaw, thereby eliminating artifacts due to head movement (i.e., if a participant moved their head, speech movements were still tracked with respect to the reference markers). Movements of the upper lip, lower lip and jaw IREDs, sampled at 250 Hz, were tracked in three dimensions by the Optotrak system as the participants spoke. An audio signal was recorded by a microphone placed approximately 8 cm from the participants’ mouth. This
signal was recorded on a digital recorder (Marantz PMD670) as well as digitized at a sampling rate of 16,000 Hz, on an A/D channel of the Optotrak system, allowing for synchronization of the movement data with the audio signal.

At the start of each experimental day, participants produced two sentences, “Mommy bakes pot pies” and “Buy Bobby a puppy” in natural conversational speaking manner. The purpose of this task was to habituate participants to speaking with the IREDs on their speech articulators. Each sentence was elicited using a visual cue (a toy puppy and a pie pan, respectively) and participants produced several repetitions of each sentence before the experiment was initiated.

2.4 Stimuli

Stimuli were four multisyllabic nonwords, composed of allowable combinations of phonemes in American English. Each nonword was associated with a closed-line, black and white, visual referent (see Table 2.1). The novel nonwords used in the present study increased in length and complexity (number of syllables, stimulus duration, phonotactic probability, and syllable structure; see Table 2.2). Visual referents for simpler words corresponded to simpler shapes and longer, more complicated words were associated with more complicated shapes.
Table 2.1 Experimental nonwords and accompanying visual stimuli

<table>
<thead>
<tr>
<th>Nonword</th>
<th>Visual stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>mabthroib</td>
<td><img src="image1" alt="Visual Stimulus" /></td>
</tr>
<tr>
<td>mabshroizub</td>
<td><img src="image2" alt="Visual Stimulus" /></td>
</tr>
<tr>
<td>mabshaytaidoib</td>
<td><img src="image3" alt="Visual Stimulus" /></td>
</tr>
<tr>
<td>mabspokweeflaib</td>
<td><img src="image4" alt="Visual Stimulus" /></td>
</tr>
<tr>
<td>Duration</td>
<td>Pre-Recorded</td>
</tr>
<tr>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td>1.018</td>
<td>0015</td>
</tr>
<tr>
<td>1.016</td>
<td>0240</td>
</tr>
<tr>
<td>1.016</td>
<td>0169</td>
</tr>
<tr>
<td>1.002</td>
<td>0022</td>
</tr>
</tbody>
</table>

Values calculated using the online phonotactic probabilthe calcualitor (Vreisleh & Lu, 2004).

Table 2.2: Nonword stimuli lengths, durations and phonotactics.
Two of the experimental stimuli selected for the present study were used in previous studies of speech motor plasticity (e.g., Walsh et al., 2006; Sasisekaran et al. 2010). All words began with the syllable ‘mab’, and ended in a bilabial (/b/). This phonemic structure enabled the selection of consistent start and end points for speech movement data extraction, based on lower lip peak opening velocities during the articulatory trajectory segmentation.

All nonword stimuli models were produced by a female, native speaker of American English and were prerecorded and digitized using PRAAT software (Boersma & Weenink, 2009). Primary stress was consistently placed on the initial syllable (“mab”) of each nonword model, with secondary stress on the second and third syllables in order to reduce the number of production errors related to stress placement. The stimuli were presented through two speakers attached to the wall behind the participants. The presentation volume was set at a comfortable hearing level, roughly the same level for all participants. The duration of each stimulus model can be found in Table 2.2, above.

2.5 Protocol and Data Collection

Participants were randomly assigned to one of two experimental groups: blocked practice or random practice. The experiment took place over the course of three days. The first two experimental sessions (Day 1 and Day 2) were completed on consecutive days and the final experimental session (Day 3) was completed two days after Day 2 (e.g., Monday, Tuesday, Thursday).

On Day 1, all participants first underwent a short familiarization phase (“prepractice”), during which the nonword stimuli were presented in blocked serial order.
Participants were instructed to pay attention to the visual referents presented on the computer screen while listening to auditory models of “made-up words” that went along with each visual. All nonword responses were practiced in the carrier phrase "Say __________ again". The phrase was printed on the slide along with the visual stimulus (see Figure 2.1 for an example), in order to decrease processing demands. Participants were instructed to read each slide aloud, substituting the shape with the spoken nonword target.

![Example of a presentation slide](image)

Figure 2.1 Example of a presentation slide

Participants were given feedback about their productions during the prepractice trials. If the participants’ productions were not accurate, the experimenter directed the participant’s attention to a specific inaccurate syllable. The majority of participants achieved two consecutive correct productions within the allowed five prepractice trials. Prepractice trials were excluded on Day 2 and Day 3.
During the experimental session, participants were instructed to produce the target nonword (in the carrier phrase) associated with the visual referent presented on the computer screen in front of them. Audio models were not simultaneously presented with the visual referent for the experimental productions. Rather, the audio model was played as feedback, only after the participant made an attempt to produce the target nonword. Participants were told that the same four nonwords corresponding to the same four shapes would be used during the course of the experiment. Participants were explicitly reminded that the simpler shapes corresponded to the simpler words and that the more complicated shapes corresponded to the longer, more complicated words. The experimenter also reminded participants that all nonwords started with the same syllable and that all ended with the same sound. Participants were instructed to produce their responses as quickly and as accurately as possible.

The presentation of stimuli varied between the two groups as follows. In the blocked practice group (BP), participants produced each nonword within the carrier phrase 15 times in a row. After 15 productions, another nonword was targeted. Order of presentation of nonwords was randomized across participants. In the random practice group (RP), participants produced the nonwords within the carrier phrase in 15 blocks with each of the four nonword stimuli pseudo-randomly presented within a block. This order was consistent across RP participants. Therefore, each nonword was produced 15 times (similar to the blocked group), but in an order that was unpredictable to the participants. After 15 productions of each nonword participants were given a short break. During the break participants read a short passage aloud and were given a chance to have a
Participants completed an additional 15 practice productions within their assigned groups after the break.

Previous studies of motor speech learning have used 15 practice productions (e.g., Sasisekaran et al., 2010). However, these earlier studies utilized nonword repetition paradigms, i.e., participants repeated nonwords after an auditory model. Because the current study imposed the additional challenge of nonword recall based on association with a visual stimulus, an additional set of 15 practice productions were completed as part of acquisition on Day 1 in order to document practice effects over a more extended period of practice.

On Day 2 of the experiment, all participants, regardless of group first completed a serial retention task. During the retention task, the nonword visual cues were presented in serial order from easiest to most complex. Ten blocks of all four nonword visual cues were presented, and no audio models were presented with visual stimuli or after participants’ responses. Following the serial retention task, participants completed additional practice and a second retention test three days from the initial day of the experiment, which were not analyzed for the present study. Only data from the practice phases on Day 1 (two 15-trial practice phases separated by a 10-minute break) and the serial retention testing from Day 2 were analyzed for the current study. Table 2.3 (below) depicts the experimental protocol.
### Table 2.3 Schematic of experimental protocol

<table>
<thead>
<tr>
<th></th>
<th>Day 1</th>
<th>Day 2: next day</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-practice</strong></td>
<td>5 blocks of serial practice</td>
<td>Serial retention:</td>
</tr>
<tr>
<td><strong>Practice Phase 1:</strong></td>
<td>15 blocks with auditory model presented after production; blocked or randomized practice</td>
<td>10 blocks; no auditory model</td>
</tr>
<tr>
<td><strong>BREAK (10 minutes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Practice Phase 2:</strong></td>
<td>15 blocks with auditory model presented after production; blocked or randomized practice</td>
<td></td>
</tr>
</tbody>
</table>

2.4 Data Analysis

To characterize acquisition, measures for each dependent variable were obtained for the first five productions of Phase 1 of practice (initial acquisition trials; TIME 1), the last five productions of Phase 1 of practice (level of performance before a break; TIME 2), the first five productions of Phase 2 of practice (level of performance after the break; TIME 2), and the last five productions of Phase 2 of practice (level of performance after the break; TIME 3).
3) and the last five production of Phase 2 of practice (end of acquisition on Day 1; TIME 4). To assess retention, measures for each dependent variable were obtained for all 10 productions of each nonword during the serial retention rest on Day 2 (RETENTION).

2.4.1 Behavioral Accuracy (Percent phonemes correct)

The experimenter phonemically transcribed all productions from Day 1 practice (Phase 1 and Phase 2) and Day 2 serial retention for each participant. Beginning at the syllable-initial boundary, phonemes of the production were matched with the target. Only accurate phonemes counted towards the total phonemic accuracy (i.e., the experimenter did not count the number or mistakes made). Dysfluent syllables and phonemes were transcribed and marked in parentheses but did not count against the phonemes correct score. The phoneme accuracy score was converted to a percent phonemes correct score for each nonword.

Inter-rater agreement for behavioral accuracy analysis for 20% of the phoneme accuracy data, was determined to be high (r = .99).

2.4.2 Kinematic Analysis (Mean nonword duration, duration variability and relative duration variability).

Kinematic analysis was completed on data from time points TIME 1, TIME 2, TIME 3 and TIME 4 (described above) for each nonword, and for the retention test on Day 2. Since inaccurate productions are part of the acquisition process, productions with errors were considered acceptable for kinematic analysis of timing measures, as long as the participant made a reasonable attempt to say the nonword (e.g., “mabshoib” instead of “mabthroib”).
Productions where the participant said only part of the nonword (e.g., “Say mab – I don’t remember the rest”) were excluded from the analysis. Self-corrections were considered acceptable for analysis (e.g., “Say mabthroid-throib again”). If the participant restarted the carrier phrase (e.g., “Say mab... Say mabthroib” again”) then only the corrected, complete nonword production was analyzed. Utterances were excluded if Optotrak data could not be accurately obtained (e.g., movement data was not synchronized with audio data [one trial of one participant = .03% of total trials]). For one participant, only the first 4 productions of the first seven (Phase 1) of nonword 2 were analyzed because he did not produce five acceptable productions.

The upper lip, lower lip, and jaw movements associated with acceptable productions were imported into MATLAB® (Mathworks, 2005) signal processing software for analysis. A custom-designed interactive program that offered a simultaneous display of the superior-inferior displacement and velocity records from the lower lip for each production was used to extract the embedded nonwords from the carrier phrase. The experimenter extracted each nonword by selecting consistent kinematic landmarks from the velocity records with a computer mouse. Starting points were chosen as the peak velocity of the opening movement for the /m/ in “mab”, and end points were selected as the peak lower lip opening velocity for the /b/ that ended the last syllable of all the nonwords. Because the speech acoustic signal was digitized at 16kHz and low-pass filtered at 7.5 kHz with an A/D unit synchronized to the Optotrak system, the experimenter was able to listen to each extracted interval to ensure that it was an acceptable production for duration analysis and that she had selected the appropriate start and end points, without inadvertently cutting off the signal. Extraction of nonwords via this method allowed for a reliable analysis of
nonword duration. The MATLAB software automatically computed mean durations and duration variability (standard deviation) of the selected extractions for each time point of analysis (TIMEs 1, 2, 3, 4 and retention).

Newell (1976) found that duration variability is related to movement duration; longer movement durations have greater variability (i.e., greater standard deviations). It was expected that longer nonwords would have greater durations and therefore greater duration variability. To minimize this confound, relative variability was calculated by dividing the standard deviation of durations (duration variability) with the mean of the durations of five productions of each nonword.

2.5 Statistical Analysis

2.5.1 Acquisition

To examine differences in performance between the blocked and random practice groups during acquisition of nonwords on Day 1, a 2 (GROUP, blocked vs. random) by 4 (NONWORD, nonword stimuli) by 4 (TIME, performance testing at 4 time points) repeated measures analyses of variance (ANOVA) was run. Separate ANOVAs were run for each dependent variable and are reported separately, below. Behavioral accuracy data (% phonemes correct) were arcsine transformed using a freely available MATLAB ® program that applies a rationalized arcsine transformation (Studebaker, 1985) before statistical analysis.

Planned comparisons were completed to assess change in performance as a function of group (GROUP x TIME effects) from: a) TIME 1 to TIME 2 (early phase of acquisition); b)
TIME 2 to 3 (performance before and after the break); and c) TIME 1 to TIME 4 (change from beginning to end of practice), in order to determine if changes in performance over time during the acquisition phase was different for the two groups. Planned comparisons were also completed separately for blocked and random groups (TIME 1 to TIME 2; TIME 2 to TIME 3; TIME 3 to TIME 4; and TIME 1 to TIME 4) in order to describe and characterize any group differences in performance change over time.

2.5.2 Retention

Repeated measures ANOVAs were computed with the between subject- variable GROUP (blocked vs. random) and within-subject variable, NONWORD (4 nonwords) at retention testing. Separate ANOVAs were run for each dependent variable and results are reported below.

Post-hoc tests (Tukey –Kramer’s multiple comparisons) were used to describe all effects of NONWORD on dependent variables. Alpha was set to .05 for all statistical tests. All p values reported for the ANOVA are Greenhouse-Geisser corrected.
3 Results

3.1 Acquisition

**Accuracy.** The overall ANOVA revealed a main effect of GROUP, $F(1,18) = 11.64, p < .01, \eta_p^2 = .393$. There were also significant main effects of NONWORD, $F(3,54) = 5.80, p < .01, \eta_p^2 = .24$, and of TIME, $F(3,54) = 63.49, p < .0001, \eta_p^2 = .78$ and a significant TIME x GROUP interaction, $F(3, 54) = 15.43, p < .0001, \eta_p^2 = .46$. The NONWORD x GROUP ($p = .08$), NONWORD x TIME ($p = .21$), and NONWORD x TIME x GROUP ($p = .79$) interactions were not significant.

Post-hoc comparisons for NONWORD (Figure 3.1) revealed that nonword 1 (“mabthroib”) was produced more accurately than nonwords 2 (“mabshroizub”) and 3 (“mabspokweelaib”). Accuracy of nonwords 1 and 4, 2 and 3, 2 and 4, and 3 and 4 did not differ from one another (Tukey-Kramer Multiple Comparisons, $p < .05$).
As evident in Figure 3.2, the BP group demonstrated higher accuracy scores overall during acquisition, $F(1,18) = 11.64, p < .01, \, \eta_p^2 = .39$. Even at TIME 1 (the first 5 trials of practice), the accuracy scores for the BP group were already significantly higher compared to the RP group, $F(1,18) = 23.69, p < .0001, \eta_p^2 = .56$. At the end of acquisition (TIME 4), the BP continued to demonstrate higher accuracy scores than the RP group, $F(1,18) = 8.89, p < .01, \eta_p^2 = .33$. 
In order to test the hypothesis that the two groups differed in performance during acquisition, planned comparisons were used to examine changes in accuracy over time, by group. It was found that the change in accuracy during the first phase of practice (from TIME 1 to TIME 2) did not vary by GROUP ($p = 0.08$). Similarly, change in accuracy from TIME 1 (beginning of acquisition) to TIME 4 (end of practice on Day 1) did not vary by GROUP ($p = .10$). In other words, the groups did not differ in the magnitude of accuracy gains during the first phase of practice and overall across practice on Day 1 (acquisition). However, the shape of the trajectory of acquisition differed between groups since the
differences from TIME 2 to TIME 3, $F(1,18) = 58.88$, $p < .0001$, $\eta^2 = .76$, and TIME 3 to TIME 4, $F(1,18) = 9.73$, $p < .01$, $\eta^2 = .35$, varied significantly by GROUP.

To further characterize the differences in change in accuracy as a function of practice schedule, planned comparisons were run for each group separately to examine changes across specific time points (1 to 2, 2 to 3, 3 to 4 and 1 to 4). Results from these planned comparisons for accuracy are presented in Table 3.1 and are graphically represented in Figure 3.3 for the BP and RP groups, separately.

Figure 3.3 Phoneme accuracy at acquisition time points for the blocked and random practice groups. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.
Taken together, these findings suggest that although the BP group demonstrated higher accuracy overall during acquisition, the groups did not differ significantly in gains made during phase 1 of practice, and across acquisition on Day 1. The trajectories of change in accuracy over time, however, differed between the two groups, such that the BP group demonstrated a significant decrease in accuracy following a short break, while the RP group continued to improve. Further, during the final phase of practice on Day 1, both groups demonstrated improvements in accuracy, but this change was found to be significant only for the BP group.

Table 3.1: Planned comparisons results for accuracy over time for both practice groups

<table>
<thead>
<tr>
<th>Time points</th>
<th>df</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blocked Practice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>(1,9)</td>
<td>228.23</td>
<td>0.00*</td>
<td>0.96</td>
</tr>
<tr>
<td>2-3</td>
<td>(1,9)</td>
<td>95.70</td>
<td>0.00*</td>
<td>0.91</td>
</tr>
<tr>
<td>3-4</td>
<td>(1,9)</td>
<td>98.10</td>
<td>0.00*</td>
<td>0.92</td>
</tr>
<tr>
<td>1-4</td>
<td>(1,9)</td>
<td>211.65</td>
<td>0.00*</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>Random Practice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>(1,9)</td>
<td>30.61</td>
<td>0.00*</td>
<td>0.77</td>
</tr>
<tr>
<td>2-3</td>
<td>(1,9)</td>
<td>8.03</td>
<td>0.02*</td>
<td>0.47</td>
</tr>
<tr>
<td>3-4</td>
<td>(1,9)</td>
<td>2.49</td>
<td>0.15</td>
<td>0.22</td>
</tr>
<tr>
<td>1-4</td>
<td>(1,9)</td>
<td>59.94</td>
<td>0.00*</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Duration. The overall ANOVA revealed a main effect of GROUP, $F(1, 18) = 39.48$, $p < .0001$, $\eta^2_p = .69$. There were also significant main effects of NONWORD, $F(3, 54) = 48.79$, $p < .0001$, $\eta^2_p = .73$, and of TIME, $F(3, 54) = 22.03$, $p < .0001$, $\eta^2_p = .55$. There was a significant GROUP x NONWORD interaction, $F(3, 54) = 6.39$, $p < .001$, $\eta^2_p = .26$ Tukey-Kramer’s Multiple Comparisons revealed that the BP group produced nonwords 2, 3, and 4 significantly more quickly than the RP group ($p < .05$; see Figure 3.4, below). There were no significant interactions between NONWORD x TIME ($p = .22$), or between NONWORD x GROUP x TIME ($p = .55$).

![Figure 3.4 Nonword durations during acquisition. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.](image)

The ANOVA for duration revealed a significant GROUP x TIME interaction, $F(3, 54) = 8.03$, $p < .0001$, $\eta^2_p = .31$. Figure 3.5 shows duration trajectories for both practice groups. Overall, the BP group had shorter nonword durations than the RP group. This difference
between the BP and RP groups was significant even at TIME 1 (beginning of acquisition), $F(1,18) = 27.71, p < .0001, \eta_p^2 = .61$. Planned comparisons revealed that the change in nonword duration from TIME 1 to TIME 2 did not vary significantly by GROUP ($p = .44$). However, overall changes from the beginning of acquisition (TIME 1) to the end of acquisition on Day 1 (TIME 4), $F(1,18) = 8.91, p < .01, \eta_p^2 = .33$, and from TIMEs 2 to 3, $F(1,18) = 13.22, p < .01, \eta_p^2 = .42$, depended on group. The difference between TIMEs 3 to 4 did not depend on GROUP ($p = .72$), as the groups made similar decreases in duration during the final phase of practice.

![Figure 3.5 Duration trajectories for both practice groups at acquisition time points (1 – 4) and retention testing (R).](image)

Figure 3.5 Duration trajectories for both practice groups at acquisition time points (1 – 4) and retention testing (R). Error bars reflect the standard error of the mean. Significant effects are marked with the * symbol.

To characterize how the changes in duration over time depend on group, planned comparisons were run for each group separately to examine changes across specific time
points (1 to 2, 2 to 3, 3 to 4 and 1 to 4). Results from these planned comparisons for duration are presented in Table 3.2 and are graphically represented in Figure 3.6.

![Figure 3.6 Duration at acquisition time points for the blocked and random practice groups. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.](image)

Taken together, these results suggest that reductions in nonword duration during the initial (TIME1 to TIME 2), and final practice phases (TIME 3 to TIME 4) did not differ significantly between the groups. Across these time points the groups made similar magnitude decreases in duration. After the break, however, the random group demonstrated a significant decrease in duration, while the blocked group demonstrated a non-significant increase in duration. Overall, while both groups became significantly faster from TIMEs 1 to 4 over the course of acquisition, the RP group demonstrated a greater
magnitude of decrease in duration than the BP group. At the end of practice on Day 1, the BP produced significantly shorter nonword durations compared to the RP group, \( F(1,18) = 17.51, p < .001, \eta_p^2 = .49 \).

Table 3.2 Planned comparisons results for duration for both practice groups

<table>
<thead>
<tr>
<th>Time points</th>
<th>df</th>
<th>F</th>
<th>P</th>
<th>( \eta_p^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked Practice</td>
<td>1-2</td>
<td>(1,9)</td>
<td>29.17</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>(1,9)</td>
<td>4.25</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>(1,9)</td>
<td>10.88</td>
<td>0.01*</td>
</tr>
<tr>
<td></td>
<td>1-4</td>
<td>(1,9)</td>
<td>57.93</td>
<td>0.00*</td>
</tr>
<tr>
<td>Random Practice</td>
<td>1-2</td>
<td>(1,9)</td>
<td>3.55</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>(1,9)</td>
<td>9.08</td>
<td>0.02*</td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>(1,9)</td>
<td>1.86</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>1-4</td>
<td>(1,9)</td>
<td>34.46</td>
<td>0.00*</td>
</tr>
</tbody>
</table>

Duration Variability (Standard Deviation of Durations). The overall ANOVA revealed a main effect of GROUP, \( F(1,18) = 20.00, p < .0001, \eta_p^2 = .53 \). There were also significant main effects of NONWORD, \( F(3,54) = 7.10, p < .0001, \eta_p^2 = .28 \), and of TIME, \( F(3,54) = 10.17, p < .0001, \eta_p^2 = .36 \). There was a significant GROUP x NONWORD interaction, \( F(3,54) = 6.22, p < .01, \eta_p^2 = .26 \) (see Figure 3.7, below). Tukey-Kramer’s Multiple Comparisons revealed that the RP group was significantly more variable in their production durations of
nonwords 2 and 3 than the BP group \((p < .05)\). Duration variability did not differ between groups for nonwords 1 or 4.

Figure 3.7 Nonword duration variability for both practice groups during acquisition. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.

A significant GROUP x TIME interaction was found, \(F(3,54) = 9.13, \ p < .0001, \eta_\text{p}^2 = .34\). As can be seen in Figure 3.8, below, the BP group exhibited lower duration variability compared to the RP group at the beginning (TIME 1), \(F(1,18) = 13.59, \ p < .01, \eta_\text{p}^2 = .43\), and at the end (TIME4), \(F(1,18) = 13.72, \ p < .01, \eta_\text{p}^2 = .43\) of acquisition. Planned comparisons were used to examine the GROUP x TIME differences in duration variability. It was found that change in duration variability from TIMEs 1 to 2, \(F(1,18) = 10.98, \ p < .01, \eta_\text{p}^2 = .38\) and
from TIMEs 2 to 3, $F(1,18) = 24.60, p < .0001, \eta_p^2 = .57$, differed significantly by GROUP.

The change from TIMEs 3 to 4 ($p = .19$) did not differ significantly by GROUP. The change in duration variability from the beginning to the end of acquisition (TIMEs 1 to 4) did not differ significantly by GROUP ($p = .33$), as the groups made decreases of similar magnitude.

![Figure 3.8 Trajectory of duration variability over time](image)

Figure 3.8 Trajectory of duration variability at acquisition time points (1 – 4) and retention testing (R) for both practice groups. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.

To characterize how the changes in duration variability over time differed based on practice schedules, planned comparisons were run for each group separately across specific time points (1 to 2, 2 to 3, 3 to 4 and 1 to 4). Results from these planned
comparisons for duration variability are reported in Table 3.3 and are represented in Figure 3.9.

Taken together, these results suggest different patterns in changes in duration variability for the two groups over time. From TIMEs 1 to 2 the BP group demonstrated a significantly greater decrease in duration variability than the RP group. After the break (TIMEs 2 to 3) the RP group demonstrated a significant decrease in duration variability while the BP group demonstrated a significant increase. The change in duration variability from TIMEs 3 to 4 and from TIMEs 1 to 4 did not differ significantly by group.

Figure 3.9 Duration variability at acquisition time points for the blocked and random practice groups. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.
Table 3.3 Planned comparisons results for duration variability for both practice groups

<table>
<thead>
<tr>
<th>Time points</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked Practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>(1,9)</td>
<td>26.95</td>
<td>0.00*</td>
<td>0.75</td>
</tr>
<tr>
<td>2-3</td>
<td>(1,9)</td>
<td>7.88</td>
<td>0.02*</td>
<td>0.47</td>
</tr>
<tr>
<td>3-4</td>
<td>(1,9)</td>
<td>7.15</td>
<td>0.03*</td>
<td>0.44</td>
</tr>
<tr>
<td>1-4</td>
<td>(1,9)</td>
<td>23.02</td>
<td>0.00*</td>
<td>0.72</td>
</tr>
<tr>
<td>Random Practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>(1,9)</td>
<td>.05</td>
<td>.84</td>
<td>0.01</td>
</tr>
<tr>
<td>2-3</td>
<td>(1,9)</td>
<td>18.61</td>
<td>0.00*</td>
<td>0.67</td>
</tr>
<tr>
<td>3-4</td>
<td>(1,9)</td>
<td>.18</td>
<td>.68</td>
<td>0.02</td>
</tr>
<tr>
<td>1-4</td>
<td>(1,9)</td>
<td>12.04</td>
<td>0.01*</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Relative Variability. The overall ANOVA revealed a main effect of GROUP, $F(1,18) = 18.63$, $p < .0001$, $\eta_p^2 = .51$. There were also significant main effects of NONWORD, $F(3,54) = 3.00$, $p < .05$, $\eta_p^2 = .14$, and of TIME, $F(3,54) = 9.34$, $p < .0001$, $\eta_p^2 = .34$. There was a significant GROUP x NONWORD interaction, $F(3,54) = 3.19$, $p < .05$, $\eta_p^2 = .15$ (see Figure 3.10, below). Tukey-Kramer's Multiple Comparisons revealed that the RP group was significantly more variable in their production durations of nonwords 2 and 3 than the BP group ($p < .05$). Relative variability did not differ between groups for nonwords 1 or 4.
A significant GROUP x TIME interaction was found, $F(3,54) = 11.28, p < .0001, \eta^2_p = .39$. As can be seen in Figure 3.11, below, the BP group exhibited lower relative variability compared to the RP group at the beginning (TIME 1), $F(1,18) = 6.54, p < .05, \eta^2_p = .27$, and at the end (TIME4), $F(1,18) = 13.34, p < .05, \eta^2_p = .43$ of acquisition. Planned comparisons were used to examine the GROUP x TIME differences in relative variability. It was found that change in relative variability from TIMES 1 to 2, $F(1,18) = 18.16, p < .001, \eta^2_p = .50$ differed significantly by GROUP, as did the changes from TIMES 2 to 3, $F(1,18) = 30.16, p < .0001, \eta^2_p = .63$, and TIMES 3 to 4, $F(1,18) = 5.29, p < .05, \eta^2_p = .23$. The change in relative variability from the beginning to the end of acquisition (TIMES 1 to 4) did not differ significantly by GROUP ($p = .71$), because the groups made decreases of similar magnitude.
To further characterize how the changes in relative variability over time differed based on practice schedules, planned comparisons were run for each group separately across specific time points (1 to 2, 2 to 3, 3 to 4 and 1 to 4). Results from these planned comparisons for relative variability are reported in Table 3.4 and are represented in Figure 3.12.

Taken together, these results suggest different patterns in changes in relative variability for the two groups. The BP group demonstrated lower relative variability overall during acquisition. The BP group made greater decreases in relative variability than the RP group during phase 1 of practice (TIMEs 1 to 2). After the break (TIMEs 2 to 3) the RP
group demonstrated a significant decrease in relative variability while the BP group demonstrated a significant increase. During phase 2 of practice (TIMEs 3 to 4), the BP group had a significant decrease in relative variability while the RP group did not. Although the BP group had lower relative variability than the RP group at both TIMEs 1 and 4, the groups made similar decreases in relative variability from the beginning to the end of acquisition.

Figure 3.12 Relative variability at acquisition time points for the blocked and random practice groups. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.
Table 3.4 Planned comparisons results for relative variability for both practice groups

<table>
<thead>
<tr>
<th>Time points</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>(\eta^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked Practice</td>
<td>1-2</td>
<td>(1,9)</td>
<td>37.65</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>(1,9)</td>
<td>12.51</td>
<td>0.01*</td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>(1,9)</td>
<td>8.71</td>
<td>0.02*</td>
</tr>
<tr>
<td></td>
<td>1-4</td>
<td>(1,9)</td>
<td>19.95</td>
<td>0.00*</td>
</tr>
<tr>
<td>Random Practice</td>
<td>1-2</td>
<td>(1,9)</td>
<td>0.28</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>(1,9)</td>
<td>20.84</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>(1,9)</td>
<td>0.00</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>1-4</td>
<td>(1,9)</td>
<td>7.87</td>
<td>0.02*</td>
</tr>
</tbody>
</table>

3.2 Retention

Repeated measures ANOVAs were computed to examine differences between GROUP (Blocked vs. Random) and NONWORD (4 nonwords) at retention testing. Separate ANOVAs were run for each dependent variable and results are reported for each variable below. All p values reported are Greenhouse-Geisser corrected.

Accuracy. The ANOVA revealed a main effect of GROUP, \(F(1,18) = 6.68, p < .05, \eta^2 = .27\). The RP group was significantly more accurate at retention than the BP group (see Figure 3.13, below). There was also significant main effect of NONWORD, \(F(3,54) = 5.57, p < .01\), \(\eta^2 = .23\). The NONWORD x GROUP interaction was not significant (\(p = .43\)).
Tukey-Kramer’s Multiple Comparisons for NONWORD revealed that Nonword 1 (“mabthroit”) was produced more accurately than nonword 2 (“mabshroizub”). Nonword 1 was produced as accurately as nonwords 3 (“mabshaytaidoib”) and 4 (“mabspokweflaib”). Accuracy of nonwords 2, 3 and 4 did not differ from one another (see Figure 3.14, below).
**Duration.** The ANOVA revealed no main effect of GROUP ($p = .13$) on duration (see Figure 3.15, below). Although the BP group demonstrated longer durations than the RP group at retention, this difference was not significant. There was a significant main effect of NONWORD, $F(3,54) = 14.12, p < .0001, \eta^2_p = .44$. Tukey-Kramer's Multiple Comparisons revealed that at retention, nonword 1 was produced more quickly than all other nonwords ($p < .05$). Nonwords 2, 3 and 4 did not differ in duration (see Figure 3.16, below). The NONWORD x GROUP interaction was not significant ($p = .87$).
Figure 3.15 Nonword duration at retention testing for both practice groups. Error bars reflect the standard error of the mean.

Figure 3.16 Nonword duration at retention testing. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.
Duration Variability (Standard Deviation of Durations). There was no significant main effect of GROUP on duration variability at retention ($p = .23$; see Figure 3.17, below). The ANOVA revealed a main effect of NONWORD on duration variability, $F(3,54) = 8.02$, $p < .005$, $\eta_p^2 = .30$. Tukey-Kramer’s Multiple Comparisons revealed that nonword 1 durations were less variable than those of nonwords 2 and 3 ($p < .05$). Nonword 1 did not differ from nonword 4. Nonwords 2, 3 and 4 did not differ in duration variability (see Figure 3.18, below). The NONWORD x GROUP interaction was not significant.

Figure 3.17 Nonword duration variability at retention testing. Error bars reflect the standard error of the mean.
Relative Variability. There was no significant main effect of GROUP on relative variability at retention $(p = .41; \text{see Figure 3.19, below})$. The ANOVA revealed a main effect of NONWORD on relative variability, $F(3,54) = 11.46, p < .001, \eta^2_p = .39$. Tukey-Kramer's Multiple Comparisons revealed that nonwords 1 and 4 had lower relative variability than those of nonwords 2 and 3 $(p < .05)$. Nonword 1 did not differ from nonword 4. Nonwords 2 and 3 did not differ in relative variability (see Figure 3.20, below). The NONWORD x GROUP interaction was not significant.
Figure 3.19 Nonword relative variability at retention testing. Error bars reflect the standard error of the mean.

Figure 3.20 Nonword relative variability at retention testing. Significant effects are marked with the * symbol. Error bars reflect the standard error of the mean.
Discussion

This study examined the effect of two practice schedules, blocked and random, on the acquisition and retention of novel nonwords in healthy young adults. Performance was assessed over the course of acquisition and during a serial retention test the next day. Duration, duration variability, relative duration variability and behavioral accuracy were assessed. These are common measures of motor skill performance in the context of limb-motor learning (Newell, 1976; Schmidt & Lee, 2005) and speech motor learning (Schulz, Stein & Micallef, 2001; Smith & Zelaznik, 2004; Walsh, Smith & Weber-Fox, 2008; Sadagopan, 2010).

Based on a consistent pattern from the limb motor learning literature, and limited results from the speech motor learning literature, it was expected that practice schedule would have differential effects on acquisition and retention. More specifically, it was expected that blocked practice would facilitate acquisition while random practice would facilitate retention. Acquisition- and retention-specific findings are discussed below.

Acquisition

It was expected that all participants would demonstrate practice effects over the course of acquisition (i.e., they would produce the nonwords more quickly, less variably and more accurately), per previous literature on motor skill acquisition (Schmidt & Lee, 2005). This hypothesis was confirmed. For both groups, at the end of the practice trials on Day 1 (TIME 4) productions were shorter, less variable in duration and more accurate than
productions during previous time points, indicating that both groups demonstrated practice effects reflecting the acquisition of novel nonwords.

Although it was hypothesized that the BP group would demonstrate better overall performance than the RP group during acquisition, an interesting finding was that significant group differences were already present at TIME 1 (the first 5 trials of practice) of acquisition on the first day of the experiment, wherein the BP demonstrated significantly higher accuracy, shorter durations and less duration variability and relative variability than the RP group. Shea & Morgan (1979) and Lee et al. (1997) found similar patterns of group differences for movement duration at the beginning of acquisition. In the current experiment, this difference in performance between the groups remained even at the end of acquisition on Day 1. In other words, performance levels of the BP and RP groups were not equivalent at the end of practice; rather, the BP demonstrated a performance advantage throughout. It is possible that the duration, duration variability and relative variability scores for the BP group were so low during acquisition because participants were actuating a single motor plan repeatedly. Increased durations for the RP group, on the other hand, may reflect increased planning, programming and execution times associated with the generation and/or retrieval of changing action plans. The high contextual interference associated with random practice (Battig & Shea, 1980) may have resulted in the persistence of poorer performance throughout acquisition.

It is unlikely that this difference in performance at the start of acquisition (TIME 1) reflects an underlying difference in ability between groups. The fact that the two groups did not differ in nonword repetition ability as assessed by the Dollaghan Nonword Repetition
Test (Dollaghan & Campbell, 1998) supports this suggestion. Rather, this finding may be attributed to a few inherent task differences between blocked and random practice, and methodological considerations. First, ‘TIME 1’ does not represent a true “baseline” because all dependent measures obtained at this time point were for the first five productions of each nonword. It is possible that both groups demonstrated similar performances levels (in terms of accuracy and production duration) on the very first production of each nonword, but that BP participants made significantly greater gains in performance for each nonword much more quickly than the RP group, such that group differences in performance were already evident at the end of the first 5 productions of practice.

Second, the BP group’s task was, arguably, substantially easier than that of the RP group. Recall that after an attempted nonword response, both groups were provided an accurate, auditory model of the target nonword production. One possibility is that this feedback may have been more advantageous to the BP group due to the experimental design used in the present study. Given an auditory model for the correct response, the BP group may have simply repeated the nonword model they had heard through the loudspeakers without relying on an association with the visual referent. In other words, the BP group’s task may have somewhat resembled a nonword repetition task, requiring relatively reduced cognitive-linguistic effort compared to that required of RP group, and involving, primarily, phonological short-term memory processes (Gathercole, 2006).

The phonological loop of short-term memory can retain information for about two seconds before the data begins to decay (Baddeley, Gathercole, & Papagno, 1998). Since participants likely began planning and producing the next utterance less than two seconds
after they heard an accurate auditory example of the nonword, representations were presumably still available in the phonological loop. Anecdotally, it was noted that some BP participants “jumped the gun” during acquisition time points. They began producing the carrier phrase and the targeted nonword before the visual cue was presented. This indicates that they were fully cognizant of which nonword they were expected to produce, and further, that they were maintaining a representation of the target in their phonological working memory. This also suggests that the action plan needed to generate the nonword was present in working memory as well (Lee et al., 1997). On the other hand, unsure of which visual cue they would see next, the RP group had to create an association between each visual referent and its corresponding nonword, store the association in working/short-term memory, recall the association, and generate the motor command for the production for a changing target. The provision of a model during practice reduced these processing demands for the BP group, similar to the effects found in Lee et al. (1997).

The suggestion that the BP group may have simply been repeating the nonword model without forming deeper, more robust associations is further supported by the consistent pattern of decrease in performance (for all three dependent variables) from Time 2 (before a short break) to Time 3 (right after the break). The BP group returned to near-“baseline” (TIME 1) levels for accuracy and duration variability. In contrast, the RP group’s performance typically demonstrated an improvement following the break.

While the BP group, as expected, consistently demonstrated higher accuracy, shorter durations and lower variability than the RP group during acquisition, the present findings demonstrated no clear patterns on group differences in the magnitude of change in
performance over time. In general, for both groups the trajectory and magnitude of change across phase 1 of practice (TIMEs 1 to 2) was similar to the change across Day 1 (TIMEs 1 to 4), with a few exceptions. During phase 1 of practice, there were significant differences in magnitude of decreases between the two practice schedules for the duration, duration variability and relative variability measures: the BP group demonstrated a significant decrease while the random group did not. On the other hand, it was found that the RP group made significantly greater decreases in nonword production duration compared to the BP group from the start of practice to the end of acquisition. There was also a trend toward a greater improvement in accuracy during the early acquisition trials (that was nonsignificant) for the RP group compared to the BP group. Whether the change in performance during the early trials or over the course of acquisition is greater for the BP vs. the RP group seems to follow no consistent pattern in previous research. For example, Adams & Page (2000), in their study of blocked and random practice on slowed speech production, found that the blocked group made larger gains during the first two trials of practice while the random practice group made greater gains over the last three trials of practice. This finding is different than those of Shea & Morgan (1979) and Lee et al. (1997) who found that the random group made the greatest changes in early acquisition trials. The difference may have to do with the nature of the task (speech motor vs. limb motor) or it may reflect a methodological difference in the number of productions constituting a “trial”, and requires further controlled investigation.
Retention

Retention was assessed one day after practice (acquisition) was completed. Participants were required to recall and produce the four multisyllable nonwords that they had practiced the day before. Ten productions of each nonword were elicited in serial order. Not surprisingly, retention performance dropped for all outcome measures (see Figures 3.2, 3.5, 3.8 and 3.11) for both groups (compared to levels of performance at the end of acquisition on Day 1), likely due to difficulty accessing previously stored phonological-articulatory representations of nonword productions (an effect documented in children in studies of word learning; e.g., Dollaghan, 1985). It was hypothesized that the RP group would outperform the BP group at retention testing. The results for behavioral accuracy support the hypothesis: the RP group was significantly more accurate than the BP group at retention. Similar trends (although not significant) were noted for mean duration, duration variability and relative variability as well (i.e., the RP group demonstrated shorter durations and lower variability than the BP group at retention). For the RP group, accuracy, duration and variability values at retention fell within the range of performance levels at the end of acquisition (TIME 4) on Day 1. The BP group, however, was substantially less accurate, slower and more variable at retention than they had been at any time during acquisition. Overall, the patterns of results indicate a reversal in performance from acquisition to retention, depending on practice type. While BP facilitated acquisition of the novel nonwords, participants in the RP group generally did better at retention. This pattern parallels the findings from the limb motor learning literature and is expected to be related to differences in the strength (Lee et al., 1997) or the elaboration and efficiency of access (Shea & Morgan, 1979) of the motor plans required for the task.
Effects of nonword complexity

Although not the primary goal of the present study, four nonwords were included to examine if differences existed between groups in performance as a function of nonword length and complexity. Nonword length and complexity have been shown to influence speech motor learning and performance in previous studies (e.g., Walsh et al., 2006; Sasisekaran et al., 2010). In addition, increased movement duration is associated with increased duration variability (Newell, 1976). To minimize the confound between nonword length and duration variability, relative variability was calculated for productions of each nonword at acquisition and retention. Similar patterns were uncovered for duration variability and relative variability for the nonwords. At acquisition, the BP group tended to have consistent variabilities across nonwords, while the RP group demonstrated significantly higher variabilities overall, especially for nonwords 2 and 3. At retention, the groups demonstrated similar patterns of for duration and relative variabilities, with nonwords 2 and 3 being the most variable.

A noteworthy finding in this study was that changes in performance over time followed similar patterns for all nonwords for both BP and RP groups. In other words, the current findings suggest that the effect of practice schedules on nonword production were similar, overall, regardless of nonword length or complexity. However, nonword 2 ("mabshroizub") was produced less accurately than the other nonword targets. Nonwords 2 and 3 tended to have longer durations which more variable in terms of both duration variability and relative duration variability. This is surprising, since these words were neither the longest nor the most complex nonwords (see Table 2.2). In this experiment, nonword complexity was defined using measures of phonotactic probability (Vitevitch &
Luce, 2004). Phonotactic probability refers to the frequency with which a given phonological segment (or string of segments) occurs in a given position in a word (Jusczyk, Luce & Charles-Luce, 1994). Higher phonotactic probability indicates a more common string. Phonotactic probability has been shown to impact word learning: nonwords with higher phonotactic probability are learned more rapidly and retained more accurately than those with lower probability (Storkel, 2001). Other phonological factors have been shown to impact word learning, such as phonological neighborhood density (the number of words that sound similar to a given word; Storkel, Armrüster, & Hogan) and phoneme markedness (the number of distinctive phonemic features between words; Rogers, Sterling, & Storkel, 2001). Future studies should further assess the interaction of practice schedule and complexity by manipulating variables other than length, phonological structures and phonotactic probability in defining complexity.
Conclusion

For decades, limb motor learning studies have shown that practice schedule differentially impacts acquisition and retention. Shea & Morgan (1979) found that blocked practice facilitates acquisition of a skill while random practice facilitates retention. Random practice schedules result in increased contextual interference during acquisition. Increased contextual interference makes acquisition more difficult but results in more flexible action plans (Shea & Morgan, 1979; Magill & Hall, 1990) or stronger memory traces for the action plan (Lee et al., 1997; Magill & Hall, 1990), which facilitate retention performance.

This study extended well-established limb motor findings regarding blocked and random practice to the area of speech motor learning. The current results support previous findings of higher overall performance, during acquisition, by participants who engage in blocked practice of a novel speech motor skill. It was also found that the participants who engaged in random practice demonstrated higher scores during retention than those engaging in blocked practice, although these effects were found to be significant only for behavioral accuracy (% phonemes correct). While the trends for differences between the two groups are clear, especially at retention testing, it is likely that a larger sample size than the one used in the current experiment (10 participants per group) is needed to demonstrate clearer significant patterns in future research, both during acquisition and retention.

Practice is an integral part of several behavioral therapeutic methods in clinical speech language pathology (Maas, 2010). The current study on healthy adults with no
history of speech, language or neurological impairments serves as a foundation for future research that expands upon and extends the current findings to clinical populations and to different age groups.
Bibliography


Psychological Assessment Resources (2003). *Computerised Wisconsin Card Sorting Task-


