THE ROLE OF “FOCUS OF ATTENTION” ON THE LEARNING OF NON-NATIVE SPEECH SOUNDS: ENGLISH SPEAKERS LEARNING OF MANDARIN CHINESE TONES

By

Ruba B. Almelaifi

BS. Physical Therapy, Kuwait University, 1993

MS. Physiology, Kuwait University, 2001

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

By

Ruba B. Almelaifi

It was defended on

October 19, 2012

and approved by

Sheila R. Pratt, Ph.D., Associate Professor, Department of Communication Sciences and Disorders

Susan Shaiman, Ph.D., Associate Professor, Department of Communication Sciences and Disorders

Susan L. Whitney, PhD, Professor, Department of Physical Therapy

Dissertation Advisor: Malcolm R. McNeil, Distinguished Service Professor and Chair, Communication Science and Disorders

THE ROLE OF "FOCUS OF ATTENTION" ON THE LEARNING OF NON-NATIVE SPEECH SOUNDS: ENGLISH SPEAKERS LEARNING OF MANDARIN CHINESE TONES

Ruba B. Almelaifi, Ph,D.

University of Pittsburgh, 2012

Focus of attention (FOA) has been demonstrated to affect motor learning and performance of many motor skills. FOA refers to the performer's focus while performing the task. The purpose of this dissertation was to assess the role of FOA in the speech domain. The research asked whether external or internal FOA would individually or differentially facilitate the learning of Mandarin Chinese tones by native English speakers. As a secondary question and experimental control, this study also examined whether the four tones were produced with the same accuracy.

Forty-two females, between the ages of 18 and 24 were randomly assigned to one of three groups: external FOA (EFOA), internal FOA (IFOA) and control (C). During the acquisition phase, the groups were instructed to either focus on the sound produced (EFOA), the vibration in the voice box (IFOA), or no related FOA instructions (control). Participants were required to repeat the Mandarin words after an auditory model. To assess learning, the participants repeated the practiced words in a retention test, and repeated similar but unpracticed words during a transfer test. The data was collected in two sessions. The dependent variables were the root mean squared error (acoustic measure) and percentage of correctly perceived tones (perceptual measure).

There was a significant difference among the four Mandarin Chinese tones for the three groups (Tones 1 and 4 were produced with significantly higher accuracy than Tones 2 and 3) before acquisition phase. There was, however, no significant difference among the three FOA groups on the dependent variables.

The results contradict the FOA effects in the literature derived from limb motor learning and oral-nonspeech learning experiments. This study represents the first attempt to test the FOA in the speech domain. As such, it is premature to draw firm conclusions about the role of FOA in speech motor learning based on these results. The discussion focuses on factors that might have led to the current results. Because FOA represents a potential factor that might affect speech motor learning, future research is warranted to study the effect of FOA in the speech domain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First and foremost, I thank God for protecting me and guiding me throughout my life.

I would like to thank Dr. Malcolm McNeil, my dissertation chair, for his support, guidance, generosity, and advice throughout this dissertation process and my whole doctoral program. I would also like to extend my gratitude to the members of my Dissertation Committee, Drs. Sheila Pratt, Susan Shaiman and Susan Whitney. Thank you for your kindness and for the invaluable suggestions and advice, all of which have greatly improved this dissertation.

I would also like to express my sincere gratitude to Mr. Neil Szuminsky for developing the *"stimulate"* program and helping me set up the experiment and organize the data and for his generous help with technical issues. I also want to thank Dr. Elaine Rubinstein for her patience and help with the statistical analysis.

I gratefully acknowledge the School of Health and Rehabilitation Sciences at the University of Pittsburgh for the financial support for this project through the SHRS Research Development Fund. I would like to acknowledge both the Department of Psychology-Research Participation—at the University of Pittsburgh, and the Clinical Translational Science Institute (CTSI) Research Participants' Registry for helping me recruit participants.

I am grateful to all the wonderful people I have been surrounded by here in Pittsburgh. Their emotional support and friendship have helped me immensely, especially through times of

loneliness. Thank you to my family in Pittsburgh, Ronna Edelstein and Morton Edelstein, for your love and support. Dr. Abreena Tlumak, thank you for your encouragement, support and friendship. I also want to extend my thanks and gratitude to all my friends in Pittsburgh and Kuwait for their love and friendship.

I am grateful to my Family. I have no suitable word that can fully describe how blessed I am to have such a loving family. I would never have made it this far without them. Special thanks go to my sisters and brother for their endless love and encouragement (Abeer, Abrar, Reham and Mohammad). I am especially grateful for my sister, Reham, and for the way our relationship has deepened and grown into a very special friendship. Thank you to my father, Barrak Almelaifi, the best dad ever – in the words of Clarence Budington Kelland, "He didn't tell me how to live; he lived, and let me watch him do it." I love you. To my mother, Amna Al-hindi, thank you for your endless love, support, and prayers that have surrounded me all the years of my life. I love you. I hope I can be as wonderful a mother to my precious children as you were and are to me. To my children, Aminah, Mohammed, Dalal, Nora, and Sarah, you are the source of my life-strength, inspiration and love. The unconditioned love, cheering and understanding I felt from each and every one of you is priceless. I am proud of you.

Last but not the least; I am indebted to my husband, my hero, and best friend, Adel Alsoulah. I would never have earned a Ph.D. without you in my life. Almost anything I could say to describe your unconditioned love and support during this time of my life and our whole time together would pale in comparison to what you have been to me. I love you with all my heart.

1.0 INTRODUCTION

The primary means of communication among human beings is speech. Through speaking, humans express their feelings and their needs. If normal speech production is compromised by pathology—resulting from impaired respiration, phonation, articulation (lingual, labial, mandibular, or velar) or fluency—the speaker's quality of life might be affected (Klugman & Ross, 2002; Yaruss, 2010). Those individuals seeking rehabilitation for a speech (and other communication) impairment often face limitation in the number of treatment sessions covered by insurance companies. This limited treatment time not only pressures healthcare providers to make the most of the available treatment sessions to optimize treatment outcomes, but also can produce suboptimal evaluations and treatments. Because speech-language pathologists strive to improve the quality of the time spent in each treatment session, researchers and clinicians have been studying potential factors that might improve the treatment outcome. For example, speech-language pathologists have successfully implemented principles of motor learning from the limb motor literature in the treatment of motor speech disorders.

The limb motor literature also demonstrates that focus of attention (FOA) is another factor that affects motor learning. The hypothesis proposed by Wulf and colleagues (McNevin, Shea, & Wulf ,2003; Wulf, McNevin, & Shea, 2001) that focus of attention (FOA)—focusing on certain aspects of the motor task—affects the performance and learning of that motor task, has garnered a significant response in the motor limb literature. Although the FOA effect has not yet been studied in the speech domain, it represents a potential factor that might affect speech motor learning. FOA has the potential to be directly incorporated during speech therapy, either through the instructions or the feedback, as both are considered integral elements in any treatment procedure.

Researchers have recently demonstrated that the focus of attention (FOA), the aspect of the task on which the performers focus during task performance, affects the performance and learning of that task. The empirical evidence from the literature suggests that an external focus of attention (EFOA)—focusing on the effect of the movement—yields a better performance during practice, retention, and/or transfer tests than an internal focus of attention (IFOA) —focusing on the movement itself (Wulf, 2007; Wulf & Prinz, 2001). In these studies, explicit instructions, the feedback provided to the performer, or both, were used to manipulate FOA. Researchers reported the superiority of EFOA on performance and learning in laboratory tasks (stabilometer, skisimulator, etc.), and in real life sports (soccer, volleyball, golf, etc.) (Wulf, McConnel, Gartner, and Schwarz, 2002; Wulf & Su, 2007), and on improving balance in individuals with Parkinson’s disease (Wulf, Landers, Lewthwaite, and Töllner, 2009).

While the empirical evidence supporting the role of FOA is available for limb motor tasks, it is not available for speech tasks. However, speech movement researchers and clinicians have recognized that FOA might affect speech production as well and have, therefore, endorsed an exploration of FOA in the speech domain (Freedman, Maas, Caligiuri, Wulf, & Robin; 2007, Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008; Schulz, 2007 as cited in Wulf, 2007). Nonetheless, no study about the effect of FOA in the speech domain has been conducted. Therefore, the aim of this study was to extend the notion of FOA from the limb motor literature by testing the role of FOA in the speech domain.

Because this dissertation represents the first step in testing the generalizability of FOA effects on the speech domain, it seemed appropriate and necessary to review the FOA literature and some speech production models as well as speech motor control models. This study, which was based on the premise that speech—as a motor task— is viable to test the FOA construct, required such a detailed review. In this dissertation, FOA refers to the aspect of the task or the movement on which the performer concentrates while performing such movement (cf. Wulf, 2007).

This study examined the effects of FOA on the learning of non-native speech sounds (Mandarin Chinese tones) by young English speaking adults.

This study asked the following questions:

- Are there significant differences in the slope for the root mean square error (RMSE) scores across the acquisition phase of the experiment among the three groups: EFOA, IFOA, Control?
- Are there significant differences in the overall RMSE scores during the retention phase of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the overall RMSE scores during the transfer phase of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the slope for the percentage of the correctly perceived words across the acquisition phase of the experiment among the three groups: EFOA, IFO, Control?
- Are there significant differences in the percentage of correctly perceived words during the retention phases of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the percentage of correctly perceived words during the transfer phase of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the percentage of correct productions among the four tones for each group, during the acquisition phase: Tone 1, Tone 2, Tone 3, and Tone 4?

1.1 OVERVIEW

The statement of purpose and the significance of this study are addressed in chapters two and three, respectively. Chapter four—the literature review—consists of many sections. The first section (4.1) discusses the evolution of the FOA concept in the motor limb literature; the theoretical account of FOA effects; the empirical evidence of the FOA effects; and the applicability of FOA in other domains. Section 4.2 discusses the viability of testing the FOA in the speech domain. Sections 4.3 and 4.4 discuss some models of speech production and of speech motor control. This study adopted a motor learning approach; section 4.5 provides a description of the Schema theory and the principles of motor learning that were incorporated into this learning study. Section 4.6 provides a description of the tonal speech task employed in this study and validates such a tonal task for the purpose of this study.

Chapter five lists the research questions for the current study. Chapter six provides a detailed explanation of the methods and procedures employed. Chapter seven presents the results obtained, while chapter eight discusses these results in relation to the literature and addresses the current study limitations. Chapter seven provides concluding remarks and suggests areas for future research.

2.0 STATEMENT OF PURPOSE

Numerous studies have reported the effectiveness of adopting an external focus of attention for limb motor learning, and the possible application of this construct to other motor systems. Once these robust findings from the limb motor system emerged, speech production researchers then sought ways to apply these findings to the oral (speech) motor system (Freedman, Maas, Caligiuri, Wulf and Robin, 2007). Despite the accumulating evidence showing that external focus of attention can assist in the learning of motor skills, no study has been conducted to investigate the effect of attentional focus on speech production. Many researchers in the speech domain successfully extended the principle of motor learning from the limb literature to the speech domain; they also demonstrated that the principles of motor learning had comparable effect in the speech domain to that reported in the limb literature (Adams & Page, 2000; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000; Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008). Therefore, it is reasonable to hypothesize that the same beneficial effects of external focus of attention, demonstrated in the limb literature, would also apply to speech motor learning. Therefore, these promising results suggest the need for a study that examines whether or not the external attentional focus construct in motor skills research could be extended to promote learning in the speech domain. Earlier studies have successfully explored and demonstrated that the principles of motor learning do indeed apply to the speech domain (Adams & Page, 2000; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000; Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008).

In order to investigate the role of the attention focus utilized in a speech production task, this study's first objective was to evaluate whether the internal and external focus of attention have differential effects on the learning of a novel speech task.

3.0 SIGNIFICANCE

The proposed research is designed to contribute to the speech motor learning literature through extending the available knowledge by studying the effect of attentional focus (unexplored factor) on speech motor learning. Moreover, this study is considered as a starting point for future investigations of this construct in speech production. The knowledge to be gained from this study will have both practical and theoretical importance. From a practical viewpoint, knowing the specific effects of the instructions and feedback provided by clinicians to their patients might help to clarify why some patients do and others do not benefit from treatment. Theoretically, answering the proposed questions would also further the understanding of how speech as a motor task responds to other factors known to affect limb motor learning. Moreover, if the proposed study demonstrates that the focus of attention affects speech learning, this finding would extend the available knowledge about the role of focus of attention to include the speech domain.

The finding might also inform instructors in second language acquisition as to how to instruct and provide feedback that would optimize learning.

4.0 BACKGROUND

4.1 FOCUS OF ATTENTION (FOA)

This section describes the evolution of research in motor skill learning. Several studies have found that an external attentional focus enhances performance and learning of motor tasks better than an internal focus (Hodges & Franks, 2000; Wulf, 2007b; Wulf & Prinz, 2001). Wulf and her colleagues not only established this finding through detailed empirical research, but have also refined the construct. Although most researchers have conducted their studies in sports motor skills, they have also suggested the potential applicability of this construct to other fields (e.g., Brydges, Dubrowski and Carnahan, 2007; Freedman, Maas, Caligiuri, Wulf, & Robin; 2007; Mornell, 2007). Before reviewing the literature, the definition of FOA and its two forms are first provided. Following the definition, the theoretical accounts proposed by researchers to explain the advantage of adopting an EFOA are discussed.

“Attention” in its broader sense means the ability to focus specifically on the performed task or the ability to concentrate in general. According to Wulf and her colleagues, the term “focus of attention” refers to the part of the task or movement on which the performer concentrates. There are two types of FOA: external and internal focus of attention. An external focus of attention (EFOA) applies to those situations in which the performer focuses on the effect of his movement or on the equipment used in the performance of the task. On the other hand, internal focus of attention (IFOA) applies to those situations in which the performer directs attention to the movement itself (its sequence, timing) or to specific body parts/or muscles that perform the movement (Vance, Wulf, Tollner, McNevin, & Mercer, 2004; Wulf, 2007b). For this discussion, the phrases "focus of attention" and "attentional focus" will be used interchangeably. What denotes an external or internal focus depends upon the task performed. These terms are not always objectively defined or used consistently among studies, and as such, subjectivity in the definitions might cause an argument as to what constitutes an external or internal FOA. For example, two studies that utilized a golf task noted that an instruction that one research group considered inducing an internal focus (Perkins-Ceccato, et al., 2003) was interpreted by Wulf and Su, (2007) as either vague or as inducing an external focus. This issue of definition of term with respect to the speech task is further discussed in the definition of terms section. The next section provides an overview about the FOA as a viable construct in motor skills learning, followed by a presentation of the proposed theoretical account for this construct.

4.1.1 FOA as a Viable Construct in Motor Skills Learning

Attentional focus has been a topic of study in motor skills for more than a century; it was first addressed by Cattell in1893 (as cited in Freedman, Maas, Caligiuri, Wulf, & Robin, 2007). Later research showed that "being too concerned with, or even just paying attention to one's movement can disrupt the performance of well-practiced skills" (Wulf & Prinz, 2001, p. 648). Referring to a personal experience, when Wulf tried to apply instructions while windsurfing, she noted that concentrating on controlling her feet movement did not help her; rather, this behavior led to poor performance and a loss of balance. The effect of the focus on the task performance attracted Wulf's attention, guiding her to formulate a hypothesis about the effectiveness of external versus internal focus of attention (Wulf & Weight, 1997). In order to test her hypothesis, Wulf, HoB, and Prinz (1998) conducted the first study that looked at whether focusing one's attention externally is more beneficial to performance and learning than focusing attention internally on the movements themselves. In their Experiment 1, Wulf and Prinz asked their 33 participants to practice on a ski simulator by moving the platform in both directions with the highest amplitude. The researchers randomly assigned the participants to one of three groups: EFOA group, IFOA group, or control group; the participants were instructed to focus on the force they exerted on the wheels located under the platform, or to focus on their feet, or they were not given any additional instruction, respectively. However, participants in all three groups were told to look straight ahead while performing the task—to prevent any confound from the visual feedback. The researchers found that the EFOA group demonstrated the largest amplitude during both the practice days and during the retention test on the third day. The performance of participants in the IFOA group and the control group was comparable. The authors interpreted these results as supporting their hypothesis that instructions to focus externally have beneficial effects on performance during both practice and during the retention test.

In Experiment 2, the researchers sought to replicate their findings using another task. Wulf et al. (1998) asked their participants, 16 university students, to balance on a stabilometer. They randomly assigned the participants to one of two groups: IFOA group or EFOA group. Participants in the internal focus group were instructed to focus on their feet, while participants in the external focus group were instructed to focus externally on the markers on the stabilometer board. The participants practiced under their assigned conditions for two days and performed a retention test on the third day. The results showed that the external focus group performed significantly better, in terms of stability of the platform, than the internal focus group only during the retention test. Wulf et al. (1998) concluded that “both experiments were clear in demonstrating that instructing learners to focus on the apparatus (i.e., the wheels of the ski-simulator in Experiment 1, the markers on the stabilometer in Experiment 2) was clearly more effective for the learning of those tasks than directing their attention to their body movements (i.e., their feet)”(p.176).

The empirical evidence that directing attention to what one is doing causes performance to decrease (e.g., Hodges & Franks, 2000; Wulf, 2007b; Wulf & Prinz, 2001). Wulf and Prinz (2001) deem these findings as “quite worrisome” because, as Wulf and colleagues noted, most instructors in sports might be giving instructions and providing feedback in ways that may actually impair learning. In her research, Wulf (2007b) found that “the accuracy and quality of the movement depends to a great extent on what the performer focuses on while executing the skill” (p. 4). She added that how fast and how well a skill is learned results from the type of the adopted focus of attention (Wulf, 2007b).

Pursuing these observations, Wulf (2007b) then demonstrated that an external attentional focus, or a focus on the “effects” that a movement has on the environment, as opposed to an internal attentional focus, or a focus on one’s own body movements, leads to optimal performance in other motor skills. In another study that utilized a stabilometer, Wulf asked participants—in the two groups—either to focus on keeping specified markers horizontal (external focus) or to focus on their feet (internal focus). The results indicated that those instructed to adopt an external focus “ demonstrated more effective learning than those provided with internal focus instructions” (Wulf, p. 8).

Since the late 1990s, Wulf and her colleagues have demonstrated the advantages of EFOA as opposed to IFOA. Additional studies, which replicated the learning advantages of external focus, also generalized these advantages to other fields related to real-world skills, such as soccer, volleyball (Wulf, McConnel, Gartner, and Schwarz, 2002) and golf (e.g.,Wulf & Su, 2007). Additional studies tried to determine whether or not external focus was generally better for most subjects, or simply a preference of some subjects over others; the results of these studies created controversy about the generalizability to performers with different skill levels (Perkins-Ceccato, Passmore, & Lee ,2003; Wulf & Su, 2007b).

In the literature, researchers have manipulated their participants' FOA either by instruction or feedback or both. Usually, instructions are given to participants before task performance begins, while feedback, which includes information about the performance of the previous trial, is provided either during or after the trial. The proposed theoretical accounts for FOA effects, along with empirical evidence, are discussed in the next section.

4.1.2 Theoretical accounts of FOA effects

Wulf and her colleagues (e.g., McNevin, Shea, & Wulf ,2003; Wulf, McNevin, & Shea, 2001) put forward the "Constrained Action Hypothesis" (CAH) to explain the superior effect of external focus of attention versus internal focus of attention when an individual performs and learns a motor skill. As described by Wulf and her colleagues, the motor system is capable of organizing itself in an automatic and optimal way only if nothing interferes with this automaticity. This is achieved when the individual focuses externally when performing a motor skill, thereby freeing the motor system to be in command when achieving the intended goal.

Conversely, when an individual tries to control his movements to enhance his performance, he actually restrains the motor system from implementing its natural organization. In other words, adopting an internal focus of attention by focusing on the movements themselves would not benefit the individual; rather, it would constrain the motor system, rendering it unable to perform in a coherent way to achieve the intended goal. Wulf et al.(2001) provided three sources of empirical evidence to support the CAH: a) external and internal foci of attention are differentiated in terms of attentional demand; b) frequency of movement adjustments differs according to focus of attention; and c) amount and extent of muscle activity differ according to the adopted focus of attention.

The first source of empirical evidence for the CAH is based on the assumption that an automatic process requires less attention. In order to test this prediction of the CAH, that acquiring an external focus of attention would lead to a more automatic performance, Wulf, McNevin, and Shea (2001) utilized a dual task paradigm—to infer the attention demand of the primary task. The participants' task was to maintain balance on a stabilometer—primary task—and to concurrently perform a secondary task— pressing a button in response to an auditory stimulus (tone) as fast as possible. The researchers randomly assigned the 28 participants to one of two groups: an EFOA group, in which the participants were instructed to focus on markers attached to the platform, or an IFOA group, in which participants were instructed to focus on their feet to maintain balance.

The results indicated that, compared to the IFOA group, participants in the EFOA group responded significantly faster to the secondary task during both the practice and the retention test. Also, the EFOA group demonstrated significantly higher accuracy in the balancing task when compared to the IFOA group. The authors interpreted these results, which illustrated that external focus of attention required less attention from the performer, indicating a more automatic control of the movement, as support for the CAH. However, the researchers in this study did not include a control group, so it is unclear whether the beneficial advantage of the external focus was due to a detrimental effect of the internal focus of attention. Including a control group would be useful in order to compare these effects.

The second kind of set of evidence provided by the researchers to support the CAH was based on the notion that if the body is capable of controlling a movement by organizing itself, it can utilize all the degrees of freedom in the motor system in order to produce optimum adjustments to the movement; Wulf and her colleagues considered this to be the case during the external focus of attention. On the other hand, controlling the motor system by focusing internally would cause the motor system to stiffen with less effective movement adjustments. McNevin, Shea, and Wulf, (2003) conducted a study to examine the advantages of external versus internal focus of attention and to test whether changing the distance of the external focus from the body affects the benefits of EFOA. The participants' task was to balance on a stabilometer (just like the primary task in Wulf, McNevin, and Shea , 2001). Forty university students were randomly assigned to one of four groups: one internal focus group and three external focus groups (in which the distance of each group point of focus was manipulated: one group with near focus and two groups with far focus points). Based upon their assigned group, the participants received instructions to focus externally or internally. The researchers directed the participants in the single internal focus group to concentrate on their feet and those in the three separate external focus groups to concentrate on markers positioned at different distances from the body (on the platform). As a result, the following occurred: 1) participants of the near external focus group focused on a marker positioned in front of their feet; 2) those in the far-inside external focus group focused on a marker positioned between their feet; and 3) those in the far-outside external focus group focused on a marker positioned outside their feet (for both the inside and outside far groups, the markers were positioned at the same distance from the feet). Although the researchers instructed the participants in all groups to focus on keeping the markers on the same horizontal level, the researchers requested that the participants look straight ahead while performing the task to prevent any confounds from the visual feedback from the markers.

The results confirmed the previous finding: the group instructed to focus internally made significantly more errors on the retention test than the group that focused externally. More interestingly, both of the far external focus groups had significantly higher frequency adjustments on the platform (to correct errors in their balance) when compared to both the near group of external focus and the internal group. The high frequency adjustment indicates a more automatic response in which the body can move with many degrees of freedom. By contrast, in an attempt to control the movement by focusing on the movements themselves, the performer will tend to constrain the body's automaticity, which will render the body too stiff. This is exactly what the researchers found in this study: the group that focused internally showed movements with lower frequency of adjustments and higher amplitude of movements, indicating less automatic control in the participants' movements. However, the researchers did not include a control group with which to compare the performance; instead, they considered the internal focus group as their control. Due to the omission of a control group, this study makes it difficult to confirm whether these significant differences are due to the beneficial effects of the external focus of attention or due to the negative effects of adopting an internal focus.

The third source of evidence for the CAH was presented from physiological measurements at the neuromuscular level. Vance, Wulf, Tollner, McNevin, and Mercer (2004) conducted a study that sought to determine the degree to which external or internal focus during performance manifest physiologically at neuromuscular level activity. In a within-subject design, eleven participants performed bicep curls while holding a bar under both external and internal focus conditions. Participants were instructed to focus either on the bar or on their arms during external and internal focus conditions, respectively. The order of the two conditions was counterbalanced among subjects. The participants' performance was evaluated in terms of integrated EMG (i-EMG) activity, which takes into account the movement time. The results demonstrated a significant main effect of focus condition; i-EMG was significantly lower during the external focus condition than during the internal focus condition. The researchers surmised that if the CAH prediction was correct, "One might expect fewer motor units to be recruited under external than under internal focus conditions for the same task requirements" (p.451), a finding that would reflect a greater economy in movement production and better outcomes. Overall, they found less EMG activity when the subjects utilized an external focus, a finding that correlates with the CAH, which posits that "an external focus promotes the use of more automatic control processes" (Vance et al., p. 450). Zachry, Wulf, Mercer, and Bezodis (2005) have replicated these findings by employing basketball free-throws as a task.

Recent evidence from an imaging study supports the CAH. Though acknowledging that the research on focus of attention is compelling, Zentgraf, Lorey, Bischoff, Zimmermann, Stark, and Munzert (2009) also noted that at present, "The neurophysiological basis of this phenomenon remains largely unknown." In order to redress this failing, Zentgraf et al. (2009) sought to link these findings with "a rather different line of investigation using neuroimaging methods to study the neural correlates of attentional modulation during motor execution" (p. 535). Thus, they attempted to link attentional focus research "derived from the sports and movement sciences and the attention-to-action studies from neurophysiology" in order to begin to establish a needed synthesis (Zentgraf et al., p. 535). Using functional magnetic resonance imaging (**f**MRI), Zentgraf et al. (2009) investigated whether adopting an internal or an external focus would have an association with activation of specific localities in the brain. The researchers asked their 31 participants to practice a finger- button press task with either EFOA or IFOA in which they were asked to focus on the button they press or on their finger movement, respectively. During the **f**MRI scanning, each group was required to recall the practiced movement under three randomized conditions among participants: 1) assigned practice condition (internal or external focus); 2) dual task condition (perform the movement while counting auditory tones); and 3) no focus condition while performing the movement—movement-only condition. The results indicated that although the participants in the EFOA group and IFOA group did not differ in terms of the duration of the button-press task, the **f**MRI scanning revealed differences between the groups. The results indicated a "higher activation in the primary somatosensory, motor, and insular cortices for an external contrasted to an internal focus" (p. 539). Zentgraf et al. (2009) interpreted these results as demonstrating that focusing attention highlights task-relevant information, and "external focus seems to enhance the processing of input from the tactile modality that mediates the interaction with the task-related environment" (p. 540). Thus, while "internal focusing may lead to a perturbation of the efficient flow of neural signals between sensory and motor areas," "external focus…implicitly provides the actor with more task-adequate signals" (Zentgraf et al., p. 540). Shifting attention results in shifts to exteroceptive sensory inputs (in the case of this study, the tactile sensations at the fingertips); this, in turn, correlates with and supports the CAH.

Wulf and her colleagues interpreted the results of these studies as validation for the CAH—that directing the performers' attention externally showed enhanced performance and efficient neuromuscular activity. Because these studies (e.g., McNevin, Shea, & Wulf ,2003; Wulf, McNevin, & Shea, 2001) only measured the effects during practice, it is not clear whether these effects are only temporary. Consequently, more studies are needed that include a retention test. Furthermore, although these results support the beneficial effects of EFOA, neither the CAH nor the results of these studies specify in detail the central processes behind the beneficial effects of EFOA or the detrimental effects of IFOA. While most researchers acknowledged that the literature has produced consistent evidence of the beneficial effects of external focus of attention on performance during practice, retention, or transfer test, some researchers argued that the CAH cannot accommodate the results of both EFOA and IFOA conditions.

Bund, Wiemeyer, and Angert (2007) questioned the CAH and its argument that external focus facilitates automatic control. The researchers inquired about two areas: 1) whether or not just acquiring an external focus would guarantee this automatic process and enhance performance diversity of motor skills, regardless of the complexity and nature of the motor skills and 2) whether this would also apply when a novice learner is performing a task in an early learning stage.

Similarly, Ehrlenspiel (2007) raised doubts about the CAH by noting that "convincing answers" have not yet been devised for exactly how internal focus interferes with the development of an automatic process. From his perspective, the CAH does not present the whole picture. Therefore, Ehrlenspiel (2007) argued that the model fails to connect the cognitive level of the automatic processes with the “movement level of disturbed movement execution” (p. 19). A hypothetical model for this effect, devised in the literature, posits that internal focus acts to break a skill down into smaller parts, which then must be activated separately “and run separately, which slows performance and…creates the opportunity for error that was not present in the ‘chunked’ control structures” (Ehrlenspiel, p. 19).

As an alternative to the CAH, Hossner and Ehrlenspiel (2007) suggested a nodal-point hypothesis (as cited in Ehrlenspiel, 2007), asserting that movements are controlled “by the anticipation of their sensory effects”. That is, when movements are chunked, the end-effects take over and control the movements, “consequently reducing the necessity of checking the attainment of intermediate effects,” a result which is perceived as automatism (p. 19). An internal focus might then well cause an inversion of the “serial chaining mechanism” insofar as attention is directed to intermediate effects overlooked by an expert performer (Ehrlenspiel, 2007). Thus, while creating a chain frees up movement by reducing muscular activity, focusing again on intermediate tasks interrupts this automaticity and takes more effort. When Hossner and Ehrlenspiel (2007) tested the prediction of their model in a basket free throw task, their results indicated that “focusing on a nodal point resulted in an increased relative muscular activity at that nodal point compared to the other nodal points that were not in focus” (P. 20).

Ehrlenspiel (2007) endorsed the nodal-point hypothesis, which provides more details about the mechanism linking the cognitive and physical elements of motor learning. Interestingly, this hypothesis does not favor either internal or external focus of attention, but instead stresses the “relevant effects that are reliably attained and thus anticipated whether internal … or external” (p. 20).

In sum, the CAH in its current formulation can only accommodate the effects of external focus of attention but not the detrimental effects of adopting an internal focus of attention. On the other hand, the nodal-point hypothesis provides a fruitful venue for future research directed toward the study of automatic motor skills. CAH, therefore, requires an enhancement to enable it to explain the effects of both external and internal focus of attention and take into account all factors that might interact with these effects, such as task difficulty, level of expertise, and the nature of the motor task.

The following section reviews the literature on the effect of FOA on motor skills.

4.1.3 Focus of Attention and Motor Skills

The following subsections reviews studies on FOA according to the method of FOA manipulation and task similarity. The first section discusses the few available studies in which the feedback directed the performers' FOA; the second section reviews the studies, in which instruction directed the performers' FOA. The final section of the literature review discusses the significance of these findings.

4.1.3.1 Focus of Attention Induced by Feedback

In the first empirical study regarding the effect of focus of attention on the performance and learning of a motor skill, Wulf, HoB, and Prinz (1998) utilized instructions to induce either external or internal focus of attention. A subsequent study by Shea and Wulf (1999) sought to determine: 1) whether feedback provided to the performer would direct the focus of attention (external or internal) and 2) whether feedback that induces EFOA would enhance performance and learning. Using a stabilometer task, Shea and Wulf (1999) asked their 32 participants to maintain their balance by keeping the platform in a horizontal position. The participants were randomly assigned to one of four groups: 1) internal focus induced by instruction, 2) external focus induced by instruction, 3) internal focus induced by feedback, and 4) external focus induced by feedback. For the two instruction groups, participants were instructed to keep the markers—positioned in front of their feet—at the same level, or to keep their feet on the same level in the EFOA and the IFOA groups, respectively. Participants in the feedback groups received the same feedback—two lines presented on a computer monitor, but the researchers told the participants the feedback either represented the markers in front of the participants' feet, or the participants' feet in the EFOA and IFOA groups, respectively. The results showed that, compared to the instruction groups, participants in both feedback groups performed the task with significantly less error during both practice and the retention test. Also, the external focus groups, whether guided by instruction or feedback, outperformed the internal groups during the retention test. Shea and Wulf concluded from these results that "the learning benefits of an external attentional focus seem to generalize to the feedback given to the learner" (P. 553). Although this study tested whether the beneficial effects of the external focus produced by instruction will also be generalized to external focus implemented by feedback, Shea and Wulf did not include groups that only received feedback; their feedback groups received both instructions and feedback. The results from this study, therefore, do not validate whether the focus conditions induced only by feedback affected the performance, or whether these reported findings represented just an additive effect to the external focus induced by instruction. It would have been helpful if this study had included two groups that only received feedback in order to separate this confounded effect.

In another experiment, Wulf, McConnel, Gartner, and Schwarz (2002) examined the effectiveness of feedback—directing the participants' to focus either externally or internally—in a realistic environment utilizing actual sport skills usually performed in volleyball and soccer. In their Experiment 1, the researchers recruited two groups of students: novices and experts. The novices had no prior experience with volleyball, while the experts were advanced players on their school team. They randomly assigned 48 participants to either an external or internal focus group, resulting in four groups: 1) novices-external group, 2) novices-internal group, 3) experts-external group, and 4) experts-internal group.

The researchers asked their participants to shoot a ball using a volleyball "tennis serve" into a target area. The feedback consisted of verbal sentences that resembled those used by coaches during training sessions; the content of the feedback was the same for both groups but differed only in wording and was provided on every fifth trial. The participants practiced for two sessions (one week apart); on the third session, which was one week later, the participants performed a retention test.

The results showed that participants in groups who received external focused feedback were more accurate in their serves during both practice and retention test independent of the expertise level than those participants receiving internal focus feedback. In terms of movement quality, although both expert groups scored higher than the two novice groups throughout the experiment, both groups that received external feedback had higher scores compared to their internal group counterparts during practice; however, during retention, the movement form of the external and internal groups was not different. A secondary interesting finding was that the withdrawal of the internal focus feedback during the retention test enhanced the performance of the novice internal focus group, which resulted in movement quality scores comparable to the external focus novice group.

This differential effect of internal and external focus feedback—provided with the same frequency—cannot be accommodated by the guidance hypothesis. This hypothesis predicts that providing feedback during practice would guide the performers to the correct response, and as a result, practice with feedback would have fewer errors. Moreover, this hypothesis suggests that, during retention, when the guiding effect of feedback is no longer available, the performance would be poor. This controversy in results led Wulf and her colleagues (2002) to attribute the differential effect of their feedback to the induced focus of attention. The internal focus—as directed by the feedback provided during practice—might have caused the performers to focus on their movement during practice; this can explain the decrease in the novice internal focus groups' performance and the enhancement seen in the internal focus group movement quality performance during retention, when the feedback was removed. In this study, the external-focus feedback made more use of metaphors or analogies, unlike the sentences used in the internal focus feedback. Wulf et al. (2002) conjecture that these metaphors are advantageous for external focus because they "…detract the performer's attention from his or her body movements and at the same time provide a mental image of a movement goal" (p. 176). Accordingly, they concluded that "attention focus induced by feedback can indeed have an effect on learning" (p. 176).

In response to these findings, Wulf et al. (2002) proposed a potential interaction between feedback frequency and attention focus. To investigate their hypothesis, the researchers conducted a second experiment in which they studied the performance of four groups according to type and frequency of feedback provided: 1) 100%-internal focus feedback, 2) 33%-internal focus feedback, 3) 100%-external focus feedback, and 4) 33%-external focus feedback. Wulf and colleagues (2002) recruited 52 students, all of whom had some experience in soccer, to practice shooting at a target. The feedback was in the form of sentences that indicated how the performers could improve their skill according to the produced movement. All participants practiced for 30 trials with feedback, and performed a retention test one week later with no feedback. The results showed a main effect of attention focus: the two external focus groups were more accurate than the two internal focus groups, regardless of the frequency of feedback. The researchers found that the interaction between feedback type and frequency was significant and confirmed their speculation about the differential effect of feedbacks. During practice, the 33%- internal focus feedback group performed with higher accuracy as compared to the 100%-internal focus feedback group; the opposite was true for both external focus groups. A non-significant trend revealed that more frequent feedback inducing external focus (100%) was more beneficial than a less frequent feedback (33%). It seems that an infrequent as opposed to frequent internal feedback was also found to give performers more opportunity to focus externally. Wulf, Chiviacowsky, Schiller and Avilla (2010) reported similar findings with children performing a soccer throw-in task.

As a result of these findings, the researchers suggested that using only the guidance hypothesis to interpret feedback frequency results in any study might be misleading by not fully explaining the nature of the relationship between feedback frequency and type of FOA. Instead, they highly recommended that the results should also be interpreted from the attentional focus perspective as well.

These findings have potentially important practical implications. Current practice in motor learning assumes the importance of the conscious control of movement and feedback on the specific movements upon which the performers focus. This practice views conscious control as essential for effective learning. According to the guidance hypothesis, feedback, regardless of its focus, helps the learners improve performance by counter-intuitively preventing practitioners from focusing on their movements and helps the learners improve their performance. This hypothesis is, however, counter to the results of the study by Wulf and colleagues (2002). In a review paper, Wulf and Prinz (2001) reported that such internal focus feedback causes performers to focus too much on their movements, thereby diminishing their performance. Overall, the underlying mechanisms of these findings suggest that thinking about one's movements or what one is learning may "inadvertently disrupt relatively autonomic processes that normally control the movement" (p. 654). These results showed how changes in feedback interpretation or wording might have a strong impact on performance and learning. This finding has great practical implications for learning and treatment.

4.1.3.2 Externality and Distance of Focus of Attention

The promising results regarding the positive role of external as opposed to internal attentional focus in motor skills performance and learning encouraged McNevin, Shea and Wulf (2003) to further speculate whether increasing the distance between the body and the external focus of attention reinforces the externalizing effect and provides additional improvement to performance. Therefore, the researchers conducted a study to test the hypothesis that "increasing the distance between the body and the action effects might further enhance the learning advantages associated with an external focus of attention" (P. 22). For this study, they randomly assigned 40 university students to one of four groups: one internal focus group and three external focus groups. The performers' task was to balance on a stabilimeter (just like the primary task in Wulf, McNevin, and Shea) (2001). They instructed the internal focus group to focus on their feet and the external focus groups to focus on markers positioned differently from the body (on the platform). The near external focus group focused on a marker positioned in front of their feet, the far-inside external focus group focused on a marker positioned between their feet, and the far-outside external focus group focused on a marker positioned outside their feet (for both the inside and outside far groups, the markers were positioned, on the platform, at the same distance from the feet).

The results confirmed the previous finding that on the retention test, the internal focus group had significantly higher RMSE scores than the external focus groups. The results also showed that while a more distant external focus did not affect performance during practice when RMSE was measured, the effect of a more distant focus of attention was obvious on the frequency of body adjustments (McNevin et al., 2003). Specifically, both the far external focus groups had significantly higher movement frequency adjustments on the platform and fewer errors as compared to both the near external group and the internal group during both practice and retention. The high frequency adjustment indicates more automatic responses by which the body can move with many degrees of freedom. In contrast, perhaps the performers attempting to control their movements by focusing on their feet tended to constrain this automaticity; their body, too stiff to adjust to any disturbance in balance, showed a low frequency of movement adjustments.

The researchers explained their findings by stating that “focusing on more distance effects results in enhanced learning by promoting the utilization of more natural control mechanisms” (McNevin et al., p. 22). The value of EFOA, then, lies in the ability of EFOA to enable more automaticity of action. This is especially true in highly rhythmic elements where external focus, unlike internal focus which constrains the motor system, allows the motor system as a whole to take over, resulting in an undisturbed system. The researchers argued that only a far-distant external focus may reinforce this, as an external focus which is still too close to the body may be considered or perceived as a kind of internal focus (the focus of the near external group in McNevin et al. study).

In a similar vein, the reported interaction between external foci of attention and the external foci's distance from the body has important practical implications. In a review paper, McNevin, Wulf and Carlson (2000) discussed the role of the following three factors in improving physical therapy rehabilitation: 1) directing the patient's focus to adopt EFOA; 2) involving the patient in setting the treatment goals, and 3) practicing in dyads. The authors related the third factor to the FOA notion by noting that the patient interacting with another person may further externalize his focus and thus improve the effectiveness of the practice. Moreover, practicing in dyads adds a competitive component to practice and causes people to set higher goals; this also benefits performance of motor skills.

Wulf and Prinz (2001) traced the origin of the relevance of remote as opposed to close external effects back to Lotze, who, in the 1850s, suggested the ideomotor principle on the relationship between representations and events (as cited in Wulf & Prinz, 2001). The ideomotor principle was updated in the 1990s as the common coding of perception and action theory. This theory argues that perception and action derive from a common representational medium with codes generated "in a commensurate way only at a distance level of representation" (Wulf & Prinz, p. 654). The idea that focusing externally on the goal of the task with anticipation of a desired outcome improves performance reinforces the importance of external focus; this view supports the idea that focusing on effects versus movements, or focusing externally rather than internally, improves performance.

As discussed above, proximal and distant or remote effects have also been explored within the scope of external focus of attention. The results generally show the similarity between the near external focus and internal focus. Overall, the literature on external attentional focus finds that “focusing on a more remote effect seems to facilitate the discriminability of the effect from the body movements that produced it and to be more beneficial than focusing on a very close effect” (Wulf & Prinz, 2001, p. 654). Researchers have also tested the effect of FOA on a person performing simultaneous motor tasks; these studies are discussed in the following section.

4.1.3.3 Focus of Attention and Supra-postural Tasks

Taking the externality of focus of attention a step further, Wulf and her colleagues inquired whether the benefits of adopting EFOA would extend to the performance of another task. The researchers wanted to know whether focusing externally on the supra-postural task while trying to maintain balance would indirectly enhance the performance on the postural or the balance task. An example of such a simultaneous motor tasks would be someone holding a cup of coffee and trying to maintain his balance while walking on an uneven floor. In such a case, holding the cup is the supra-postural task, while trying to maintain balance is the postural task.

McNevin and Wulf (2002) wanted to know whether adopting an external or internal focus of attention while performing a supra-postural task would affect the performance of the postural task. For this purpose, they conducted a study utilizing a within-subject design that included 19 volunteers. The participants’ task was to stand still (with their eyes closed) while lightly touching a hanging sheet with their fingertips. The researchers instructed their participants to minimize the movement of the hanging sheet under three conditions: 1) the external focus condition, in which the participants were instructed to focus on reducing the sheet movement; 2) the internal focus condition, in which the participants were instructed to focus on reducing their fingers movement; and 3) the control condition, in which the participants did not receive further instruction, and no sheet was available under this condition (the participants' task was to simply stand still). The practice consisted of performing three trials per each condition; the order of these conditions was counter-balanced among subjects. The results demonstrated a better performance in the external focus condition compared to the similar internal and control conditions. The advantage of the external condition was shown by higher movement adjustment frequency and lower amplitude values of the participants' postural sway. These findings demonstrated that external focus of attention enhanced postural stability, despite an unexpected finding of increased postural sway (McNevin & Wulf, 2002). The researchers attributed the increase in the postural sway found during the external focus condition to the limited postural sway during the control condition. This finding is not surprising because during the control condition, the participants were instructed to stand still; the absence of a sheet hanging in front of them during the control condition might have changed the task in the control condition to a standstill task, which is totally different and not comparable to the other two experimental conditions.

In another similar study, instead of utilizing a static posture task, Wulf, Weigelt, Poulter, and McNevin (2003) used the stabilometer task to examine whether the beneficial effects of external focus of attention on a supra-postural task would enhance the performance and learning of this dynamical balance task (postural task). In Experiment 1, the researchers randomly assigned 18 participants to one of two groups: an external focus group and an internal focus group. The researchers required the participants to balance on a stabilometer while holding a hollow bar in a horizontal position—within which a table tennis ball was enclosed. The researchers instructed the external focus group to focus on the bar, while the researchers instructed the participants in the internal focus group to focus on their hands. In this experiment, holding the bar was the supra-postural task, while balancing on the stabilometer was the postural task.

The results showed that, compared to the internal focus group, the participants in the external focus group demonstrated better balance as indicated by fewer errors in both the supra-postural task and the postural task. This effect was significant not only during practice, but also during the retention and transfer tests. The researchers interpreted their results as indicating that adopting an external focus on the supra-postural task not only improved the performance of that task but also enhanced the performance of the postural task. The researchers acknowledged that this experiment did not include a control group with which to compare the effects of external focus. Moreover, the researchers noted that the presence of the tennis ball in the bar might have altered the participants' performance, because the ball created a sound when it contacted the end of the bar. Perhaps the higher degrees of errors in the internal focus group were derived from the distraction of the ball sound rather than from the focus they adopted (Wulf, et al., 2003). To control for this possibility, the researchers conducted a second experiment.

In Experiment 2, the researchers included a control group and also removed the tennis ball from the bar. The procedures were the same as in Experiment 1; before the second trial, the researchers instructed the participants according to their assigned groups. However, unlike Experiment 1, participants did not receive instructions before the first trial to ensure that groups were equivalent in their performance. The results of Experiment 2 confirmed and extended the findings of Experiment 1: 1) all groups performed in a similar way on the first trial (ruling out any sampling bias); 2) the external focus group outperformed both the control group and the internal focus group; 3) on the retention test, participants in the internal focus group had a higher error score, indicating a lower performance, than those in the control group; and 4) on the transfer test, the performance of the internal focus group did not differ from that of the control group. The researchers interpreted the results from both experiments as showing that external focus of attention improved performance in both the supra-postural and postural tasks as compared to an internal focus of attention, which was detrimental to performance when compared to the control group. To explain the advantage of adopting an external focus of attention during the supra-postural tasks for the performance of the postural task, the researchers referred to the decreased attentional demand when the supra-postural task is performed with EFOA (Wulf & Prinz, 2001); this allowed more attention for the dynamic balance task (Wulf, et al., 2003).

In an attempt to further refine the nature of the relationship between supra-postural and postural tasks in terms of the adopted FOA, Wulf, Mercer, McNevin and Guadagnoli (2004) conducted a subsequent study. In a within-subject design, Wulf et al. asked their participants to balance on an inflated disk—the postural task—while horizontally holding a pole in their hand—the supra-postural task. All participants performed three trials under the following four conditions: 1) EFOA on the supra-postural task—focusing on the pole; 2) IFOA on the supra-postural task—focusing on their hands; 3) EFOA on the postural task—focusing on the disc; 4) IFOA on the postural task—focusing on their feet. The participants were instructed to reduce the movement of the pole, their hands, the disc, or their feet, in the above four conditions, respectively. RMSE measured the participants' postural sway. Mean power frequency (MPF) was also measured to indicate the frequency of movement adjustments, with a higher MPF indicating better adjustment. For the supra-postural task, both the MPF and RMSE measured the stability of the pole.

As expected by the researchers, results indicated that external focus of attention in both tasks (focusing on the pole or the disc) caused significantly less postural sway when compared to the internal focus conditions. When the priority of focus was on the supra-postural task, the participants under EFOA conditions performed with increased postural adjustments compared to the internal focus condition. When the priority of focus was on the postural task (either by external or internal focus), the participants showed higher postural adjustments (higher MPF). Interestingly, the postural adjustment was significantly higher when participants focused externally on the supra-postural task than when they focused primarily on their postural task. Wulf and colleagues interpreted these results as validating the advantage of EFOA over IFOA, and also as indicating that by focusing externally on the supra-postural task, participants maintained their postural movements, and the task "resulted in spontaneous reductions in postural fluctuations to facilitate the achievement of the supra-postural goal" (Wulf et al., p. 189). Because the effect of external attentional focus on the supra-postural task was more influential on the postural task, the researchers also concluded that the relationship is one-directional.

In all the above reviewed studies, it is apparent that the emphasis was either on the differential effect of external versus internal focus of attention on diversity of motor skills (Wulf, 2007; Wulf & Prinz, 2001), or on how acquiring an external focus of attention is beneficial in enhancing performance (McNevin, Shea, & Wulf, 2003; Wulf, McNevin, and Shea, 2001). Instead of studying the effect of external focus of attention by using a dynamic task or supra-postural task, Vuillerme and Nafati (2007) wanted to know how internal focus of attention holds back the automaticity of the motor system and hinders performance during quiet standing. Employing a within-subject design, Vuillerme and Nafati asked their 16 healthy young adult participants to stand upright on a force platform in two conditions: IFOA condition and control condition. For the IFOA condition, the researchers instructed the participants to actively control their posture and reduce their body sway, while the researchers instructed the participants in the control condition to simply stand still. The order of these conditions was counter-balanced between participants who performed three trials under each condition. During both the IFOA and control conditions, Vuillerme and Nafati asked the participants to respond verbally to auditory stimuli. The participants' reaction time was measured to infer the participants' attention demands during each condition as a manipulation check to ensure that the participants followed the instructions. Moreover, the participants provided a "subjective rating" of how much they think they were actively engaged in both conditions. Performance on postural control was measured by "center of gravity vertical projections (CoGv)" (an index of postural performance) and by the difference between the center of foot pressure (CoP) and "the centre of gravity vertical projection (CoGv)" (CoP-CoGv); this difference score is considered as a measure of undisturbed stance control reflecting the neuromuscular activity involved in balance in terms of ankle joint stiffness.

The results showed that compared to the control condition, the reaction time was significantly higher during the IFOA condition; the reaction time results also matched the participants' reported amount of task engagement. Vuillerme and Nafati considered these findings as indications that their instruction induced the IFOA condition and that participants did actively engage in the task. Nonetheless, no significant difference on postural performance (CoGv) occurred between the conditions. Vuillerme and Nafati attributed the non-significant difference between the conditions to the nature of the task (quiet standing), which they think was not challenging enough for the young participants in this study. However, the attention and control conditions differed significantly in terms of the undisturbed stance control (CoP-CoGv) which was higher during the IFOA condition, indicating that the body was stiffer (increased ankle joint stiffness) when the participants were actively controlling their posture.

Based upon these results, Vuillerme and Nafati concluded that focusing internally on body sways “promoted the use of less automatic control processes and hampered the efficiency for controlling posture during quiet standing” (p. 192). Because focusing on one’s body sway represents an internal as opposed to an external attentional focus, the study results are consistent with the FOA literature findings that internal attentional focus hampers the performance of motor skills tasks. Paying attention to something as simple as body sways represents a highly discrete source of evidence supporting the CAH (Vuillerme & Nafati, 2007). Although the task used in this study might not have been challenging enough, as the authors acknowledged, they still succeeded in the following two ways: 1) they justified the number of trials used in their study; and 2) they used a secondary probe reaction time to measure the attentional demand in the two conditions to ensure that their instructions manipulated the independent variable (FOA). Moreover, Vuillerme and Nafati also asked the participants to rate on a numerical scale how much they were involved in actively trying to control their posture as they wanted to make sure the participants really complied with the instructions. Checking the independent variable manipulation is not mandatory in each study (especially if the results do not suggest any alternative explanation and the conditions differed on the dependent variable); nonetheless, the manipulation check benefited this study because it demonstrated that the independent variable was indeed manipulated.

In contrast, Wulf and her colleagues extensively manipulated the focus of attention either through instructions or feedback; however, except for reminding their participants about the instructions between trials, they did not provide any other manipulation check for their independent variable. This issue raises an important point that instruction and feedback have been powerful in directing the participants' FOA. Therefore, in providing instructions and feedback, instructors should carefully word their directions in order to optimize performance during practice and retention test.

In summary, the above reviewed studies suggest that the advantage of acquiring external focus of attention when performing a supra-postural task would not only enhance the performance of that task but also would indirectly enhance the performance of the postural task as well. The assumption behind this phenomenon is that "the motor system seems to be able to automatically adjust posture to the demands of other tasks that it subserves" (McNevin & Wulf, 2002, p.195).

The next section reviews studies that examined the role of FOA on novices and experts as they perform motor tasks.

4.1.3.4 Focus of Attention in Novices and Experts

This section will review studies that tested the generalizability of the effectiveness of EFOA on performers with different expertise levels. Wulf, McConnel, Gartner, and Schwarz (2002) examined whether the focus of attention would show differential effects depending on the performers' skill level (novices or experts). In Experiment 1, the researchers assigned their novices and advanced volleyball players to either an external or an internal focus group. The participants' task was to perform a "tennis serve" in a volleyball setting. The researchers manipulated the focus of attention in this study by feedback sentences provided to the performers in response to the performers' serve.

The results showed that relative to internal focused feedback, performers—regardless of their expertise level—who received external focused feedback were more accurate in their serves during both practice and the retention test. Nonetheless, the novices and experienced players differed in terms of their movement forms, which judges rated according to standard criteria. For the experienced players, players in the external focus feedback group received higher scores during both practice and retention as compared to players in the internal focus feedback group. For the novice performers, the movement forms of the external focus group, during practice, were scored higher as compared to those in the internal focus feedback group, during practice. However, during the retention test, when no feedback was provided, the novices in the internal feedback group performed with high movement forms; the enhancement in their performance was attributed to the absence of the degrading effect of the internal feedback during the retention test. This finding was replicated in Experiment 2 of Wulf, et al. (2002) that used a "lofted soccer pass" as a task.

Accordingly, Wulf et al. (2002) concluded that these results mirrored the findings of Shea and Wulf (1999) on the advantage of external focus of attention during both practice and retention. They further asserted that the external focus of attention is beneficial for all performers, regardless of their skill level. However, in this study, the enhancement in the novices' performance during the retention test might be due to an advantage of the internal feedback that caused them to correct their movements with practice.

Another study recently examined the role of FOA on experts during a golf shot performance (Wulf and Su, 2007). Six expert golfers participated in the researchers' second experiment. The researchers defined the level of expertise by utilizing the players' handicap scores. The task of the players was to hit a golf ball into a target hole. The performance was measured in terms of the players' accuracy in hitting this target. In order to make this task challenging for the experts, the size of the target was smaller than the standard golf hole. In a within-subject design, all participants performed under three conditions: EFOA, IFOA, and a control condition. Participants were instructed to focus either on the club motion, their hand movement, or on what they normally focus on when they play golf during the EFOA, IFOA and control conditions, respectively. The golfers completed 20 trials under each condition. Results indicated that the different conditions had substantial effect on performance: the golfers showed significantly more accuracy under the EFOA condition than under the similar IFOA and control conditions. Wulf and Su concluded from these results that the advantage of external focus also generalizes to experts. Because this study lacked a manipulation check, the researchers assumed that the golfers during the control conditions adopted their normal focus during their practice. This assumption was not beneficial to the study because it raises questions whether or not the experts usually focus on the most advantageous aspects of their sport, and if they do not, about how their strong performance (judged by their handicap score) can be explained.

Contrary to the findings of Wulf and Sue (2007), Perkins-Ceccato, Passmore and Lee (2003) found that FOA had differential effects, depending on the performers' skill level. Perkins-Ceccato Passmore and Lee (2003) examined whether or not the attentional focus would have different effects on the performance of novice and experienced golfers. In a within-subject design, the researchers compared the performance of 10 novice golfers to the performance of 10 experienced golfers under two attentional focus conditions: external and internal focus of attention. The participants in the external focus condition were instructed to focus on hitting the ball into the hole, while participants in the internal focus condition were instructed to focus on applying appropriate force during the golf swing. The researchers reminded the participants of their current focus condition by repeating the instructions before every fourth trial. The order of the conditions was counter-balanced between subjects; in each condition, the participants practiced 40 shots. The participants' accuracy was measured by calculating both average and variable error scores for the landing points of the ball from the target. After each trial, the participants were requested to rate their performance on the pervious trial on a five point scale by guessing how far the ball landed from the target in the external condition and how much force they thought they applied in the internal condition. The results showed that the novice performers had significantly higher average error scores as well as higher variable error scores as compared to the experienced golfers. More interestingly, the FOA showed a differential effect on performance, depending on the skill level. This effect was indicated by a significant two-way interaction between the skill level and the FOA. That is, the highly skilled golfers performed with lower variable error scores under the EFOA condition, whereas participants with low golf skills performed with lower variable error scores under the IFOA condition.

Perkins-Ceccato, Passmore, and Lee (2003) interpreted these results as demonstrating an interaction between FOA and the performers' skill level in golf; the skilled performers benefited from an EFOA, while the novice performers benefited more from an IFOA. Wulf and Su (2007) attributed the inconsistency between their findings and the results of Perkins-Ceccato, Passmore, and Lee (2003) to the vague instructions provided to the participants in the Perkins-Ceccato study. These unclear instructions might have created “a confound between information content and attentional focus” (Wulf & Su, 2007, P. 386).

The above discussed studies are among the few studies that looked at the possible interaction between skill level and the type of focus of attention adopted during practice. Although both studies included participants with different experience levels, almost all of the participants in the Wulf and Su study had a handicap score of zero, while those in the Perkins-Ceccato et al. had handicap scores of four for the highly experienced golfers and 26 for the golfers with low skills. In addition, the two studies differed in what dependent variable was used to evaluate the performance. Wulf and Su measured the accuracy scores, while Perkins-Ceccato, Passmore and Lee measured the participants’ variability error scores. Moreover, no control group was tested in either study. Due to these methodological differences and the different expertise levels of the participants in these studies, the nature of the interaction between skill level and the type of focus of attention is inconclusive. Therefore, more research is needed to clarify the nature of this important potential interaction.

4.1.3.5 FOA in other motor skills

More recently, Porter, Nolan, Ostrowski, and Wulf (2010) wanted to examine whether the EFOA benefits would generalize to include a task that requires agility. Agility was defined in this study as “the ability to change the direction of the body rapidly using a combination of strength, speed, balance, and coordination“ (Porter, et al., p. 2). As a secondary objective, the researchers sought to know whether the participants followed the FOA instructions; for this purpose, the researchers utilized a questionnaire as a manipulation check. In a within-subject design, the researchers required 20 young adults to perform an agility task under three conditions: control condition, EFOA condition, and IFOA condition.

The participant task was to perform the agility “L” test; for this test, the participants had to run a course of two five-meter long paths, connected at a right angle, as quickly as they could and with maximum effort. The researchers instructed their participants in all conditions to run the path with maximum speed and effort; however, the researchers provided extra instructions during the EFOA and IFOA conditions. In the EFOA, the researchers instructed the participants to focus on running towards the cones as fast as possible and to focus on pushing the ground with maximum effort; in the IFOA condition, the researchers instructed the participants to focus on moving their legs as fast as possible and to focus on planting their feet with maximum effort. All participants first performed under the control condition with no extra instructions. The order of the other two conditions was counterbalanced among participants.

The participants performed 15 trials under each condition, and they performed each condition in one day. The practice was completed in three non-consecutive days. Before each condition, the researchers asked the participants to repeat the instructions for their condition. During the practice, after each trial within each condition, the participants took a short break during which they were asked to answer the following question: “What were you focusing on during the previous trial? If you did not focus on anything particular during the trial, please leave the question blank” (Porter et al., 2010, p. 3). Responses to the questions were collected for qualitative analysis. The participants’ movement time was measured as the dependent variable.

The findings of this study demonstrated that the movement time was significantly shorter —illustrating that the participants were faster—under the EFOA condition than under both the control and IFOA conditions, the results of control and IFOA conditions were not different from one another. According to the researchers, this result reinforced the benefits of adopting EFOA on performance and extended the generalizability of the beneficial effect of EFOA to an agility task. For the qualitative analysis of the questions, the researchers aggregated the responses of all participants in each condition. The results showed that participants complied with the instruction in 67% of the trials for the EFOA and in 76% of the trials in the IFOA. In the remaining trials, the participants' responses indicated a switch between focusing internally and externally or focusing on other aspects of the task. During the control condition, the participants' responded that they focused externally on 13% of the trials, focused internally on 10% of the trials, and focused on other aspects of the task on 77% of the trials. The researchers found these results to be interesting since in both the literature and their study, the performance on the dependent variable was comparable in both the internal focus condition and the control condition. One proposed hypothesis for the comparable results was that when the participants are not instructed to focus either externally or internally, they would choose to focus internally. As conjectured by Porter et al. (2010), "[t]he results of the current study suggest that this may not be the case; in fact the current results revealed that participants in the CON [control] condition focused internally only 10% of the time" (p.6)**.** Thus, the researchers emphasized the role of manipulation check in understanding what the participants focused on during their assigned conditions. The study included a questionnaire to investigate what the participants focused on during performance following the instructions; this study is only one of a few FOA studies (Fasoli, et al. 2002; Vuillerme & Nafati, 2007) that have attempted to employ a manipulation check on the independent variable.

Most of the researchers investigated the effect of FOA on movement outcome. However, Lohse, Sherwood, and Healy (2010) sought to study the effect of FOA on movement kinematics and on movement outcome in an attempt to provide a detailed picture of the effect of FOA on performance. Twelve students volunteered to participant in this study. In a within-subject design, the participants were required to throw a dart towards a bristle dart board under three conditions: acquisition phase, external focus condition, and internal focus condition. During the acquisition phase, the researchers familiarized the participants with the task and instructed the participants to perform the task as “accurately and consistently as they could to the center of the board” (Loshe et al., p. 548). During the external and internal focus conditions, the researchers instructed the participants to either focus on the flight of the dart or on the movement of their arm. The order of the external and internal conditions was counterbalanced between participants. These instructions were repeated after every block of trials. During each condition, participants performed seven blocks, three trials each.

Lohse et al. (2010) measured the accuracy of the participants’ performance by calculating an error measure, the distance from where the dart landed and the dart board center. Researchers also measured the time between the two trials to indicate the preparation time. Furthermore, from the onset of the arm movement to the dart release, the researchers utilized surface electromyography EMG (Integrated EMG) to measure the muscular activity of each participant’s biceps, biceps brachii and triceps muscles. The researchers employed Dart Fish Connect Pro-Motion Analysis software for the movement kinematic analysis of shoulder angle, elbow flexion, throwing time, and angular velocity of the dart.

The results indicate that, during the external focus condition, the participants 1) had significantly less error and 2) required significantly less preparation time than those in the internal focus condition. Although integrated EMG activity was less in both the biceps and the triceps muscles during the external focus than during the internal focus, this difference was not significant. No significant difference in terms of kinematic measures between conditions emerged. The researchers interpreted their results as being consistent with previous findings in terms of the improved accuracy and decreased EMG activity (in the agonist muscle) in the external focus condition compared with the internal focus condition. The researchers attributed their non-significant findings in the EMG and the kinematic measures to methodological issues and the task employed.

As the above reviewed studies suggest, most of the empirical evidence regarding the effectiveness of EFOA on motor skill performance and learning has been derived from studies of young and healthy adults learning motor skills. Researchers have recently extended the FOA literature by investigating the FOA notion in special populations and in rehabilitation settings.

4.1.3.6 Focus of Attention and Special Populations

Chiviacowsky, Wulf, and Wally (2010) investigated whether the FOA effect would extend to the older adult population; the researchers included thirty-two participants in the age range of 60-85 years old. Employing the stabilometer task, the researchers asked their participants in the EFOA group to focus on keeping the markers on the platform on the same horizontal level and those in the IFOA group to focus on keeping their feet on the same level. The researchers, who required their participants to look straight and only focus their attention as instructed, reminded the participants about these instructions before each trial. The participants practiced for 10 trials and received feedback on their performance as time in balance after each trial. One day after practice, the participants performed a retention test of five trials with no reminder of instructions or feedback. The results showed that, as compared to the IFOA group, the EFOA group maintained balance for a longer time during the retention test. Chiviacowsky et al. concluded that these results agree with and replicate previous studies on the beneficial effect of EFOA and extend the findings to include the older adult population.

Wulf and colleagues also extended the scope of their research on attention focus to include special populations. Wulf, Landers, Lewthwaite, and Töllner (2009) wanted to find out whether their findings would extend to people with Parkinson's disease (PD) and to discover whether external focus of attention would help this population maintain a good balance. Therefore, they examined 14 participants diagnosed with idiopathic PD, stages II or III according to Hoehn and Yahr stages of PD, without any dyskinesia and with minimal to moderate balance impairment. Although the participants were independent in ambulation, seven out of the 14 had a history of falling in the past year. Wulf et al. assessed the participants' ability to follow instructions and directions prior to the experiment during the interview and the consent procedure. They then asked the participants to balance on an "inflated rubber disk" under three conditions: 1) IFOA condition—in which the participants were instructed to focus on minimizing their feet movement; 2) EFOA condition—in which the participants were instructed to focus on minimizing the movement of the rubber disc; and 3) control condition—in which the participants were instructed to focus on standing still. Participants performed three trials under each condition; the participants' postural sway RMSE was measured. The advantage of adopting external focus of attention was apparent, as the participants showed significantly lower postural sway under the EFOA condition when compared to both the IFOA and the control conditions—the results of which did not differ from each other.

Wulf et al. (2009) concluded that their results reinforce the previous findings and generalize the benefits of EFOA to people with PD; the researchers also recommended the implementation of this focus during balance training in rehabilitation settings. In this study, patients in the control condition were told to focus on standing still. In a previous study, Vuillerme and Nafati (2007) noted that increased attention to swaying of one's body in the context of quiet standing "promoted the use of less automatic control processes and hampered the efficiency for controlling posture during quiet standing" (p. 192). Hence, the similarity of the of results in the IFOA condition and the control condition might be overstated in Wulf et al. (2009) because participants may have adopted the same focus of attention under both the internal and control conditions. Without a questionnaire or direct assessment, the researchers had little evidence for what the participants focused on in each condition.

Researchers have also acknowledged the importance of engaging participants' in EFOA during rehabilitation. For example, in their review article, Wulf and Prinz (2001) noted that occupational therapists have begun to incorporate the notion of EFOA in their treatments. Studies have found that patients recover more quickly when engaged in purposeful activities as opposed to exercise, because purposeful activities direct the performers' attention to the effect of their movement, not to their movements. According to the theory of purposeful activity, using real events in which to retrain function is better than using meaningless exercises and simulation. The purposeful activity theory agrees with the notion that "motor control is facilitated by placing attention on the goal, rather than on the movements themselves" (Wulf & Prinz, p. 657). Since 1985, when Gliner reported that it was better to focus the actor's attention "to a particular object rather than to the internal aspects of the act" (as cited in Wulf & Prinz, 2001, p. 657), occupational therapists have supported the concept of purposeful activity.

Fasoli, Trombly, Tickle-Degnen, and Verfaellie (2002) investigated the effect of EFOA and IFOA instructions on performing three functional reaching tasks—utilized during occupational therapy—by individuals with cerebrovascular accident (CVA). The researchers included 16 participants who had CVA that affected an upper limb and 17 age-matched adults; they then compared the performance of the two groups. Before starting the experiment, the researchers screened the potential participants with CVA as follows: 1) the administration of the shortened version of the Token Test for auditory comprehension to ensure that they could understand verbal instruction; 2) the administration of the Florida Apraxia Screening Test to ensure that the participants did not suffer from motor planning impairments; 3) the administration of the Modified Ashworth Scale to evaluate upper limbs for spasticity; 4) the administration of the letter cancellation and visual extinction test researchers to screen potential participants for visual neglect; and 5) the administration of the Perception of Joint Position Sense Test to assess the participants' awareness of the involved limb. Only 16 participants who scored within normal range in the screening tests participated in the study: six participants with left-CVA, nine participants with right-CVA, and one participant with bilateral infarcts; their post-stroke time ranged from six months to 32 years.

The researchers required the participants to perform three functional reaching tasks: 1) "removing a can from a shelf and placing it on the table"; 2) " taking an apple off a shelf and putting it into a basket"; and 3) " moving an empty coffee mug from the table onto a saucer". The participants performed the three tasks under two conditions: EFOA and IFOA; the order of these conditions was counterbalanced among the participants. Therefore, this experiment consisted of four groups according to the participants' group whether or not the participants had CVA and the order of the conditions: 1) participants with CVA (EFOA-IFOA); 2) participants with CVA (IFOA-EFOA); 3) participants without CVA (EFOA-IFOA); and 4) participants without CVA (IFOA-EFOA). The order of the three tasks was randomized; the participants practiced the tasks in the same order under both conditions.

Fasoli et al. (2002) analyzed the quality of the participants' movement. The researchers attached a light-emitting diode (LED) to the participants' hand (the affected hand in the participants with CVA and the same hand in the age-matched control group) to analyze the participants' movements utilizing an OPTOTRAK 3020. The dependent variables in this study included: movement time (to measure the overall speed of the movement); peak velocity (to indicate the force produced during the reach); movement units (a higher number of movement units indicates a more guided movement, while a smaller number of movement units indicates a smoother movement); and the percentage of time to peak velocity (a percentage in the range of 33%-55% indicates a continuous movement and also indicates that the movement was pre-planned).

To ensure that the participants understood the task, all participants performed one baseline trial of each task, during which the researchers provided only general instructions about the goal of the movement. After completing all baseline trials, the participants answered questions regarding whether they attended to their movements or to the task itself. Following the completion of the baseline trials and the questionnaire, the researchers instructed the participants according to their assigned condition. For the EFOA condition, the researchers instructed the participants to pay attention to the size and weight of the can and the can's position on the shelf. They instructed the participants in the IFOA condition, to pay attention to their arms and to think about how their elbow straightens and how their wrist and fingers move. The researchers reminded the participants about these instructions during the practice. The participants performed eight trials for each task (three reaching tasks) under the two conditions (EFOA and IFOA) for a total of 48 trials. After the participants had completed their trials (24 trials) under one experimental condition, the researchers re-administered the same questionnaire that the participants had answered after the baseline trial, as a manipulation check.

To analyze the data, the researchers used three-way ANOVAs with the group and sequence of conditions as the between factors and the order of the tasks as the within factor. For each group, the results indicated that participants in the EFOA, when compared to those in the IFOA condition, demonstrated the following: significantly shorter movement time (faster movement), significantly greater peak velocity (more forceful movement), and fewer movement units (smoother movements during the can task). Although the percentage of time to peak velocity did not differ between the two conditions for the participants with CVA, the percentage of time to peak velocity in the control group was significantly higher during both the can and the mug tasks; this indicates more preplanned movements in the EFOA conditions.

When the researchers compared the performance of the two groups, they found that participants in the CVA group performed the task with significantly longer movement times (were slower), lower peak velocity (less forceful movement), and more movement units (fewer smooth and more guided movements) than those participants in the control group.

Fasoli et al. (2002) interpreted these results as replicating the previous findings about the beneficial effect of adopting EFOA on performance. The results showed that not only did the beneficial effects extend to participants with CVA, but also that the performance of a common reaching task, which is used frequently in occupational therapy, could benefit from the adoption of EFOA. The researchers further noted that, while the FOA seemed to affect some aspects of movement (the movements' speed and the force applied), the FOA did not affect the percentage of time to peak velocity. Consequently, the researchers suggested that further investigation is needed to clarify such differences in FOA effects on movement characteristics.

These two studies (Fasoli et al., 2002; Wulf et al., 2009) examined only performance, not learning. It would have been more enlightening if the researchers had extended their measures to include a retention test. Furthermore, although the literature is consistent about the beneficial effects of adopting an EFOA in young healthy people, these studies highlight that there is insufficient evidence to apply this construct to populations such as those with various neurological disorders. This area of research needs further investigation to clarify the generalizability of findings and whether or not FOA would have clinical importance.

McAlister (2006) recently applied the notion of attentional focus in a simulated occupational therapy rehabilitation setting, young adults using prosthesis. McAlister study was motivated by the fact that most occupational therapists either have patients focus on the quality of the movement of prostheses or use videotape to provide feedback. While the literature suggests these internally focused methods may actually hinder recovery. McAlister (2006) argues that while the impact of focus on learning motor skills has been extensively studied, no studies "have specifically studied the influence of internal or external focused instructions…upon amputees who seek to learn how to use a prosthetic device" (p. 62).

In the McAlister study, thirty college students were randomly assigned to one of two groups: EFOA group or IFOA group. Using prostheses, they had to perform a novel task—pouring cereal into a bowl. The instructions, presented on a video tape, directed the participants to focus externally or internally according to their group assignment. The results showed that, in a novel task such as using upper limb prosthesis, "the external focus group spilled significantly less cereal during both skill acquisition and skill retention sessions" (McAlister, p. 7). These results confirm the advantages of EFOA over IFOA on the learning of a novel task (pouring cereal in a bowl with prosthesis) in a simulated occupational therapy setting. McAlister supports the purposeful activity theory by stating that "if tasks are meaningful to patients and can be applied to daily living, their attention is focused on the meaning of the task as opposed to the required internal movements" (p. 55). According to McAlister, this finding has clinical implications for the amputee population because half of all subjects with prostheses are unable to use their prosthesis effectively (McAlister, 2006). McAlister (2006) surmised that "instructions that are based on goal achievement could be given to rehabilitation patients when learning a motor task" in order to improve the rehabilitation outcomes (p. 43). Overall, the study offered evidence that external attentional focus may improve outcomes for learning novel tasks in a simulated therapy session, extending the findings from balance and sport motor skill learning.

Because the data have yielded promising results that support EFOA, more researchers have indicated an intention to explore the use of FOA in their everyday clinical practice. For example, in a physiotherapy treatment, Laufer, Rotem-Lehrer, Ronen, Khayutin, and Rozenberg (2007) examined the effect of FOA induced by instructions during dynamic balance training in individuals with an ankle sprain. Forty young adults with a sprained ankle participated in this study. The researchers randomly assigned the participants to either an EFOA or IFOA group. The participants' task was to balance on their injured limb in a dynamical balance task, using a Biodex Stability System - BSS. Those in the EFOA group had to maintain their balance by stabilizing the platform, while those in the IFOA group had to maintain their balance by stabilizing their body.

The participants practiced ten trials per day for three days. Two days after completing their practice trials, the participants performed a retention test. The participants' postural stability was measured by three stability indices: 1) Overall Stability Index (OSI), which indicates the overall variation of the platform—a high value indicates greater movement and less stability; 2)

Anterior/Posterior Stability Index (APSI), which measures the platform stability in the sagittal plane; and 3) Medial/Lateral Stability Index (MLSI), which measures the platform stability in the frontal plane. The results indicated that the EFOA group was more stable during practice; this higher stability was maintained during the retention test as indicated by the APSI and the OSI indices. Laufer et al. (2007) interpreted the study results as being in agreement with previous studies. That is, there was an advantage for both performance and learning of adopting an EFOA. Moreover, the results of this study extend this advantage of EFOA to physical therapy— when treating participants with a sprained ankle.

Durham, Van Vliet, Badger, and Sackley (2009), in a similar setting, examined the FOA notion in the context of feedback given by physiotherapists to their clients. More specifically, the researchers sought to quantify the use of feedback and instructions during a real treatment session and to look at the communication content in terms of attentional focus information. The study included evaluation of the treatment session of eight physiotherapists and eight clients. The clients had a hemiplegic arm and were receiving physiotherapy treatment at the time of the study. The researchers collected their data by: 1) videotaping the treatment session; 2) interviewing physiotherapists and clients; and 3) asking the therapists to complete a questionnaire. The verbal communication between the therapists and their clients was analyzed and categorized as being instruction, feedback, or motivation sentences. Furthermore, the feedback statements were categorized as external, internal, or mixed in terms of what attention of focus information they contained. Collectively, a total of 1,914 statements were identified from the eight treatment sessions. The results revealed the following: instruction was the most frequent type of communication; motivation statements were the second most frequent type of communication; feedback statements were used the least. The collected statements showed a quantitative composition of 54%, 33% and 13%, respectively. Surprisingly, almost all the feedback statements (96%) were classified as being internally focused, while the remaining feedback statements (4%) were ambiguous with a mixed focus. Similarly, 75% of the instructions provided to the clients were internally focused. From these results, the researchers concluded that the motivational statements provided by the physical therapists were the primary sort communication during treatment; the instructions or feedback the therapists provided were usually internally focused. Because the purpose of this study was not to determine the effectiveness of these communication methods on the treatment outcome, it is not clear whether the therapists were attaining positive results according to their communication. Given that abundant studies in this area to date have indicated that external focus is more beneficial to both performance and learning, Durham et al. (2009) conjectured that the physical therapists "may be providing less than optimal feedback by not making more use of the information with an external focus" (p. 88). However, they noted that "it is unclear if results from healthy subjects apply to the neurologically impaired" (p.78). Although this finding cannot be generalized to other settings and therapists, this study reflected some bias in utilizing internally focused statements (instructions and feedback) in treatment sessions (Durham et al., 2009). Therefore, more research that is directed toward studying this phenomenon in people with neurological diseases is needed.

Like physical therapists, speech-language pathologists extensively use feedback and instructions during therapy. Therefore, the bias toward internal focus communication reported by Durham et al. (2009) during a physical therapy setting might apply to speech therapy sessions as well.

During treatment sessions, the speech pathologists usually make use of many therapeutic approaches, such as articulatory placement cues, phonetic derivation and sound contrast practice, etc., to help their clients improve their speech production. As surmised by Schulz (2007), “Such therapy techniques can be grouped together by their ‘internal focus’” (as cited in Wulf, 2007a, p. 184). She also added that given the inconclusive treatment effects of using the above techniques, it is worth mentioning that a study designed to examine the effects focus of attention notion in the speech domain is needed.

Although speech pathologists extensively use instructions and feedback during treatment sessions, they often overlook the effects from the focus of attention perspective. At present, no study has addressed this issue, despite the emerging evidence that learners benefit from external focus of attention instructions and feedback in terms of performance and learning in the acquisition of motor skills. The likelihood that results regarding attentional focus in motor limb studies can be extended to include speech tasks is heightened by the fact that research on the speech motor control system has begun to blur the boundaries between the limb motor literature and the speech production system. The speech motor control system has been conceptualized as sharing a motor system with the non-speech oral-motor system (e.g., Ballard, Robin, & Folkins, 2003; for contrary view, see; Ziegler, 2003a, b).

Taking the concept of the attention focus in a new direction, Freedman, Maas, Caligiuri, Wulf, and Robin (2007) argued in favor of extrapolating ideas from the limb motor learning literature into the speech motor learning domain. In order to confirm this connection, Freedman et al. (2007) studied the effects of external versus internal focus of attention on both limb motor and oral-motor movements to establish whether the attentional focus notion could be applied to the oral-facial effector system underlying speech production. The researchers sought to determine whether findings from the limb literature could be extended to the oral-facial system of the tongue, “as [a] first step in understanding the role of attentional focus in producing and learning oral movements, and eventually speech” (Freedman et al., p. 135). The researchers randomly assigned their 46 participants to either an EFOA group or an IFOA group. The participants’ task was to perform both a manual pressure task and an oral-motor pressure task using the Iowa Oral Performance Instrument (IOPI). The researchers instructed the participants to either focus on the pressure they apply on the IOPI’s rubber bulb, or to focus on the pressure they apply with their hand or tongue in the EFOA and the IFOA groups, respectively.

The results showed that the instructions had a substantial effect on performance: the performance of the EFOA group was significantly more accurate and significantly less variable than the performance of the IFOA group. Moreover, the effector had a main effect: the participants’ performance on the tongue task was significantly more accurate and significantly less variable than the participants’ performance on the hand task.

Overall, the study supported the constrained action hypothesis underlying support of external attentional focus by “replicating previous limb findings (hand task) and extending them to the oral-facial system (tongue task), in that the performance for both effectors was enhanced with an external focus of attention” (Freedman et al., p. 135).

Conceding that they only studied a non-speech tongue movement task, Freedman et al. (2007) also conjectured that the results may indicate that attentional focus “is an important variable to consider in treatment of speech disorders” (p. 135). The results of Freedman and his colleagues are encouraging as they demonstrate that attentional focus might be a potential factor that should be considered during speech treatment. However, the Freedman et al. study did not include a control group with which to compare both groups. Moreover, their conclusion was based only on performance measures, so it is not clear whether the results would be maintained after the practice phase. Therefore, the researchers recommended further research on how external attentional focus can help in learning a novel speech task.

Although the researchers do acknowledge that the future implications of their research might focus on performing a speech task, currently no study has attempted to examine the connection between speech production and focus of attention. Therefore, the objective of this study was to examine the role of EFOA and IFOA in learning of a novel speech task.

The next section provides a summary of and highlights some issues noticed in the reviewed literature.

4.1.4 Summary and Discussion

This section highlights some methodological issues observed in the reviewed studies. Moreover, this section also identifies research areas that require additional investigation in the FOA literature.

The extensive research about the role of focus of attention on motor skills has established convincing findings regarding the advantage of acquiring an external focus of attention on the performance and learning of motor skills (e.g., McNevin, Shea, & Wulf, 2003; Shea & Wulf, 1999; Wulf, HoB, & Prinz, 1998; Wulf, McNevin, & Shea, 2001; Vance, Wulf, Tollner, McNevin, & Mercer, 2004). External focus of attention had a superior effect on diverse of motor skills (McNevin, & Wulf, 2002; Wulf, McConnel, Gartner, & Schwarz, 2002; Wulf & Su, 2007; Zachry, Wulf, Mercer, & Bezodis, 2005). Having established convincing findings, researchers now seek to extend these findings in motor skills to other fields and to special populations (Brydges, Dubrowski, & Carnahan, 2007; McAlister, 2006; McNevin, Wulf, & Carlson, 2000;

Wulf, Landers, Lewthwaite, & Töllner, 2009). However, in reviewing these studies and the theoretical account proposed to explain the FOA construct, the following points were noted:

First, in order to advance knowledge about how FOA affects motor performance and learning, there should be consensus among the researchers on the operational definition of the levels of the independent variable (what is considered as externally or internally directing the participants' FOA). As discussed earlier, an instruction that might have been considered by one research group as inducing an internal focus (Perkins-Ceccato, et al., 2003) was interpreted as inducing an external focus or vague by others (Wulf and Su, 2007).

Second, the focus of attention can be classified as a conceptual variable that is induced by either an instruction or feedback. This situation requires the researchers to check whether their participants obtained—or understood—that particular FOA according to their group or condition. Almost all studies lacked a manipulation check; the researchers based their conclusion on the assumption that the participants adopted the intended condition. An independent variable manipulation check was only present in few studies (Fasoli, et al. 2002; Porter et al. 2010; Vuillerme & Nafati, 2007). In the Vuillerme & Nafati study (2007), considering a manipulation check was helpful, as it was the only way to determine that the independent variable was manipulated. As a manipulation check, Vuillerme and Nafati (2007) asked their participants to report how much they engaged during their performance according to the instructed FOA. The researchers also measured the participants' probe reaction time under both conditions (IFOA and the control condition). Bund, Wiemeyer, and Angert (2007) commented on this issue by noting that in some of the experimental tasks studied, "It is not clear which focus of attention the participants really adopt," and thus they call for more validation of this point (p. 17). Perhaps a manipulation check did not seem to be very important in most of the studies in the FOA literature, specifically in those studies in which the predicted effect of the independent variable (FOA conditions) on the dependent variable, as the performance or learning was demonstrated with no alternative explanations. Future studies would benefit from the inclusion of some kind of manipulation check for the independent variable.

Third, it was also noticed that several studies focused on the differential effect of FOA, depending on the skill level, with inconsistent findings among studies (e.g., Perkins-Ceccato, et al, 2003; Wulf, McConnel, Gartner, & Schwarz, 2002; Wulf & Su, 2007). Therefore, more studies are required to resolve this discrepancy and to clarify whether adopting an external focus of attention is beneficial primarily for expert performers, or for both novices and experts.

Fourth, although most of the studies reviewed above reported significant difference between an external and internal focus of attention conditions in terms of performance—during either practice, retention, or both, only few studies reported an effect size. An effect size helps the researchers as well as research consumers to decide whether the magnitude of the effect of the significant results is important and meaningful. In this literature, reporting an effect size is especially important, as researchers highly recommended the application of this FOA construct in rehabilitation (McNevin, Wulf, & Carlson, 2000), sports training, and other clinical disciplines (e.g., Brydges, Dubrowski, & Carnahan, 2007; Freedman, Maas, Caligiuri, Wulf, & Robin, 2007).

Fifth, the majority of the studies did not include a control group with which other group results could be compared (e.g., Wulf et al., 1998, Experiment 1; Wulf & McNevin, 2003; Wulf, Weigelt, Poulter, & McNevin, 2003); As such, "It seems to be unclear whether the benefit of an external versus internal focus is due to an advantage of external focus or a disadvantage of internal focus" (Zentgraf & Munzert; 2009, p.524). Even in those studies that included a control group, the results showed an advantage of EFOA and illustrated that IFOA did not differ from the control group. Although these results demonstrated that an external focus enhances motor performance, it is still not clear how adopting an EFOA of attention enhances learning, how an IFOA hampers performance, or in what way IFOA and control conditions can be comparable. This unresolved issue might be due to: 1) the type of measures (outcome measures) empirical utilized in this literature or 2) the proposed constrained action hypothesis—CAH—to account for this phenomenon, which lacked an explicit explanation at this point. These two issues are discussed next.

In the FOA studies, outcome measures—such as the calculated accuracy in reaching a target or measuring postural sway to indicate balance—have been used to make inferences about what the performers learned. Research that only utilizes such outcome measures to infer learning is based on a "product oriented definitions" of motor learning. As described by Robert (1997), research conducted with "product oriented definitions concentrate[s] on what motor task was learned and tend[s] to neglect how it comes about" (p. 20). Therefore, in order to understand how focusing externally or internally affects the learning process, future research should be geared toward investigating, in more details, any associated processes with learning. "Such task properties might consist of the motor program and associated parameters necessary for performance" (Robert, 1997, P. 21). While a handful of studies have more closely analyzed how the performers' movement differed in both EFOA and IFOA conditions (e.g., Lohse, Sherwood, & Healy, 2010; Zentgraf, & Munzert, 2009) , it is still unclear what mechanism affects learning under both EFOA and IFOA conditions. Not only did the studies neglect to specify this aspect, they also the proposed a CAH explanation for this notion. As stated by Zentgraf and Munzert (2009), "At present, the issue of when and how attentional foci affect motor performance is unresolved for numerous reasons. One problem is that the CAH lacks a clear notion of the characteristics of ‘natural’ motor-control processes, and these are rarely investigated in Wulf’s studies…..” (p. 521).

Furthermore, studies by Wulf and others have established that an EFOA is more effective than an IFOA in healthy individuals. However, most of these studies employed a controlled laboratory environment. Although the utilization of a laboratory task in a controlled environment has the advantage of more experimental control, it has the potential disadvantage of a limited generalizability to the rehabilitation setting or to situations outside of the experiment setting. As noted by Durham et al. (2009), “It is unclear if results from healthy subjects apply to the neurologically impaired” (p. 78). Although research has been conducted to investigate the effect of focus of attention on people with Parkinson’s disease, results are inconclusive. Moreover, as Durham et al. (2009) pointed out in their study, patients are provided with internal focus instructions during physical therapy rehabilitation. Therefore, more studies are needed that look at whether instructions given to clients in treatment or rehabilitation settings enhance or hamper their learning of motor skills.

In sum, this section has discussed areas that need future research in the FOA literature. In spite of the shortcomings, these results are interesting and important because they demonstrate the impact of changing the wording of the instructions or feedback has on performance and learning. As such, this dissertation was motivated by these results and examined the effect of EFOA and IFOA on learning in the speech domain. The next section addresses some characteristics of speech that render it a good candidate to examine the generalizability of the FOA effects.

4.2 THE VIABILITY OF TESTING THE FOA IN THE SPEECH DOMAIN

The objective of this proposed study was to investigate the differential effects of EFOA and IFOA on learning a novel speech task. The initial step necessitates considering whether speech serves as a viable candidate for testing the generalizability of FOA effects to the speech domain. Shriberg and Kent (2003, p.5) asserted the following:

> Physically, speech is both a pattern of the movement of the speech organs and a pattern of acoustic vibrations. Speech is most conveniently studied in physical terms by observing the movements of the speech structures (tongue, lips, jaw, and so on) and by recording the acoustic signal that the speech structures generate. Therefore, the study of sounds in a spoken language generally includes a description of how individual sounds are formed and information on the acoustic or auditory properties of a sound. (p. 5).

According to this description, it is obvious that the motor component of speech is acknowledged, however, the description also describes the difference between speech movements and most other movements.

Similarly, Clark, Yallop, and Fletcher (2007) described how speech can be approached from analysis viewpoint, as follows:

> Once we decide to begin an analysis of speech, we can approach it on various levels. At one level, speech is a matter of anatomy and physiology: we can study organs such as tongue and larynx and their function in the production of speech. Taking another perspective, we can focus on the speech sounds produced by these organs—the units that we commonly try to identify by letters, such as a 'b-sound' or an 'm-sound.' But speech is transmitted as sound waves, which means that we can also investigate the properties of the sound waves themselves. Taking yet another approach, the term 'sounds' is a reminder that speech is intended to be heard or perceived and that it is therefore possible to focus on the way in which a listener analyzes or processes a sound wave. (p. 1).

This understanding of speech and its many components validates the belief that speech can be comparable to other motor tasks studied in the FOA literature and thus can benefit from a study of attentional focus. The instructions provided to the performers to direct their attention to their movement or the effect of their movement can be applied when a speaker performs a speech task. In speech, an IFOA instruction would require the speaker to focus on the movement of the articulators—for example, movement of the tongue, the vibration of the vocal folds, or movement of the lips, while an EFOA would require the speaker to focus on the effect of these movements—the produced sound or acoustic signal or focus on the consequences for the movement by focusing on the receiver of the spoken words. This characteristic of speech enables the instructor to choose the suitable FOA when instructing the speaker.

Nonetheless, speech goes beyond being a simple motor task to being a complex task which requires coordinated movements of many muscles and joints. Moreover, additing to the motor complexity, speech should achieve a communicative goal as well. In the words of Brunner, Hoole and Perrier (2007) “Speech production is at the same time a semiotic and a motor task. As such it has to reach communicative objectives, while respecting the same constraints and rules of other skilled motor tasks carried out by humans” (p. 1). The current study examined the effect of FOA by utilizing both acoustic and perceptual analysis.

The above cited descriptions view speech as a multi-faceted activity. For example, in one speech study, participants may be instructed on how to move their articulators (tongue, lips, or larynx); in another study, the participants may be required to concentrate on achieving the acoustic output of the speech task. A third study may emphasize the communicative goal of what the individuals produced.

Wulf, Toller, and Shea (2007) suggested another point to consider. In order for FOA effects to manifest themselves, the task should be complex and pose a challenge to the performer. However, some studies did not demonstrate results consistent with the literature on FOA effects perhaps because these studies failed to utilize a task that met the required task complexity, such as using quiet standing task or walking (gait) task in normal subjects (e.g. Cohen, 2010; Vuillerme & Nafati, 2007). This belief that a task should be complex applies to speech. “…under most circumstances, speech is produced with an ease that belies the complexity of the operations underlying it” (Duffy, 2005, p. 3). Although speech emerges as a well-practiced task that resembles an automatic motor task, speech instead is an extremely complex task that entails a specific coordination between many systems, such as the respiratory, phonatory, articulatory and resonatory subsystems. Adding to this complexity is the required coordination and interaction of the motoric level with the linguistic and sensory information. All these complex processes must take place extremely fast in order for speech to be articulated in a correct way and to effectively convey its communicative goal though complex acoustic events. (More details about Speech motor control is discussed later).

Although the average speaker usually produces speech automatically, the speaker sometimes needs to approach speech in a more conscious and controlled way. For example, a person might produce a novel sound in a second language or the person might have a compromised speech motor control, requiring the speaker to relearn the movement pattern or an acoustically acceptable alternative. Such speech processing can be either automatic or controlled. Schneider & Shiffrin (1977) have defined these two types of processing. Major issues distinguish automatic and controlled processing. Automatic processing, which is not volitional, occurs at high speed, is independent (not affected by other tasks happening at the same time), has unlimited capability, is inflexible, and is parallel in nature. On the other hand, controlled processing is volitional (can be stopped), slow, easily interrupted, demands a lot of attention, has restricted capacity, is flexible, and is serial in nature. At the beginning stage of learning or performing a new task, the processing of that task is slow and attention-demanding. The major benefit of controlled processing is its flexibility that can enable modification of the response to accomplish the task goal. Automatic and controlled processing can be viewed as occupying two opposite extremes on a continuum (Schneider & Shiffrin, 1977).

When performing a new task, a person might utilize a controlled form of speech processing. With practice, however, the person can carry out the task with less attention; through more practice, the individual shifts to an automatic processing mode, as in the case of an expert performer. An example of an automatic processing task is driving a car. Because driving is so automatic for most people, they can easily talk, while driving, without interfering with the way they drive.

Therefore, learning to produce a novel speech sound or sequence of sounds, utilizing a controlled processing, within a motor learning paradigm provides an effective way to study the role of FOA on learning a speech task. Importantly, during this learning process, the speaker's FOA can be directed (by the instruction provided by the examiner) to focus either on the movement of the articulators or to focus on the effect of the movement of the articulators—the produced sound.

The current study utilized a complex novel speech task—for native English speakers—that requires finite control of the vibratory patterns of the vocal folds. Moreover, the objective of this study was to examine whether the instructions provided to the learners—to direct their FOA externally or internally—would differentially affect their learning.

The primary objective of this study was to investigate the role of FOA in the speech domain. Therefore, the next section briefly describes how the abstract intended message that the speaker wants to communicate, is transformed and eventually handled by the speech motor system to be produced as sound.

4.3 MODELS OF SPEECH PRODUCTION

In seeking to apply ideas from motor skill learning to speech production, existing models of speech production must be accommodated. Speech is an act performed by humans to communicate information. In order to achieve this communicative goal, many processes must transform the intended message from a more abstract representation to a less abstract one that can be handled by the motor system (Dell, 1986; Levelt, Roelofs, & Meyer, 1999). Although different models of speech production agree that the intended abstract message goes through many levels in order to become less abstract, these models diverge when it comes to explaining the details of these processes (Dell, 1986; Garrett, 1982; Levelt, 1989, Levelt et al., 1999; Rapp & Goldrick, 2000, 2004; Shattuck-Hufnagel, 1979, 1992). Specifically, speech production models differ in the following three ways: 1) the details of the explanation provided; 2) whether or not these processes—levels—interact or communicate (via feedback or activation); and 3) the scope of the model's explanation (whether it explains normal speech production only or can also account for speech in different pathological conditions). Although this discussion provides a comparison of the general characteristics of these models, it is beyond the scope of this paper to describe every model. Only models considered as theoretical framework to the current study were discussed in detail.

4.3.1 The Levelt Model

Levelt's (1989, 1999) model is a comprehensive model which specifies four levels of processing: conceptual preparation, lexical selection, morphophonological encoding and syllabification, phonetic encoding and articulation. Levelt's model is best described as a multilevel, feed-forward model (unidirectional). The different levels communicate by a feed-forward and serial communication, in which each level's input, is transformed within that level and then this level output would become the input for the next level. According to this serial theory, the proposed steps would occur in successive, non-overlapping stages.

The first—highest—of the four levels is the conceptual preparation. At this level, the intended message that is to be spoken is translated into a lexical concept. Levelt defines the second level as the lexical selection, during which “retrieving a word, or more specifically a lemma, from the mental lexicon” occurs (Levelt, Roelofs, & Meyer, 1999, p.4). Lemma is the meaning or conception of an item. In other words, the lexical selection is a process of searching the mental lexicon for the most relevant (highly activated) lemma that would best describe the intended lexical concept. At this stage, the lemma is syntactically defined. The selected lemma then enters the third level, which is morpho-phonological encoding and syllabification. As an output of this level, the lemma is defined phonologically.

The morpho-phonological encoding and syllabification level defines the lemma by specifying its morphological makeup (the lemma morpheme is retrieved from the mental lexicon), metrical shape (lemma defined in terms of its number of syllables, stress patterns—only if uncommon, and segmental make-up (the spelling out of the word’s morphemes with labels to indicate their order). Levelt and his colleagues (1999) call attention to the fact that although the morphemes are spelled out at this stage, no syllabification occur at this level. Rather, syllabification takes place later in the process and is not stored in the mental lexicon. In order to create a maximally pronounceable syllable, the context within which the word appears is required before syllabification takes place. Up to this point of processing, the end result of the above levels is a word that is syntactically and phonologically defined with all its segments spelled out, but in its current abstract form cannot yet be handled by the speech motor system. The fourth level of levelt’s model is phonetic encoding, a critical step, which translates the abstract phonological encoding into gestural scores (goals to be achieved by the articulators) that could be performed by the articulators.

Although Levelt acknowledged that his theory does not explicitly explain the phonetic encoding level, he strongly suggested that an important computation takes place at this level; this computation is the gestural score for the phonological word. Levelt proposed the idea of the mental syllabary, which he defines as a "repository of gestural scores for the frequently used syllables of the language" (Levelt et al, (1999), p. 5). He assumed that during the phonetic encoding of a frequent syllable, the mental syllabary is accessed to retrieve the gestural score—which is assumed to be stored as a program in the memory—of the frequent syllable (Levelt, Roelofs, & Meyer, 1999). On the other hand, Levelt noted that the gestural score for a new or less frequent syllable is not stored in the syllabary but is assembled from its segments.

These gestural scores can overlap in time and be combined according to the context in which they appear. The end result of the phonetic encoding is a gestural score that can be handled by the articulators (the gestures can overlap and they are context dependent). Levelt acknowledged that it is beyond his theory to specify how the articulators function in translating the gestural scores into speech. Although Levelt used a feed-forward model, feedback plays an important role in self-monitoring to detect errors. This feedback originates from two levels in the model: 1) the level of phonological word (at the internal speech state before it is articulated) and 2) a feedback from the speaker's overt speech. As previously stated, models of language production can be compared and contrasted in terms of the levels they attempt to clarify in speech production processes and the extent of communication and interaction between these levels (whether levels are independent or influence each other and in which direction).

Garrett developed a multilevel model for the speech production system. His model also consists of four processing levels: functional, positional and phonetic before the final articulation can take place (Garrett, 1980; 1984). Unlike Levelt's model, whose goal is the production of a string of syllables defined in terms of their gestural scores that can be handled by the articulators, the goal of Garrett's model is to explain the processes involved in producing a larger unit (sentence production), not only syllables. In addition, some models were developed with a more narrow focus, which consider only one of the levels of speech production—the phonological level, such as that of Dell (1986) and Shattuck-Hufnagel (1979; 1987) models of speech production. Although the Dell model only explains the phonological level, it has incorporated many stages within this level (semantic, morphological, and phonological processing stages). Because all of these stages communicate with each other in both directions by feedback and feed-forward mechanism, the Dell model can be described as an interactive model. This is not the case in Levelt's model, in which the levels communicate in a feed-forward manner and feedback utilization is limited.

The other example of a narrow focused model is the Shattuck-Hufnagel (1979; 1987) model. This scan-copier model of phonologic assembly is composed of a serial ordering mechanism that consists of two independent levels of representation: slots and units or segments to fill the slots. The model also contains monitoring devices that aid the "copier" in keeping track of which units have been copied and in detecting errors. This tracking can be contrasted to how Levelt proposed tracking of the phoneme order in his model; according to Levelt, a segment or phoneme becomes available with a labeled link (with number) to indicate its correct order.

Moreover, unlike the Levelt model, which is based on reaction time studies (Levelt, 1989; Levelt et al., 1999), Dell and Garrett models of speech production share the characteristic of both being developed from speech error data (Dell, 1986; Garrett. 1982).

Although Levelt model did not specify further how the motor system handles the gestural scores or how articulation takes place, Levelt suggested that his model could be linked to a dynamical systems perspective (Browman & Goldstein, 1992; Kelso, Saltzman & Tuller, 1986; Kent, Adams, & Turner, 1996) to handle the gestural score and explain speech motor control.

When compared with other language production models, Levelt's model provides greater details on the pre-motor levels of speech production. Consequently, due to its comprehensive explanations, the Levelt model serves as the framework for this discussion—to account for pre-motor processes. The next section discusses some speech motor control models that accommodate the motor level of speech production.

4.4 MODELS OF SPEECH MOTOR CONTROL

Unfortunately, models that deal with the pre-motor aspect of speech production do not specify how the pre-motor processes could connect with the lower motor act of speech. This might give the false impression that language processing proceeds independently from speech production, which is not the case. Accordingly, this section provides an overview of models that examine how speech might be controlled.

People utilize speech as a major way of communication among them. As was mentioned before, "Under most circumstances, speech is produced with an ease that belies the complexity of the operations underlying it" (Duffy, 2005, p. 3). Although speech seems to flow smoothly and effortlessly, speech is instead a very complex performance that requires precise coordination among subsystems. Speech involves the activation of many muscles (in all the subsystems). Each of the involved muscles is capable of moving in diverse ways. Moreover, all the muscles are also capable of combining with each other in diverse ways in terms of temporal coordination, speed and direction. The involvement of a large number of muscles results in higher degrees of freedom. On the one hand, a system with a high degree of freedom is a flexible system that can accommodate different situations. However, extreme degrees of freedom challenge the control system that guides all the muscles to perform the desired action. It should be noted that the degree of freedom problem is common in the motor system and not unique to speech (Kent, Adams, & Turner, 1996).

Theories of motor control—from two different perspectives—agree that the degrees of freedom problem can be solved by limiting the freedom within which each muscle can move. These theories assume that some movement components can be controlled as single unit. Such an assumption reduces the need to restrain each muscle in order to move in the required direction for a specific movement (e.g., Abbs, Gracco, & Cole, 1984; Keele, 1968). Nonetheless, theories of motor control differ on how the controllable single unit emerges. The nature of this single unit has been explained from two perspectives: the motor program perspective and the dynamical system perspective. From the motor programming perspective, the single unit constitutes a specific direction for each muscle involved in that movement. It is defined and stored in the memory before the movement begins (e.g, generalized motor programs in Schema Theory as discussed later). If the performer initiates that movement, a central drive is responsible for activating this single unit, which controls the specific parameters of the muscles. However, according dynamical system theory, no such representation of a single unit exists in the memory. Rather, a single unit emerges naturally as the dynamical system moves to produce the movement. As the dynamical system interacts with the environment to produce a movement, the involved muscles form a single unit—coordinative structure that works together to achieve the movement goal (Kelso, Saltzman, & Tuller, 1986). Each perspective suggests a plausible solution for the degrees of freedom problem. Although the dynamical system perspective seems capable of providing a solution to the degrees of freedom problem that has strong supportive empirical evidence, it has little to say about motor learning (e.g., Schmidt, 2003; Schmidt & Lee, 2005). From the motor programming perspective, the Schema Theory (Schmidt, 1975) makes explicit prediction about motor learning. Consequently, the following sections discusses in more details the motor programming perspective; specifically the Schema Theory.

The input to speech motor control, which "refers to the systems and strategies that control the production of speech" (Kent, 2000, p.391), are phonological representations (i.e., phonemes and syllables) of the message the speaker would like to produce. Researchers have proposed two different reference frames that the neural control system might utilize to control speech (Kent, 2000): the gestural and the acoustic target. According to the gestural target perspective, the control system specifies the location and degree of constriction along the vocal tract to produce the intended speech sound (Saltzman & Munhall, 1989). The proponents of the acoustic reference frame assert that the vocal tract is shaped according to the intended acoustic goal or auditory goal (Guenther, 1995; Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico, Wozniak, & Guiod, 1997).

While many models were proposed to account for speech motor programming (Gracco & Abbs, 1986; Sternberg, Knoll, Monsell, Wright, 1988; Van der Merwe, 2008), the Van der Merwe model (2008) and the DIVA model of speech production are discussed in detail because they explicitly specify the levels of speech motor control and associate these levels with areas in the brain.

4.4.1 Four level speech production model

In an attempt to explain the transformation of the intended message to speech from a brain behavior viewpoint, Van der Merwe (2008) proposed a theoretical framework of speech sensorimotor control. As the name of the framework implies, sensorimotor integration is a key concept in this model. Van der Merwe appreciated the importance of an open loop system (a system which depends on a central program to control its movement independent of any feedback) in the control of speech production, but she also emphasized the value of the role of afferent feedback (tactile, proprioceptive or auditory feedback) in controlling and modifying the speech movements. Van der Merwe also acknowledged that “the exact nature of sensorimotor interface during all of these phases is not yet known” (Van der Merwe, 2008, p.6). Her model, dissimilar from the traditional view of three-stage model (consisting of linguistic encoding, programming, and execution), introduces a fourth stage (motor planning) at a higher level than the motor programming stage. The motor planning stage plays an essential role in guiding speech movement by constructing the general decisions about the motor goal. The motor programming stage then adds some specification and modification to accommodate these plans in a context dependent manner. In this model, programming and planning are two distinct stages and the two terms are not used interchangeably. Also, by differentiating the motor levels, Van der Merwe attributes different neurogenic speech disorders to specific levels in her model. This model also makes possible the correlation between dysfunction in speech disorders and a hypothetical brain regions because the model suggest the neural structures involved in each level.

This model divides the speech production system into four hypothetical levels (linguistic-symbolic planning, motor planning, motor programming, and execution) with a high degree of interactivity. The linguistic-symbolic planning, the first level of processing, is the only pre-motor level in this model. Although Van der Merwe acknowledged all the processes that the intended message goes through before it is phonologically defined, she assigned only the first level to all the pre-motor preparation of the intended message. It is not uncommon for speech motor control models to underspecify or totally neglect the linguistic (pre-motor) stages. The model attributes errors produced by individuals with aphasia to dysfunction at the linguistic-symbolic planning level.

While Van der Merwe assigned all linguistic processes to the first level in the model, she used the next three levels of her model to specify the motor aspects of speech production.

The second level in the model is a motor planning stage, the highest motor level. The input to this level is the phoneme. According to Van der Merwe, the first step is to transform this phoneme (unit of planning in her model) into a code that the motor system can handle. In order to achieve this transformation, “motor planning entails formulating the strategy of action by specifying motor goal” (Van der Merwe, 2008, p. 9). The motor goal for each phoneme in the utterance is specified in terms of the temporal and spatial characteristics of that phoneme which are considered to be invariant (core) motor plans.

Producing a phoneme requires 1) retrieving the phoneme motor plan from sensorimotor memory, 2) successively ordering the motor plans according to their position in the utterance and 3) adapting the motor plans in a context dependent manner (in the case of coarticulation or speech rate effect on the duration of the segments). As described by Van der Merwe the motor planning level is goal-oriented (to prepare a motor goal that would result in a specific speech sound that achieves an acoustic goal); in order to achieve the intended goal, an internal feedback is continuously utilized to monitor the extent of the adaptation to make sure that the adaptation will not lead to sound distortion of the planned phoneme. The end result of motor planning is a collection of smaller motor goals, which are articulator specific (lip rounding, jaw depression), sequentially arranged, and context adapted. According to Van der Merwe, dysfunction at the motor planning level would manifest itself by imprecise motor plans that might lead to sound distortion, as is the case in the speech of individuals with apraxia of speech (AOS); AOS impairment has a phonetic-motoric origin (McNeil, Pratt, & Fossett, 2004).

The motor programming level (the second motor level in the model) then transforms these broad motor plans from articulator specific to more detailed muscle specific motor programs. Van der Merwe assumes that the motor program provides detailed information about “the muscle tone, movement direction, velocity, force, range, as well as mechanical stiffness of the joints” (Van der Merwe, 2008, p.13) which should be ready before movement initiation. Moreover, motor programming is also responsible for sequencing the muscle specific motor programs in order to achieve the larger articulator goal (lip rounding) and also responsible for the initiation of these motor programs. The motor programs at this level can be controlled by internal feedback (from other brain regions) or updated by sensory or auditory feedback. The model attributes the difficulty in initiating speech movement noticed—in individuals with dysarthria associated with Parkinson’s disease—to dysfunction at the motor programming level.

The lowest level in the model is the execution (carrying out) of these motor programs. At this level the motor programs “is finally transformed into a non-learned automatic (reflex) motor adjustments” (Van der Merwe, 2008, p. 16) which are transferred to the final common pathways to activate the particular selected muscles. The model assumes that lower motor neuron dysarthria (Flaccid Dysarthria) is attributed to dysfunction of the execution level.

The ascription of specific neurogenic communication disorders to dysfunction at specific levels of the model is theoretical; Van der Merwe acknowledges that the interaction between different brain regions and the hierarchical nature of the regions' performance may result in dysfunction that manifests itself at more than one level. She provides a comprehensive and detailed description of brain areas that are involved in each processing level and shows how these areas might interact with each other (for more details, see Van der Merwe, 2008). While most speech production models discuss the speech production processes in intact systems, Van der Merwe also extends her explanation to account for pathological population such as individuals with aphasia, apraxia of speech, and dysarthria. Importantly, this model emphasizes the significance of sensory feedback in correcting and adjusting speech through its effect on motor programs (which is the notion of motor learning). The modification of the motor programs takes place when the system changes to closed loop and depends on feedback to modify its movement to meet the new requirement in the intended goal. In her words, "during motor learning the control mode is presumably predominantly based on feedback control, which aids in optimizing accuracy." (Van der Merwe, 2008, P. 5).

The four-level framework seems to specify in more detail the motor processes—planning and programming—of speech motor control. Moreover, the hypothesized planning and programming stages of this model can be directly related to the Schema Theory notion of GMP and its parameters (McNeil et al., 2004). Nonetheless, this four-level model has relatively little to say about speech motor learning—the paradigm that the current study utilized to study the role of FOA on speech production and learning. Therefore, a recent model of speech production—The Directions Into Velocities of Articulators (the DIVA model) is discussed next.

4.4.2 The DIVA model

The DIVA model is a neural network model of speech production, which simulate speech acquisition and production (e.g., Guenther, 1995; Guenther, Ghosh, & Tourville, 2006; Guenther & Perkell, 2004). The model describes the motor and sensory processes involved in speech production and learning. Moreover, the model simulates the interconnection among areas of the brain assumed to be involved in speech production. The DIVA model provides a specific framework that can be tested and simulated.

The DIVA model describes the complex transformation of a string of phonemes into an acoustic signal resulting from movements of the modeled articulators. Though, Guenther and his colleagues simplified this complex transformation (of the phonemes into articulatory movements) in their model, they demonstrated that the DIVA model can account for imaging results as well as the following speech phenomena: motor equivalence, coarticulation, and speech rate effects.

The primary purpose of the DIVA model is to attempt to account for how infants produce, acquire, and learn speech. According to the DIVA model, the speech sound is specified as convex regions in auditory/orosensory spaces (Guenther & Perkell, 2004). The representation of a speech sound as a convex region, rather than a point in the auditory/orosensory space, addresses the acceptable variability in the produced speech sound from both the articulatory and acoustic perspectives.

The model proposes four reference frames: 1) acoustic frame, 2) phonetic frame, 3) orosensory frame, and 4) movement (articulatory) frame. The acoustic frame consists of the sounds to which the infant (listener) is exposed. The phonetic frame comprises a transduced acoustic signal from the acoustic frame that the model learns to produce. The orosensory frame, also called a somatosensory frame, includes signals from the tactile and proprioceptive receptors from the vocal tract; this frame utilizes this sensory feedback to detect the configuration of the vocal tract and to determine which sound the speaker is producing. The articulatory frame describes the specific motor commands to the articulators' muscles. The DIVA model explicitly specifies how these four hypothetical frames correlate with the brain regions involved in speech motor control. Parameters govern the interaction among these frames; the interaction between two frames is designated "mapping." Each frame represents a collection of cells; synapsis connects the cells of separate frames. The weight of the synapsis at one point directs the interaction—"the mapping" —between the frames. The DIVA model also incorporates feedback and feed-forward communication among these frames.

According to Guenther (1995), the DIVA model assumes that babbling elicits learning in the model. During the babbling phase, parameters for the first two mappings are tuned: 1) orosensory-to-articulatory mapping and 2) phonetic-to-orosensory mapping. In the model's simulation, babbling is modeled by a random activation of the articulators though the activation of the articulatory frame. Receptors (tactile and proprioceptive) sense changes in the vocal tract configuration; they signal these changes as a feedback to the orosensory frame. This relationship between the motor and sensory parts of the babbling tunes the first mapping: the orosensory-to-articulatory mapping. After this relationship is established, the orosensory-to-articulatory mapping converts the desired target—specified in the orosensory space—to an acceptable articulatory movement. The Direction Into Velocities of Articulators model derives its name from this first mapping, The DIVA model emphasizes that it is the target directions—in the orosensory space—that lead to the movements of the articulators through the orosensory-articulatory mapping.

During babbling, infants are capable of recognizing speech sounds of their language—saved in the acoustic frame; as a result, each speech sound will be coded in the phonetic frame, also called the “speech sound map.” When the infant produces a speech sound during babbling, the cell representing this speech sound will be activated in the phonetic frame. As the vocal tract moves to produce the sound, the orosensory frame receives the sensory inputs. This relationship between the produced speech sounds and the sensory feedback from the vocal tract tunes the second mapping in the model: the phonetic-to-orosensory mapping. As the infant produces the sound over and over again, the orosensory frame will expand its acceptable orosensory space for that speech sound.

The establishment of these two mappings set the model for the performance phase. The performance phase of the DIVA model can be described as follows: First, the phoneme strings fed to the model activates cells in the phonetic frame “speech sound map.” Second, through the phonetic-to-orosensory mapping, the orosensory frame activates the orosensory targets learned for that phoneme. Third, through the orosensory-to-articulatory mapping, the orosensory targets transforms to appropriate movements of the articulators to produce the speech sound. Then, the orosensory frame receives input about the current configuration of the vocal tract and compares the feedback with the target configuration; if any discrepancy is detected, an error signal is sent back to the articulatory frame. Finally, after the model completes the production of the first phoneme, the cells in the phonetic frame that correspond to the next phoneme are activated, and the previous steps are repeated. Once the DIVA model learns the sounds through repeated production, a feed-forward control will govern the previous steps, independent of feedback.

The DIVA model is innovative in the way it connects its components and synapses to hypothesized areas in the brain; neurophysiological and neuroanatomical evidence indicates that these areas are involved in speech production. The correlation between the model components and the brain regions furthers the understanding of how speech is controlled. For example, Guenther proposes that the phonetic frame—speech sound map—is located in the pre-motor area that stores the motor plans of the speech sounds. Projections from the pre-motor area to the sensory areas—auditory cortical areas in the temporal gyrus and the orosensory areas in the somatosensory cortex—carry information about the anticipated sound pattern and from the anticipated sensory consequences of the articulators' movements.

During speech production, the auditory and somatosensory areas receive feedback about the actual sound pattern produced and about the sensory consequences of the vocal tract movements. Any discrepancy between the intended goal and the actual information would cause the sensory areas to produce an error signal; the error signal represents the preferred direction of movements in the vocal tract that will correct this error. Once the sound is learned, a feed-forward projection from the pre-motor area to the primary motor area will take over the control. This feed-forward control functions independently of any feedback. Nonetheless, while using the feed-forward control, a more frequently generated error signal would indicate a mismatch between the intended and the current state of the vocal tract. This mismatch necessitates an update of the parameters governing the mapping between different frames in the model in order to match the current change in the vocal tract. According to the DIVA model, this situation of parameters update is expected to occur as a result of normal developmental changes of the vocal tract size as a child grows. Also, persistent constraints or changes in the hearing status of the individual will generate such frequent error signals. As the error signal becomes zero, the updated parameters eventually become incorporated into an updated feed-forward control (e.g., Guenther & Perkell, 2004).

Although the model learns to produce each phoneme independently, the model is capable of producing the articulatory movements of the phoneme string in a context dependent manner. Moreover, while the model does not encounter any constraints during the learning phase, the model can automatically handle any constraints encountered during performance.

Guenther and colleagues relate the notion of the orosensory-to-articulatory mapping to the notion of the coordinative structure in the dynamical system perspective (Saltzman & Kelso, 1987). Similarly, researchers have acknowledged that "several key concepts of Schema Theory (motor programs, schema-type relations) are also incorporated in the recent DIVA model of speech production" (Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008, p. 280). According to the DIVA model, each speech sound is represented as a convex region—described in an auditory/sensory space—in the phonetic frame "speech sound map." The feed-forward model is described as a projection from the pre-motor area to the primary motor area; this notion mirrors the motor program concept as proposed in the Schema Theory. Also, the DIVA model integrates the notion of the recall schema of the Schema Theory—the rule governing the assignment of the most appropriate parameter based on the initial condition and the intended outcome. In the DIVA model, each speech sound is represented as a region in an auditory/sensory space. This large area of representation, which causes one-to-many multidimensional outputs, allows the speaker to produce the same phoneme or the syllable by utilizing a variety of vocal tract configuration and acoustic dimensions. In order for the model to produce the most appropriate phoneme in this situation, the model should consider the initial condition and the intended phoneme outcome, and then choose the best parameters to map the phoneme to the best vocal tract configuration in this context.

Moreover, the notion that the pre-motor area projects to the sensory and the auditory areas and provides these areas with information about the expected sensory consequences based on the current configuration of the vocal tract (initial conditions) and the intended phoneme outcome coincides with the recognition schema in Schema Theory (Schmidt, 1975).

From the previous discussion, it seems that the DIVA model is capable of addressing both speech acquisition and learning as well as accounting for many phenomena in the speech domain. However, as stated by Guenther (1995), the DIVA model addresses “ those aspects of infant development relevant to the acquisition of the motor skills necessary for the production of speech sounds independent of any underlying linguistic meaning or syllabic structure”(p,596). Guenther’s observation points out a common shortcoming, as models of speech production, unfortunately, describe the processes involved in speech production independently from their underlying language processes.

In conclusion, although the DIVA model only addresses speech from a motor perspective, it does offer a detailed account of how speech sounds are produced and learned. Moreover, in the DIVA model the goal of the speech sound is an acoustic goal that accord with the goal of the novel tonal speech sound utilized in the current study. Therefore, the DIVA model is the most suitable model to account for speech motor control in the current study.

The current study examined the role of EFOA and IFOA on performance and learning in the speech domain. Because this study is framed within the Schema Theory, the next section discusses the Schema theory as a foundation for the experiment to test the role of the FOA construct on speech motor learning.

4.5 THEORIES OF MOTOR LEARNING

Two major theories of motor learning are discussed in this section: A closed loop theory and an open loop theory. Both theories have had a great impact on motor learning research. Although both theories assume the existence of an abstract representation for any movement, the two theories differ in the way they value the role of feedback.

4.5.1 The closed loop theory: Adam's theory (1971)

Adam (1971) put forward the closed-loop theory. According to a closed-loop system, motor learning depends on an error detection and correction mechanism. Adam's theory (1971), suggest that peripheral feedback plays an essential role in motor learning; without the feedback, no movement can be performed and no learning can take place. Movements are controlled by comparing a stored reference copy about the movement's correctness with the incoming information about the ongoing movement. This incoming information is in the form of feedback received from the moving structure. Any detected deviation from the stored reference copy results in an error signal. This error signal is sent to a control center, which will then initiate corrective action for the next attempt of that movement. Two types of memory representations are assumed to be in charge of organizing the movement: "memory trace" and "perceptual trace." The memory trace is an abstract representation of skills that are learned from past experiences. This memory trace is responsible for starting the movement by choosing the path of the action and initiating the movement (Adams, 1971).

Once the movement is initiated, the perceptual trace takes over. The perceptual trace evaluates this movement by comparing it to the reference copy. The perceptual trace results from feedback received during previous practice attempts. Based upon Adam's theory, the perceptual trace is strengthened only through correct movement execution. It is the responsibility of the perceptual trace to detect errors and generate error signals. The error signals are eventually sent to the control center to cause a corrective action for the next attempt of the movement. Adam's theory assumes that any error produced during practice is detrimental for learning; only correct trial feedbacks strengthen the perceptual trace. Adam based his closed loop theory on "slow, linear-positioning movements." This closed loop mechanism seems suitable for controlling and learning slow movements; the slow movement allows for ample time to receive and process feedback before the movement ends. However, this theory does not explain the learning or production of a rapid movement that is performed in a shorter period of time, which is usually less than the time required for feedback. Another shortcoming of the closed loop theory is its inability to explain how a movement can be performed in the absence of incoming sensory information, for example, in the case of a deafferentation. That is because Adams constructed his theory with the notion that without feedback, no movement can be performed and eventually learned. These limitations led Schmidt (1975) to formulate his Schema theory as a major revision of Adam's theory. The Schema theory emphasizes the role of an open loop processing in motor learning to account for the following: rapid, ballistic movement execution; production of novel movements; and as an explanation to the storage problem (as discussed in the following section).

4.5.2 The open loop theory: Schema theory (Schmidt, 1975)

In his Schema theory, Schmidt considers motor programs to be a plausible explanation for the control of fast movements without the dependency on feedback. According to Schmidt (1975), the idea of the motor program entails a plan for every movement the organism needs to perform. These plans (motor programs) are stored in the brain. As a result, these programs are retrieved before movements are initiated. When the notion of the motor programs was first proposed, Scholars criticized it as it introducing a storage problem in the brain. They attribute the storage problem to a consequence of the space needed to store the immeasurable number of the motor programs, which govern all the movements the organism is capable of performing. Moreover, when first proposed, the motor program concept did not account for how a novel action would be controlled. As a consequence, Schmidt introduced the notion of "generalized motor program" (GMP). GMP is an abstract representation of not one but a whole class of movements that shares invariant characteristics among them. As such, one GMP can be utilized to plan many movements as long as they share invariant characteristics; this notion would decrease the number of motor programs that the brain needs to hold. This addresses the novelty problem by explaining how one can perform a novel movement that one was incapable of previously performing. Schema theory assumes that the most appropriate GMP will be retrieved from memory and then get adjusted to meet the requirements of the novel outcome.

Generalized motor programs and parameters are two key concepts in Schema Theory. The GMP is an abstract representation about invariant information for a class of movements, such as the relative timing of when particular events should take place or the relative force required to produce a movement. Parameters are the values assigned to the GMP; these parameters allow individuals to adjust a movement pattern to meet specific environmental demands, such as performing a movement with a shorter or longer duration, while applying more or less force or defining the muscle group that will perform the movement. Individuals are able to produce different versions of a movement governed by the same GMP by assigning different parameters. To produce any movement, the performer retrieves from memory the GMP that governs that class of movements and then applies specific parameters. Rules developed from previous experiences govern the parameterization of the GMP.

Schema theory proposes that after movement execution, four types of information will be available to the learner. The four types of information include: 1) the initial condition of the environment or the structure's position (for example, the position of a limb before the movement, or the position of the tongue before producing a sound); 2) the parameters specified for the generalized motor program; 3) the outcome of the movement; and 4) the sensory consequences of the movement (how the movement felt, sounded). From this information, two abstract schemas (rules) are formulated and stored. Schema is a concept borrowed from Head (1926) and Bartlett (1932) (as cited in Schmidt & Lee, 2005), and defined by Schmidt as "an abstract memory representation thought of as a rule, concept, or generalization" (Schmidt & Lee, 2005, p.413).

The two schemas are the recall schema and the recognition schema. The recall schema, responsible for movement production, stores the relationship among the initial conditions, the movement outcome, and the parameters that were specified to the GMP to produce the movement. The recall schema deals with providing the GMP with the appropriate parameters according to the initial conditions and the intended outcome. On the other hand, the recognition schema is responsible for movement evaluation after execution. The recognition schema stores the relationship among the initial condition, the movement outcome, and the sensory information received as a result of the movement. The recognition schema plays an important role in error detection by comparing the predicted sensory consequences from the memory with the actual sensory consequences after the movement is completed (Schmidt, 1975; Schmidt and Lee, 2005). For example, before producing any movement, the GMP for that class of movements is first retrieved from the memory. Then, based on the initial environment (the weight and size of the ball or the position of the tongue or the jaw before producing a speech sound) and the intended outcome (putting the ball in the basket or producing the intended sound to convey the intended word), both the recall schema and the recognition schemas are activated. The recall schema is responsible for assigning the most appropriate parameter, acquired from previous experience, to the GMP in order to achieve the intended outcome. Also based on previous experience, the recognition schema specifies the expected sensory consequences of this movement; once the movement is carried out, the recognition schema evaluates the movement by comparing the actual sensory consequences after the movement with the stored sensory consequences for that movement in the memory.

The recall schema and recognition schema in Schema theory operate in somewhat the same way as the memory trace and the perceptual trace in Adam's theory, respectively. According to the Schema theory, both recall and recognition schemas should be strengthened through practice for motor learning to take place. Motor learning is defined as a permanent change in the person's ability to produce specific movement (Schmidt and Lee, 2005). Motor learning in Schema theory is directly related to the strength of both schemas. The strength of these schemas cannot be measured directly, but it can be inferred from the person's ability to produce a certain behavior. After practice, motor learning is usually measured from the performance during either retention or transfer test, or both. The retention test, requires the individual to reproduce the practiced task either immediately after completing the practice sessions (immediate retention) or after some delay that could be hours or days (delayed retention). The retention test shows the performer's ability to maintain the practiced behavior by applying the same practiced parameter he utilized during acquisition; the performance reflects the strength of the recall schema. On the other hand, the transfer test necessitates the individual to produce an untrained task that is similar to the trained task; the performance reflects the strength of the recall schema in choosing the most appropriate parameter in this novel situation (that is how the learned behavior can be generalized to other related task).

Schmidt and Lee (2005) emphasize the distinction between the performance during practice (acquisition) and the performance during learning tests (retention or transfer); they recommend that learning should only be inferred from the performance during learning tests (retention/transfer). Furthermore, the participants' performance during acquisition cannot predict the participants' learning as many factors, such as motivation, presence of feedback, context within which the task is practiced, and fatigue, might transiently affect performance during practice. Learning is best inferred by retention or transfer tests when the transient factors are eliminated. Proponents of Schema Theory argue that there are many variables that influence motor learning. These variables are designated "principles of motor learning" (Schmidt and Lee, 2005; Shea and Wulf, 2005). The practice and feedback variables are among the most influential factors of motor learning. Schema theory makes explicit predictions about how manipulation of these principles affects motor learning. Researchers in the limb motor learning literature have investigated how manipulation of feedback and practice variables influences motor learning (Guadagnoli and Lee, 2004; see Schmidt and Lee, 2005; Wulf and Schmidt, 1989). Different practice and feedback conditions and their effect on motor learning are discussed in the following sections.

4.5.2.1 Conditions of Practice

Practice is the most significant principle that influences motor learning (Schmidt and Lee, 2005). The motor learning literature describes practice from amount, intensity, and order perspectives. A brief description of each practice condition and how it affects learning will be discussed in the following sections.

Amount of practice

Amount of practice refers to the amount of time a performer devotes to the practice of a task (number of trials before retention or transfer are tested). Through practice, the performer has a better opportunity to experience different information (e.g., initial conditions, intended outcome, the feedback about the actual outcome, sensory consequences) both before and after executing the movement. This experience strengthens and updates the recall and the recognition schemas. Evidence from the limb literature has demonstrated that an increased amount of practice enhances learning (e.g., Schmidt and Lee, 2005; e.g., Park & Shea, 2003, 2005) and eventually enables the learner to perform the task in an automatic mode (Schmidt and Lee, 2005). Although schema theory is based on the assumption that schema are strengthened through practice, it should be noted that the effect of practice in the motor learning literature is not always obvious; practice can interact with other practice conditions (Giuffrida, Shea, and Fairbrother, 2002; Shea & Kohl, 1991) as well as with different feedback conditions, which might obscure its effect in some studies. On the contrary, researchers reported that massive practice can deter learning.

Keetch, Schmidt, Lee, and Young, (2005) designed a study to compare the effects of massive practice on a basketball shooting task. A group of basketball players participated in this experiment. The experimental task involved shooting a basketball from different distances. One of these distances was a massive practice—shooting from the foul line—due to the foul line importance in basketball. A regression analysis on the basis of the shot distance from the basket calculated the predicted shot success of each distance. If no beneficial effect of massive practice from the foul line resulted over the other distances, the researchers predicted that the actual success of all shots would not significantly differ from those predicted by the regression line. The results showed that the actual score for the players from the foul line was significantly higher than the predicted value by the regression line. However, all shooting scores from the other distances were not significantly different from the predicted value in the regression line. The researchers interpreted these results to indicate that massive practice at one parameter (foul line) did not enhance learning for that class of movements; rather, it resulted in the emergence of an a special skill that did not enhance movements within the same class.

Similarly, the findings of Park and Shea (2003) indicated that two amounts of practice conditions revealed differential effects on retention and transfer tests. Two groups of participants practiced a continuous force exertion task by their dominant arm for either 200 or 800 trials. On the retention task, the 800 trial group outperformed the 200 trial group. This finding is consistent with the Schema theory prediction. However, the increased practice hampered the performance on the transfer task. The 200 trial group outperformed the 800 trial group during the two transfer tests, which required the participants to repeat the task with their 1) non-dominant hand and 2) dominant hand but using other muscle groups. These findings suggest that increasing the amount of practice might cause the parameter (the dominant limb in this study) to become an integral part of the GMP, which would then hamper the movement execution by any other effector (Park & Shea, 2003).

Park and Shea (2005) also reported this negative effect of increased practice on a transfer test when they utilized another task (moving a lever to a target in a sequential order). They reported that although extended practice benefited the performance on the retention test, it was detrimental on the transfer test performance. Although Schema theory advocates the importance of increasing practice trials, no consensus on what might be considered an optimal amount of practice in motor learning. In the motor learning literature, differences in the tasks utilized and the complex interference of the amount of practice with other studied variables make it difficult to define an optimum number of practice trials.

Variability of practice

Practice variability is another practice schedule that affects motor learning. The motor task can either be practiced in constant or variable context. In constant practice, only one version of the movement is performed (e.g., shooting the ball from the same distance); in this case, the task is performed with one parameter for the whole practice. In variable practice, on the other hand, different versions of the movement are performed, that is the same shooting task is practiced from different distances (practicing movement variations by assigning different parameters to the same GMP).

According to Schmidt's (1975) variability of practice hypothesis, the recall schema is strengthened as a result of practicing the task from different initial environmental conditions, parameters, or movement outcomes. This variable practice would develop a more valid schema that would allow the performer to apply the most appropriate parameters when required to perform unpracticed novel task (Schmidt & Bjork, 1992; see Van Rossum, 1990, for a review). Accordingly, Schema theory predicts that practicing the same task by assigning many parameters would lead to better retention of these movements and a better transfer to unpracticed movements governed by the same GMP. It should be noted here that the variability of practice hypothesis does not specify the selection of movement GMP; rather, the Schema theory is only concerned with assigning different parameters within the selected GMP for that task. Motor learning literature has extensively studied the effect of variability of practice on learning. Results from these studies provided support for the beneficial effects of variable practice (Giuffrida, Shea, & Fairbrother, 2002) and the detrimental effects of constant practice (Shea & Kohl, 1991; see Van Rossum, 1990, for a review) on motor learning. Also of importance, the variable and constant practice has been shown to have a differential effect on GMP and parameters learning (e.g., Giuffrida et al., 2002; Lai & Shea, 1998; Lai, Shea, Wulf, & Wright, 2000; Shea, Lai, Wright, Immink, & Black, 2001). Specifically, constant practice benefits learning of a relative timing (GMP), while a variable practice is more useful in enhancing parameter learning (e.g., Giuffrida et al., 2002; Lai & Shea, 1998; Lai, Shea, Wulf, & Wright, 2000; Shea, Lai, Wright, Immink, & Black, 2001).

Lai, Shea, Wulf, and Wright (2000) examined the effect of different practice conditions in an attempt to detect a practice condition which will enhance both GMP and parameters learning. Participants, in their study, performed a key-depressing task in one practice session. The practice session was divided into two halves, and the participants' practice condition either changed for the second half or remained the same during the practice session, according to the participant's assigned group, as follows: Constant-Variable; Variable-Constant; Constant-Constant; Variable-Variable. In the constant-condition, the participants practiced one variant of the task, while in the variable practice; the participants practiced many versions of the task which differed only in the absolute timing. The findings indicated that the Constant-Variable practice was optimal for learning both relative timing (GMP) and absolute timing (parameters) relative to other practice conditions. Empirical evidence indicated that practice variability interacts with other factors, such as feedback frequency and practice schedule (e.g., Giuffrida et al., 2002; Shea & Kohl, 1991; Wulf & Shea, 2004). Interestingly, not all variable practices appeared to be equally useful, especially if different GMPs governed the practiced movements.

Practice organization (schedule)

Unlike the variability of practice, in which one or more versions of the task are practiced (in which all movement variations belongs to the same GMP), practice schedule refers to another condition in which different movements governed by different GMPs are also practiced in the same practice session. Researchers demonstrated that the order in which these different tasks trials are presented in the practice session has a differential effect on learning. Different tasks can be either presented in a blocked, random, or serial order. In a blocked practice, participants practice one task in subsequent order before moving on to another task; therefore, the participants practice trials that require the same GMP without being interrupted by another task.

In random schedule, however, the learners practice a different movement task—that requires a different GMP—on each trial; that is, the learners usually do not practice the same movement on two successive trials. A serial practice schedule is similar to a random practice in that the learner practice a different movement target on each trial, but the order in which the trials appear in the serial practice is constant and predicted (e.g., if the practice consists of four tasks, a serial order might be 1234, 1234, 1234). As such, in a serial schedule, the target of the next trial can be predicted because the order of the tasks is repeated throughout the practice. Many researchers reported the benefits of random practice over blocked practice on retention and transfer tests (e.g., Lee & Magill, 1983; Lee & Magill, 1985; Shea, Kohl, & Indermill, 1990; Shea & Morgan, 1979; Wright, Black, Immink, Brueckner, & Magnuson, 2004; Wulf & Lee, 1993).

In other studies, random and serial schedules yielded comparable results (Shea, Lai, Wright, Immink, & Black, 2001). The Schema Theory cannot accommodate these results because it does not explicitly predict this differential effect of blocked relative to random practice. Therefore, researchers proposed the elaborative processing hypothesis and the reconstruction hypothesis as alternative explanations for this differential effect of blocked and random practice on learning. According to the elaborative processing hypothesis, Shea and colleagues (Shea & Morgan, 1979; Shea & Titzer, 1993; Shea & Zimny, 1983) argued that the more processing the learners engage in during practice, the better the learning. Random practice allows more elaborative processing between the tasks, which are held in the working memory during practice; as a result, the learners become more aware of how the practiced tasks differ or resemble each other during practice. Because the learners actively engage in the random practice condition, they will most likely outperform those learners in a blocked practice on retention and transfer test. This discrepancy in performance stems from the fact that no such processing can take place in a blocked practice since all practiced trials in a block are the same. According to the elaborative processing hypothesis, it is this engagement in the elaborative and distinctive processing that strengthens the ability of the learners to distinguish the requirement of each task and thus improve learning.

The reconstruction hypothesis (Lee & Magill, 1983; Lee & Magill, 1985), assumes that the repeated reconstruction of the action plan for each trial strengthens the learners' engagement in constructing the plan of action for each task. An important assumption in the reconstruction hypothesis is that the working memory cannot hold more than one action plan. Random practice requires a new action plan to be constructed for each trial, because the previous action plan is not useful for the current trial and should be discarded before a new one can be planed. Based on this assumption, every trial will require the retrieval and construction of its action plan. On the contrary, in a blocked practice, the same action plan which was constructed for the first trial in a block can remain in the working memory for subsequent trials within that block. This blocked condition would decrease the chance of reconstructing the action plan, thus hindering the learning process.

Although the previous two hypotheses offered plausible explanations for the beneficial effects of random practice over constant practice on learning, researchers reported findings that appear to contradict the beneficial effect of random practice. During the practice sessions, participants in the blocked practice condition outperformed those in the random practice condition during acquisition. However, during retention and transfer, the learning of participants in the random practice condition was better than the learning of those in the blocked condition (Lee & Magill, 1983; Lee, Magill, & Weeks, 1985; Shea & Morgan, 1979). That again emphasizes why in Schema Theory performance during practice should not predict learning (as was discussed earlier).

Using a contextual interference (CI) perspective, researchers have proposed an explanation for these differential effects of a random versus blocked practice on learning (Lee, & Magill, 1983; Sekiya, Magill, Sidaway, & Anderson, 1994). Contextual interference (CI) is created from the context within which the trials are practiced. This CI notion argues that practicing more than one task (each task governed by a different GMP) in a random order causes high contextual interference. Random practice requires the participant to perform a different task variation on each trial. On each trial, the participants should retrieve a different action plan (GMP) and apply an appropriate parameter. During a blocked practice, the participants are required to perform the same task variation in a block before practicing another variation. As such, within the block, the participants do not have to retrieve the action plan (GMP) for that block; rather, the participants can produce and modify their response only by applying different parameters. According to contextual interference hypothesis, the extra planning from trial to trial results in a high contextual interference (as in random practice) that affects performance during acquisition but_enhances performance during retention or transfer test. The CI hypothesis asserts that the CI effect is more obvious when different GMPs, not the same GMP, govern the task variations (Hall, & Magill, 1995; Lee, Wulf, Schmidt, 1992; Sekiya, Magill, Sidaway, & Anderson, 1994).

Hall and Magill (1995) sought to find out whether the beneficial effect of variable practice can be attributed to schema enhancement (as predicted by the variability of the practice hypothesis) or to contextual interference. While they manipulated the practice schedule (random or blocked), Hall and Magill had their participants practice variations of a multi-segment task, and either the same or a different GMP governed the task variations. This manipulation yielded four experimental conditions: 1) “same relative time—blocked practice schedule”; 2) “different relative time—blocked practice schedule”: 3) “same relative time—random practice schedule”; and 4) “different relative time—random practice schedule”. The findings of this study demonstrated the following: Although practice schedule (random or blocked) did not appear to affect the performance when the task variations belonged to the same movement class (governed by the same GMP), practice schedule effects occurred when different GMP governed the practiced tasks. When different GMP governed the tasks, participants in the random practice group outperformed those in the blocked group in both retention and transfer tests. However, participants assigned to random practice were less effective in their performance during acquisition. When the same GMP governs the practiced task variations, no difference was found between random and blocked practice groups.

Hall and Magill took their findings to indicate how both CI and Schema theory speak to the variability of the practice effect when the practiced movement variations are governed by the same or different GMPs, respectively. The CI effect can be summarized as follows: When the contextual interference is high, such as in random practice, the performance during practice (acquisition) will decrease; this decrease is transient because an enhancement of performance occurs during retention tests. When the contextual interference is low, as in blocked practice, the performance during practice (acquisition) will improve, but this low contextual interference will diminish retention.

Not only does the practice schedule produce a differential effect on performance during acquisition and learning tests, but this differential effect of the practice schedule has also been reported for GMP and parameters learning. Researchers have reported that parameter learning (e.g., learning absolute timing) is enhanced in a random practice relative to blocked practice, and that learning a GMP (e.g., relative timing) is enhanced in a blocked practice condition (Shea, Lai, Wright, Immink, & Black , 2001). The stability hypothesis was proposed (Shea et al., 2001) to account for these findings. From the stability hypothesis perspective, a relative timing (GMP) learning is enhanced in a practice condition that maintains constancy between trials, as in a blocked practice. Compared to random practice, blocked practice is more stable because the absolute timing requirements do not change from trial to trial; this stability allows the learners to extract the invariant relative timing characteristic (GMP) of the movement and thus enhances GMP learning. Random practice which is less stable due to changes in the parameters requirements, from trial to trial, enhances the learning of absolute timing of the task (parameters). In the Shea et al. study, all the practiced movement variations differed in terms of their parameters (absolute timing requirements), but were all governed by the same GMP (they all had the same relative timing).

Practice distribution

Although Schema theory advocates that practice enhances learning, not all structures of practice are equal. Schema theory makes explicit predictions about the effects of the amount of practice and the variability of practice (constant or variable) on motor learning. In addition to the previous discussed variables, researchers of the motor learning have demonstrated that the pattern or distribution of trials (massed or distributed) affect motor learning. During massed practice, all the training trials (or sessions) are performed one after the other (back-to-back), with only a short period of rest in between. During distributed practice, the rest periods between practice sessions are usually much longer (longer than the time utilized for practice; or might also extend to days). Evidence from the limb motor learning literature demonstrated that distributed practice enhances motor learning (Baddeley & Longman, 1978; Shea, Lai, Black, & Park, 2000). Robertson, Pascal-Leone, & Miall (2004) attributed the beneficial effects of distributed practice, when practice sessions were distributed, to a better memory representation of the movement. Empirical evidence, which demonstrated that allowing time between practice sessions enhanced learning, confirmed this prediction (Lee & Genovese, 1989; Shea, Lai, Black, & Park, 2000).

All the above mentioned practice variables interact in a complex way during practice. Giuffrida, Shea, and Fairbrother (2002) sought to study the effect of three practice schedules (constant, blocked, and serial) under two practice conditions: low (54 trials) and high (162 trials). The participants were randomly assigned to one of six groups. They were required to practice a multi-segment timing task. The participants practiced one version of the task—with the same absolute timing (one parameter)—in the constant practice group and three versions of the task—with different absolute timing—in both the blocked and serial groups. Half of the participants in these three groups practiced in a low practice condition (54 trials), while the other half practiced in a high practice condition (162 trials). During acquisition, feedback was provided after each trial. Retention and two transfer tests measured learning. In the retention test, participants were required to perform the same version of the practiced task. In the transfer test, participants were required to perform two unpracticed task variations: one task governed by the same GMP, and the other task version governed by a different GMP. The results indicated that the amount of practice interacted with the practice schedule. In the retention test, the participants in the high practice group outperformed learners in the low practice group when they practiced the task in a blocked or serial practice schedule; however, the low practice group outperformed the high practice group in the constant schedule group (Giuffrida, Shea, & Fairbrother, 2002).

Shea & Kohl (1991) reported a similar detrimental effect of increased constant practice on learning. Lee and Magill (1983, 1985) attributed the detrimental effect of constant practice on learning to the fact that during constant practice the learner utilizes only one GMP; as such, this GMP might have been present in the memory during the practice period and not reconstructed for each trial (which would have strengthened the schema and resulted in more automatic performance). Also, practicing the same movement over and over again would strengthen the schema only for that specific version of the task (applying only one parameter); this would enhance retention of the same task but be detrimental in a transfer task (Park & Shea, 2003, 2005).

In summary, the limb motor learning literature supports the beneficial effect of random practice when learning is inferred from both retention and transfer tests. It should be noted that this random practice, like other practice conditions, interacts in a complex way during acquisition. Researchers have also determined that random practice is beneficial in the speech domain (as will be discussed later).

4.5.2.2 Feedback variables

The second principle of motor learning that has been extensively studied in the motor learning literature is feedback. Participants usually receive different kinds of information after performing a movement. This information can be either received from within the body, or provided to the performer, these two types are designated intrinsic and extrinsic feedback, respectively. The intrinsic feedback refers to information about how the movement felt or sounded to the performer. The extrinsic feedback either refers to the feedback that the performer receive by looking at the movement outcome (e.g., the performer can see if he missed the target), or to feedback provided to the performer. The external feedback that is provided usually by the examiner or coach to the performer is designated as augmented feedback.

Because of the inherent difficulty in manipulating the intrinsic feedback, researchers have chosen to manipulate the augmented feedback in order to determine its effect on motor learning.

Researchers have manipulated augmented feedback either by its form, amount, or time of presentation (for reviews, see Swinnen, 1996; Wulf & Shea, 2004). Effects of different manipulations of augmented feedback will be discussed in the following sections.

Forms of Augmented feedback

Augmented feedback has been categorized into two forms; knowledge of result (KR) and knowledge of performance (KP) (Schmidt & Lee, 2005). KR is the information provided to the performer (by the instrument or instructor) about the movement outcome in relation to the intended goal (e.g., telling the performer he was too fast if the goal is to perform the movement in specific speed, or simply saying he was correct). KR is usually provided after the movement is completed. KP provides the performer information about the pattern of the movement such as "your elbow wasn't flexed enough" or about the quality of the movement. Both KR and KP which have a positive influence on error corrections on subsequent trials, eventually enhance performance and learning (Winstein & Schmidt, 1990). Although KP has been found to be more beneficial in some studies when the performer was not familiar with the task goal or did not know the intended goal of the movement (e.g., Newell, Carlton, & Antoniou, 1990), KP was found to be just as effective as KR when the movement goal was known (e.g., Swinnen, Walter, Lee, & Serrien, 1993). Therefore, KP might be more useful if the performer must learn a novel task such as proposed in this study.

As was discussed previously, Schema Theory claims that the availability of the information after the movement plays an integral role in schema development; and without it, schema cannot be developed or updated. This perspective highlights the importance of the augmented feedback in making the outcome information available. Since both the recall and recognition schemas require this information, both will benefit from such feedback. Moreover, it might be presumed that the more frequent this feedback, the more the benefit for the learner; such an assumption might strengthen schema development. Nevertheless, though increasing the availability of feedback was beneficial in some situations (during acquisition), it was detrimental to retention and transfer in most studies. The augmented feedback frequency effect is discussed in the following section.

Feedback Frequency

The feedback frequency refers to the proportion of trials on which feedback is given, to the total number of trials. In the motor learning literature, the feedback frequency is typically presented as a percentage. During practice, the augmented feedback can range from lowest - no - frequency of 0% to the highest frequency of 100% and anything in between, on a continuum. Previous studies have demonstrated improved learning with low frequency KR relative to high frequency KR (Wulf, Schmidt, & Deubel, 1993; Winstein and Schmidt 1990; Wulf, Lee, & Schmidt, 1994; Anderson, Magill, & Sekiya, Ryan, 2005).

Winstein and Schmidt (1990) examined the effect of high versus low feedback frequency on the learning of a lever-positioning task. Two groups performed exactly the same task but differed in the frequency in which KR was presented, with 100% in one group and 50% in the other. Although both groups performed in a similar way during acquisition, this study showed that the group that practiced with 50% feedback frequency was more accurate on a retention test than the other group. The researchers used the guidance hypothesis to interpret the retention test results (e.g., Salmoni, Schmidt, & Walter, 1984; Schmidt, 1991). This hypothesis proposes that, although the feedback provided to the performers has a beneficial effect in guiding the performers to the correct movement (which contributes as input to update the schema about outcome information), the feedback may have detrimental effects on learning. The interpretation is that frequent feedback deters learning if the performers become dependent on it to guide them; by doing so, the performers do not involve themselves in any information processing of their own internal feedback to detect errors. As a consequence, the performers fail to produce the correct movement on the retention test when the feedback is suspended. In the original study (Winstein and Schmidt, 1990), the similarity between groups during acquisition might be attributed to the faded feedback schedule utilized during practice. The 50% feedback group received the feedback in a faded schedule; that is, the feedback frequency decreased as practice continued. Apparently, this fading feedback schedule might have optimized the effect in the low frequency group during practice.

Wulf, Schmidt, & Deubel (1993) sought to determine whether a low frequency feedback would differentially affect the learning of generalized motor programs and movement parameterization. They randomly assigned participants to one of two practice groups according to 100% or 63% feedback frequency. The participants' task was to perform lever-patterning movements that required the same GMP but differed in their parameters. An independent error measure, separately calculated both the GMP and parameter accuracy. The findings showed that high frequency feedback during practice degraded GMP learning, but enhanced parameter learning, as indicated by both transfer and retention tests. On the other hand, the group that received low frequency feedback during practice was more accurate in its GMP but was not accurate with parameters during delayed transfer and retention tests.

The Schema Theory perspective explains the detrimental effect of low feedback frequency. According to Schema Theory, parameter learning occurs as a result of practicing the movement by assigning different parameters to achieve movement outcomes. However, in the case of reduced feedback frequency, the outcome information will not be available on every trial. With part of the information missing, no schema update can take place; therefore, parameter learning is diminished.

The stability hypothesis perspective explains the enhancement of GMP learning in a low frequency practice. As discussed before, the stability hypothesis (Lai & Shea, 1998; Shea, Lai, et al., 2001) claims that any factor that induces stability during practice allows the performer to stabilize the performance. This stabilization then enables the performer to extract the invariant characteristics between task variations and to generate a more stable GMP. Unlike high frequency feedback that causes the performance to change from trial to trial, the low frequency feedback stabilizes performance, thereby enhancing GMP learning (Lai & Shea, 1998; Shea, Lai, et al., 2001). Researchers have found that the interaction between the feedback frequency and practice conditions might either obscures or enhances the feedback frequency effect. For example, the beneficial effects reported for low frequency feedback emerge more in a variable practice condition relative to a constant practice condition (see Wulf & Shea, 2004). Moreover, in their review, Wulf & Shea (2002) demonstrated how principles derived from studying simple motor tasks might not induce the same effect when applied to a more complex task. Therefore, a high frequency feedback, rather than a low frequency feedback, might be more beneficial when learning involves a complex task (e.g., Swinnen, Lee, Verschueren, Serrien, & Bogaerds, 1997).

Of particular interest to the proposed study is the interaction between the feedback frequency and the focus of attention reported by Wulf, McConnel, Gärtner, and Schwarz, (2002). Wulf and colleagues (2002) examined the effect of two feedback frequencies as a function of the focus of attention that the feedback induced. The participants practiced a “ lofted soccer pass”, and were randomly assigned to one of four groups that differed in terms of the frequency and type of feedback received: a) 100%-internal focus feedback, b) 33%-internal focus feedback, c) 100%-external focus feedback, and d) 33%-external focus feedback. The feedback was presented in the form of sentences that indicated how the performer can improve his/her skill according to the produced movement. The frequency of the feedback was provided according to the group condition. In the internal focus group, the sentences referred to the participants' body parts, while in the external focus group, the reference to the body parts was avoided and the sentences referred to the effect of the participants' movement. The results showed a main effect of attention focus with the external focus groups performing more accurately than the internal focus group, regardless of the frequency of feedback. Interestingly, the interaction between feedback type and frequency was significant; during practice, the group with 33%- internal focus feedback performed with higher accuracy when compared to the group with the 100%- internal focus feedback. The opposite was true for the groups with external focus feedback. That is, a more frequent feedback- inducing external focus (100%) was more beneficial than less frequent feedback (33%). The infrequent internal feedback also gave performers more opportunities to focus externally. Because the guidance hypothesis perspective cannot explain these results, the researchers interpreted their findings in terms of the constrained action hypothesis, which suggests that adopting an internal focus of attention causes the performer to consciously control the movement. This attempt to consciously control the movement freezes the otherwise automatic motor control processes, which eventually disrupts performance and learning (see Wulf & Prinz (2001) for a review). As a result, the researchers suggested that using only the guidance hypothesis perspective to interpret feedback frequency results might be misleading and not provide insight into the complex interaction between feedback frequency and FOA; instead, they recommended that the results should also be interpreted from the attentional focus perspective as well.

Feedback temporal manipulation

Feedback temporal manipulation refers to the time when the feedback is presented to the performer relative to the trial. The augmented feedback can either be presented during or after the movement. That is, as concurrent or terminal feedback, respectively. Terminal feedback can be given either as 1) immediate feedback that is presented after the movement completion without delay or 2) after the movement with some delay (Schmidt & Lee, 2005). Compared with concurrent feedback, researchers have demonstrated that terminal feedback is more beneficial to learning (e.g., Schmidt & Wulf, 1997; Vander Linden, Cauraugh, & Greene, 1993). Although concurrent feedback enhanced performance during practice, it hampered learning when measured by retention and transfer tests (e.g., Park, Shea, & Wright, 2000; Schmidt & Wulf, 1997; Vander Linden et al., 1993). Moreover, terminal feedback provided a few seconds after movement termination proved more beneficial than an immediate terminal feedback (e.g., Swinnen, Schmidt, Nicholson, & Shapiro, 1990). The period after the trial completion is the time when the internal feedback can be processed.

Augmented feedback provided immediately after the trial was determined to have less value than when it is delayed. That is because the immediate augmented feedback might prevent the performer from processing the internal feedback from the body as results of that trial (Salmoni, Schmidt, & Walter, 1984). For example, in the Swinnen, et al, (1990) study, the researchers manipulated the delay before presenting KR in order to test the effect of this period on learning a motor timing task. The participants were randomly assigned to one of two groups.

The first group received the KR immediately after the movement (no delay group), while the second group received the KR after a 3.2 seconds delay (delay group). In addition, the second group was required to estimate their error before the KR was presented. The findings indicated that the group that received the KR after 3.2 seconds along with their error estimation after movement termination outperformed the immediate group during the delayed retention performance. Interestingly, this advantage of practicing with delayed KR persisted in a second retention test four months after practice. Using a similar logic, Anderson, Magill, Sekiya, and Ryan (2005) reported that delaying the feedback after two intervening trials –and provided as a summary feedback - led the participants to depend on their intrinsic feedback as a measure of their error.

In sum, the studies reviewed above highlighted the effect of several feedback manipulations on motor learning including the nature, frequency and time of presentation of the feedback. Moreover, this section further indicated how principles of motor learning interact in a complex ways to affect learning.

Though the existence of other motor control theories is well acknowledged (e.g., Kelso & Tuller, 1981; Saltzman & Munhall, 1989; Thelen & Smith, 1994), the proposed study will use the Schema Theory perspective because it provides a theoretical framework for explaining the principles of motor learning as discussed above. Furthermore, researchers in speech production and speech motor learning have incorporation of the principles of motor learning, derived primarily from the limb literature, into speech motor learning. The next section reviews some of these studies to provide evidence about the applicability of the Schema theory to speech motor learning.

4.5.3 Schema Theory and Speech Production

The interest in utilizing principles of motor learning derived from the limb literature in speech motor learning is increasing due to the influential effect of such principles on motor learning in general. In order to achieve the best results for speech motor learning, it is necessary to define what might constitute GMP and parameters in speech. As researchers acknowledge, the issue of specifying what GMP might represent in speech remains uncertain (e.g., Ballard, Granier, & Robin, 2000). Researchers have proposed many speech units as possible candidates that a speech GMP might encompass. The suggested speech units include: the phoneme (Van der Merwe, 2008; Rogers & Spencer, 2001); the syllable (Levelt et al., 1999; Aichert & Ziegler, 2004); the stress group (Sternberg, Monsell, Knoll, Wright, 1978); the word (Klapp, 2003), and the phrase (Varley, Whiteside, Windsor, Fisher, 2006).

In general, the GMP notion entails an invariant relative timing, relative force among involved muscles, or both. For example, Ballard, Maas, and Robin (2007) suggested an important timing in speech– the voice onset time (VOT). The VOT describes a relative timing of events between oral (upper airway articulators) and laryngeal (vocal folds) movements. As such, Ballard and colleagues speculated that the VOT might be an integral part in the speech GMP. As a consequence, they assumed that different GMPs govern voiced and voiceless sounds. Moreover, as regards to the invariant relative force which characterizes GMP, Ballard and colleagues assumed that a plosive speech sound, which requires more force to accomplish a complete closure at the level of the articulators, would be governed by a different GMP than a fricative speech sound, which requires less force to produce an incomplete closure. They base their assumption on the findings that no generalization (transfer) between plosive sounds and fricative sounds occurred in a previous speech treatment study (Knock, Ballard, Robin, & Schmidt, 2000). When applying the same notion of parameters from the Schema theory to speech, parameters might include the absolute duration of the utterance, the absolute force applied by the articulators' muscles.

Researchers have studied the applicability of Schema Theory to the speech domain by investigating the effects of principles of motor learning manipulation on learning a novel speech task or on re-learning speech in participants with normal speech or with speech disorders, respectively. Kim (2007) investigated the effect of two practice conditions and two feedback conditions on the acquisition and learning of a novel speech task. The practice was performed in one session, and the participants performed two retention tests: one day and one week after the training. The two groups of participants, both monolingual English speakers, practiced the production of ten short Korean sentences either for 25 or 100 trials and received feedback (either on 20% or 100% of the trials) on their performance. Native Korean speakers judged perceptually the participants' speech for intelligibility, naturalness, and precision. The results indicated no main effects for the practice trials; the performance - on the two retention tests involving both the 25 trials and the 100 trials groups—did not differ. However, the Kim study showed a significant interaction between the two independent variables (feedback frequency and the amount of practice). Only when the 100 trials practice group received a 20% feedback frequency did the sentences produced by participants in the 100 practice trials condition receive higher scores (from the judges) than the sentences produced by participants in the 25 trials condition.

One strong assumption of the Lee Silverman Voice Treatment (LSVT) (Fox, Morrison, Ramig, & Sapir, 2002), is based on the notion that increased practice benefits treatment outcomes. LSVT is an intensive speech program that is especially designed to improve some of the speech characteristics in the speech of individuals with Parkinson's disease (PD). The treatment consists of intensive practice of four treatment sessions/week for four weeks (i.e., 16 sessions in four weeks) in addition to home practice. Spielman, Ramig, Mahler, Halpern, and Gavin (2007) sought to compare the effectiveness of the LSVT intensive practice with a modified version of the LSVT, which consists of a more extended practice. The extended LSVT version involved the same 16 treatment sessions, but the sessions were distributed as two sessions/week extended over eight weeks. A group of 12 participants diagnosed with idiopathic PD participated in the study. All 12 participants were assigned to the extended LSVT treatment group as the researchers planned prior to their study. This allowed the researchers to compare the results of their study to previously collected data available from another study, which included 14 individuals with PD who received the intensive LSVT practice (Ramig, Sapir, Fox, & Countryman; 2001). The findings of both the intensive practice and the extended practice of the LSVT program were comparable; that is, the extended practice did not demonstrate any benefits over the intensive practice. The results from the above mentioned studies indicate that the amount of practice did not differentially affect the speech characteristics. The amount of practice effect might have a great effect if it interacts with other practice variables.

Similar to the limb motor learning literature, practicing a speech task in a random practice condition has been found to enhance learning when compared to blocked practice. For example, Adams and Page (2000) examined whether the effectiveness of reduced feedback frequency, variable practice and random practice while learning a novel speech task. The participants were randomly assigned to one of four groups: 1) Group1—constant practice-feedback on every trial; 2) Group 2—constant practice-feedback on every fifth trial; 3) Group 3—blocked practice-feedback on every trial; and 4) Group 4—random practice-feedback on every trial. The findings indicated that, “specific variations in feedback and practice schedules (reduced feedback, random practice, and variable practice) were found to have a significantly greater effect on the retention of a novel slow rate of speech” (Adams & Page., 2000, p. 219). This study showed that the same reported effect of practice and feedback conditions, as has been demonstrated in the limb motor learning literature, was replicated in the speech domain. Nonetheless, it should be noted here that, although Adams and Page (2000) did not study the amount of practice as a variable in their study, their results showed that the amount of practice did not improve the learning in the constant group relative to the variable group. Variable practice was not only demonstrated to be effective with normal speakers, but it also has been demonstrated to be effective in treatment for speakers with Apraxia of Speech (AOS). AOS is a motor speech disorders which is described as involving a disruption at the level of motor planning (Van der Merwe, 2008), or both motor planning and programming (McNeil, Robin, and Schmidt, 1997; Ziegler, 2002).

Ballard et al., (2007) examined the effectiveness of a variable practice on the relearning of voicing control in speakers with AOS. In a single subject design, two participants were treated to “re-control” their voicing when producing different phonemes. During the acquisition sessions, the target phonemes were practiced within different phonetic contexts (the phonemes shared the same manner of articulation). Learning was inferred from the participants’ performance on a retention test (maintenance) and a transfer test (ability to generalize the practice task to other untrained but similar tasks). The findings indicated that both participants improved their performance during practice and were able to maintain the practiced voicing control for more than four weeks post-treatment. Ballard and her colleagues also reported that the results of their study showed that variable practice not only enhance learning but also prevented overgeneralization of the practiced phoneme (producing the practiced sound in place of other sounds). The Ballard et al., (2007) study demonstrated that the random practice and decreased feedback frequency are beneficial for re-learning speech in speakers with AOS. Although this study is not the first to apply principles of motor learning to the speech domain (see Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, Schmidt, 2008, for review), the Ballard et al., (2007) study provided evidence that the manner of articulation is an integral part of GMP. The researchers based their evidence on the participants' generalization to another phoneme within the class of the manner of articulation (within plosive sounds), but not across manners of articulation (e.g., no generalization from practiced plosive to fricative sound). Likewise, a single subject alternating treatment design (Knock, Ballard, Robin, & Schmidt, 2000) demonstrated a better retention of a speech tasks, governed by different GMPs, under random practice conditions relative to blocked practice for two speakers with AOS.

The principles of motor learning concerning feedback manipulation have also been applied to the speech domain. For example, speech learning studies have reported the beneficial effect of low frequency feedback on learning. Steinhauer and Grayhack (2000) examined the differential effects among three feedback frequencies on the learning of vowel nasalization, a novel motor task, in young adults with normal speech. The participants were randomly assigned to one of three groups according to feedback frequency received during practice: 0% feedback (no feedback), 50% feedback, or 100% feedback. The findings from Steinhauer and Grayhack (2000) study demonstrated that, compared to low frequency feedback, a high frequency feedback during acquisition was inferior to acquisition, retention and transfer.

Similarly, Adams, Page, and Jog (2002) sought to study the effect of feedback frequency by providing feedback after every fifth trial—another way of reducing feedback frequency. In this study, they replicated their previous study (Adams and Page, 2000) with a group of speakers with PD. Consistent with their previous findings from normal speakers, summary feedback delivered after every fifth trial, enhanced retention of a novel speech task when compared to the group receiving feedback after every trial. Recently, Austermann-Hula, Robin, Maas, Ballard, and Schmidt (2008) employed a single-subject alternating treatments design to examine the effects of the frequency and timing of feedback on the learning of speech tasks in speakers with AOS. This design entails exposing the participants to all treatment conditions (such as the high and the low feedback frequency in this study). In order to separate the effect of the treatment conditions, the researchers also assigned a set of the speech stimuli to both treatment conditions. Therefore, participants can serve as their own control. In both Experiments 1 and 2, the researchers assessed the participants' performance during the treatment sessions and after the completion of the treatment sessions by retention and transfer tests.

In Experiment 1, the researchers compared the relative effects of 100% and 60% feedback frequency on re-learning speech tasks in four participants with AOS. Participants' speech production was scored online, and the knowledge of results feedback (i.e., "correct" or "incorrect") was provided in a frequency according to the treatment condition, on 60% or 100% of trials. The results from this study showed that low frequency feedback might enhance speech relearning as two out of the four participants demonstrated enhanced performance in retention and transfer tests; the unaffected performance in the other two participants might be attributed to the task complexity, as speculated by Austermann-Hula et al. (2008).

In their Experiment 2, Austermann-Hula et al. (2008) examined the effect of two feedback timings—immediate and delayed feedback—on the relearning of speech tasks in two speakers with AOS (who also participated in Experiment 1). The immediate feedback was presented as soon as the trial ended, while in the delayed feedback condition, the feedback was presented five seconds after the trial was completed. In both feedback conditions, the participants received knowledge of result feedback on 100% of practice trials. The feedback was presented visually on a computer monitor with a red or green light to indicate incorrect or correct responses, respectively. The findings indicated that at least one participant benefited from the delayed feedback and showed better performance during retention and transfer. Similar to the first experiment, issues related to task complexity might have affected the outcome in the other participant. The researchers reported that, although the participants in both experiments did not show similar results, some participants showed enhancement in the low feedback frequency condition and the delayed feedback condition. As in accord with motor learning, no participants in this study seemed to benefit from the high frequency or the immediate feedback condition.

Moreover, Katz, Carter, and Levitt (2007) illustrated that implementing principles of motor learning improved treatment outcomes of non-speech oral gestures in one participant with buccofacial apraxia (BFA). As defined by Katz and his colleagues, BFA is “the inability to perform voluntary movements of the larynx, pharynx, mandible, tongue, lips, and cheeks, while automatic or reflexive control of these structures is preserved.” (Katz et al, 2007, p. 1230). For this single subject design study, the researchers assigned two oral gestures to receive structured motor practice and one oral gesture to receive visual augmented feedback during practice. The findings showed that, compared to the oral gestures treated in a conventional motor treatment, the oral gestures treated with the augmented feedback indicated more consistent improvement during acquisition and more maintenance, when tested six weeks after treatment. Katz et al. interpreted these results as preliminary evidence of the effectiveness of implementing principles of motor learning (especially, augmented feedback) in the treatment of BFA.

Recently, McNeil, Katz, Fossett, Garst, Szuminsky, Carter, and Lim (2010) reported the relative effects of providing online—concurrent—augmented visual feedback, in two feedback frequencies, on learning a target speech movement in two participants with AOS. Both participants received two forms of feedback: Kinematic feedback about their tongue movement and another feedback based on the clinician's perceptual judgment of the correctness of the produced sound. Both the kinematic and clinician's feedbacks were provided visually. In a single subject design, the participants practiced producing the sounds while receiving the visual feedback. A transfer (generalization) test and a retention (maintenance) test, given one month after treatment measured the learning. The feedback was presented either 100% or 50% of the time (randomized across treatment targets). The findings indicated that the augmented visual feedback provided during treatment had positive effects on both acquisition and generalization to the untreated but similar sounds. However, in this study, the feedback frequency effect was confounded due to its limited randomization order. Moreover, it was difficult to isolate the effect of the visual online kinematic feedback because this feedback was provided together with the clinician's feedback as discussed by McNeil et al. (2010).

Although findings from the previous review studies are promising, the results should be interpreted with caution because most studies in this section - utilized a single-subject design (see Adams, Page, and Jog, 2002; for exception) which has limited external validity (generalizability of the results to other individuals). This choice resulted from the difficulty of conducting a group design study when the participants showed diversity in severity and when few participants were available to study.

In sum, the reviewed studies provide evidence that the principles of motor learning have a similar effect on learning speech tasks as they do in the limb literature. Specifically, the review of the studies highlights specific manipulations in both practice and feedback (variable practice, random practice, and low feedback frequency, especially if provided in a fading schedule) that enhanced learning in both the limb and the speech domain.

As much as they emphasize the importance of practice and feedback conditions as major influential factors in learning, Schmidt and Lee (2005) have also acknowledged the importance of utilizing the pre-practice period to prepare the learner to the learning process. A major goal to achieve during this pre-practice period is to motive the learners to practice the task.

Motivation is an important factor that makes people eager to learn. The instructor in any learning situation plays a significant role in motivating the learner to be more eager to learn by (a) helping the learner to envision the importance of the to be learned task for that person in particular such as: winning a game or being more intelligible; (b) involving the learner in setting a specific goal to be achieved as was suggested by McNeil et al., 1997 in speech motor learning literature; (c) introducing the task to the learner by providing simple explanation and verbal instructions and expected outcome. Schmidt and Lee (2005) also added that modeling the task for the learner might also help in showing how the task is performed and how the task sounds when producing a speech task. Moreover, during this period, providing the learner with information about the acceptable performance would help to learn the task.

This dissertation study utilized a novel tonal speech task to assess the effects of FOA in the speech domain. Therefore, the next section describes the tonal tasks and reviews the relevant literature.

4.6 GENERAL DESCRIPTION OF TONAL SPEECH TASK

Mandarin is a tonal Chinese dialect. A major characteristic of Mandarin and other tonal languages is the use of frequency contours to indicate lexical differences. Yip (2002) defined tone as the use of pitch (frequency) to contrast word meaning phonologically at the segment or the syllabic level. Accordingly, any language that uses such tone contrasts is designated as a tonal language. Pitch is the psychological correlate of fundamental frequency (F0)—the rate of the vibrations of the vocal folds. As a result, words with the same phonemic structures can convey different meanings according to their accompanying tones at the syllable or the word level (Ng et al., 1998; Yiu et al., 1994). Mandarin Chinese has four tones. These tones are usually described and contrasted by their pitch register (high, middle, low) and by the pitch contour—how the F0 changes over time (e.g. level, rising, falling). A description of pitch register and pitch contour for each Mandarin Chinese tone is shown in Table 1. Tone-1 can be best described as a high-level tone; Tone-2 as mid-rising; Tone-3 as low-falling-rising and Tone-4 as high-falling (Chao, 1948 as cited in Wang, Jongman & Sereno, 2003).

Table 1 A description of the pitch register and Pitch contour of the four Mandarin Chinese tones

Tone	Pitch register	Pitch contour
Tone-1	High	Level
Tone-2	Mid	Rising
Tone-3	Low	Dipping, falling-rising
Tone-4	High	Falling

Figure 1 illustrates canonical pitch contours for the four tones –when produced in isolation; the four tones differ in terms of their F0. Figure 1 also illustrates that another feature that varies between the tones is the dynamic range (the difference between the highest and lowest value of F0). For example, Tone 4 has the largest F0 range, while Tone 1 has the smallest range. These differences between the tones serve as essential cues for speech perception in tonal languages. Among these differences, F0 register and F0 contour (the pattern of F0 changes over time) are considered the primary acoustic parameters (Howie, 1976). In addition, tone duration, amplitude, and turning point are helpful perceptual parameters of tones (Lin, 1965). The correct production of these tones requires a specific control at the level of the vocal folds, though not only at this level of articulation (as will be discussed later). Native Mandarin Chinese speakers do not acquire these four tones at the same rate due to a hierarchy of difficulty among these tones. Many researchers proposed that the level tone (i.e. Tone 1) is the easiest to acquire; Tone

4 (high-falling) is the second easiest to acquire; while Tone 2 (mid-rising) is acquired before Tone 3 (low falling rising) (Li & Thompson, 1977; Yue,1980, as cited in Shen, 1989).

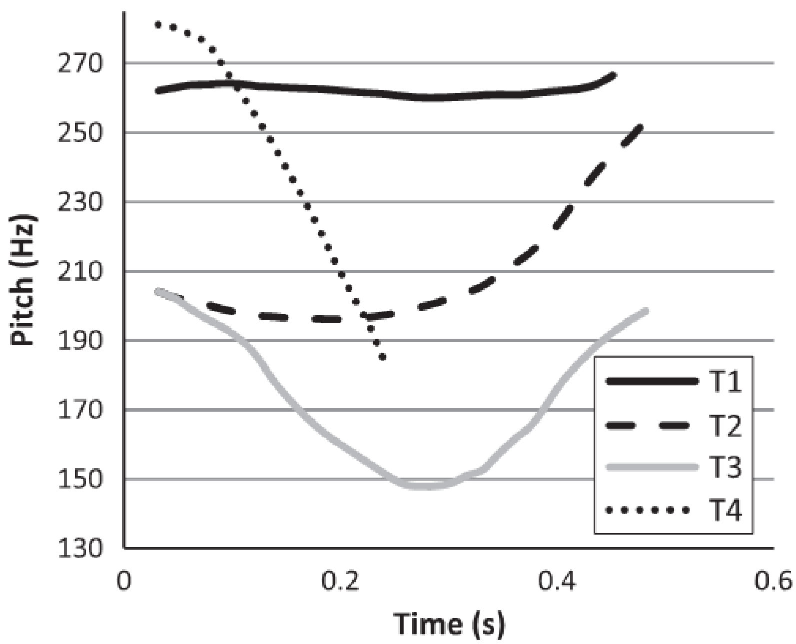


Figure 1 Pitch Contour of the four Mandarin Chinese tone produced in isolation. Reprinted from "Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers", Journal of phonetics, 40, 269-279. by Yen-Chen (2012). With permission from Elsevier.

Not only is the change in pitch during speech production unique to tonal languages, but the way pitch changes in a short period of time is what makes Mandarin Chinese and other tonal languages special and unusual for non-native speakers. The same syllable may mean four totally different monosyllabic words according to how the listener perceives the speaker's F0; the same Ma syllable may mean either: mother, horse, hemp or scold when produced with Tones 1, 2, 3, 4, respectively.

In contrast to tonal languages, English is a stress timed language. In American English, the pitch changes as a function of the sentence level. This type of change at the sentence level is called intonation. American English speakers do not use intonation to make phonemic distinction, but they do use intonation to convey the pragmatic meaning (the speaker intention or expression). For example, a speaker using monotone or a voice with more inflection would still state the same words, but the intonation would express a message that reflects his/her psychological state (boredom, depression, excitement, sarcasm). More importantly, intonation in American English can be used to differentiate syntactic form such as between a statement and a question that will have falling or rising pitch, respectively. Unlike English speakers, the native speakers of Mandarin Chinese tend to have more control over their vocal folds. Their speech is characterized by a larger fundamental frequency range within shorter period of time when compared to English speakers (Chen, 1974; White, 1981).

The non-native Chinese Mandarin speaker's unfamiliarity in the way the F0 is associated with syllables and words results in the tonal task presenting a great challenge and difficulty at both the perceptual and production levels. Thus Mandarin Chinese is recognized as one of the most difficult languages for English speakers to acquire as a second language (Ross, 2001). Several researchers with interest in second language acquisition have analyzed the tone production in an attempt to identify the cause of the tone acquisition difficulty (Chen, 1974; Chen, 1997; Shen, 1989; Wang, 1995; White, 1981).

Li and Thompson (1977) reported a similar tone acquisition order in native Mandarin Chinese speaking children (as T1, T4, T2, T3, moving from easiest to most difficult). By contrast, this hierarchy of difficulty was not found in American English speakers who studied Chinese for four months (Shen, 1989). In this study, Shen specified that the speakers demonstrated an error rates as follows: 16.7% error rate for Tone 1; 8.9% error rate for Tone 2; 9.4% error rate for Tone 3; 55.6% error rate for Tone 4. Shen attributed the high error rate in Tone 4 production to the L1 interference – as it is less marked for English speakers. In other studies that analyzed acoustically the tone production errors produced by non-native speakers, the researchers demonstrated that the signal deviated from that of native speakers from many perspectives. Chen (1974) found the pitch of the non-native speakers to be limited in range. In Chen's study, Mandarin speakers utilized a pitch range that was 1.5 times broader than the range utilized by American English speakers. As a result of this finding, Chen (1974) recommended non-native speakers to work on expanding their pitch range to the level that is required to correctly produce the tones. White (1981) reported a similar limited pitch range. Miracle (1989) reported that when learning Mandarin Chinese, non-native speakers showed a different pitch register (either too high or too low) or pitch contour when compared to native speakers.

Researchers have also examined the learning of Mandarin Chinese from the perspective of second language acquisition, either in educational settings such as classrooms or through an interaction with a computer program designed to improve pronunciation. The tonal production task is an ideal candidate to be trained as a novel speech motor task. That is because it requires control over the articulators as well as a specific demonstration of F0 fluctuation, in which the speaker must decrease, increase, or maintain the pitch at a certain level over the course of the syllable or the word. This F0 fluctuation is essential to produce the target word. Increasing the tension in the vocal folds, which increase the vocal folds rate of vibrations, leads to a higher pitched voice. Although people can increase or decrease their rate of the vocal fold vibration on demand, the structural properties of vocal folds (the length and thickness of the vocal folds) are gender dependent. Due to shorter and thinner vocal folds, women generally have voices with higher fundamental frequencies. Nevertheless, speakers of the same gender might also exhibit significant variability due to age differences. It is well-documented that normal speakers of non-tonal languages show changes in their pitch when speaking. However, native tonal language speakers, aware of the importance of how tones phonemically distinguish words, tend to have a better control of how their pitch changes in specific time within an utterance. For example, native speakers of Mandarin Chinese understand that learning tonal patterns is essential when learning each word. That explains the sensitivity of native speakers to perceptual cues that distinguish the tones, and it also explains why non-native speakers face a challenge when listening or producing tones for which they lack sensitivity. Gandour (1983) conducted a study about perceptual processing of tones and compared native tonal language speaker with non-native tonal language speakers (English speakers). He found that the English speakers assigned more weight to unrelated perceptual cues (F0 height) and paid less attention to the more important perceptual cues such as the pitch contour. Therefore, instructors who teach non-native tonal language speakers to acquire this new language might need to do more than providing only auditory model and feedback, which depends on the listener’s perception (specially at the beginning of the training period), to guide them to the correct production. One approach suggested by researchers to aid second language learners through their learning process is the use of a visual feedback. The idea of utilizing visual feedback during speech training is not new; rather, it can be traced back to Vardanian (1964), who used visual feedback when teaching English to non-native speakers. Other researchers also used visual feedback to teach intonation to deaf speakers (Anderson, 1960; Abberton and Fourcin, 1975).

More recently, many researchers advocate visual feedback as a way to teach intonation of a second language (Leather, 1990; Stibbard, 1996; Weltens & De Bot, 1984). Many investigators who taught a second language intonation to speakers with different language backgrounds (for discussion of the use of the Visi-Pitch, see Albertson, 1982) reported the effectiveness of such feedback. In one of the first attempts to use visual feedback, De Bot (1983) investigated the effectiveness of two modes of feedback (visual vs. auditory) and practice duration on learning English intonation (variation of pitch) by Dutch speakers. The researcher randomly assigned the participants into one of six groups: 1) a control group (pre-test + post-test, with no practice sessions); 2) another control group (pre-test + received instructions about intonation+ post-test, with no practice sessions); 3) a visual + auditory feedback group (45 minutes of practice); 4) a visual + auditory feedback group (90 minutes of practice); 5) an auditory feedback group (45 minutes of practice); and 6) an auditory feedback group (90 minutes of practice. The improvement score, measured as the difference between pre-test and post-test scores. The results showed that participants in the visual + auditory feedback group outperformed participants in the auditory feedback group. However, no difference in improvement score emerged between groups based on the practice duration. De Bot (1983) concluded that a visual feedback is useful in learning intonation of a second language because it encourages learning in a different way. De Bot (1983) also noticed that compared to the participants in the auditory feedback group, those in the visual + auditory group demonstrated the following: they chose to repeat the sentences more often; made more attempts to correct their sentence production; and made more efforts to match the model. De Bot attributed this behavior to the motivation that the visual feedback provided to the participants; visual feedback served as a reference of correctness for these individuals. In the De Bot study, the participants practiced the sentences at their own pace for a predetermined session duration; however, they were not limited by a set number of practice trials.

With advances in computer technology, the use of visual feedback is becoming more available to learners as a self-monitoring resource. Verdugo (2006) used a computer program to assist Spanish speakers to learn English intonation. Using a pre/post-test design, this study investigated the effectiveness of a multiple sensory approach (including auditory, visual) and also speech production practice to increase the awareness of the Spanish speakers to English intonation. The researcher randomly assigned the participants into one of two groups: 1) practice group—enrolled in 10 weeks practice, received information to increase their awareness about English intonation, practiced producing sentences with different intonation, and received visual feedback on their production; and 2) control group—received no practice. The results of the acoustic analysis of the learners' speech showed that, although both the control and the experimental groups did not significantly differ on the pre-test, the groups differed significantly on the post-test. On the post-test (performed during the last session of practice), the experimental group's speech showed more variety in inflection which resembled more closely the intonation of native speakers. The native English speakers' perceptual judgment indicated higher intelligibility and improvement in the experimental group's speech but not in the control group's speech, confirmed this finding.

The researcher concluded that the intonation awareness and practice with a visual display "has served to draw learners' attention to the prosodic organization of speech and the function of intonation in communication" (Verdugo, 2006, p.153). In the Verdugo study, the acoustic analysis consisted of deciding if the learners' and the native speakers' utterances coincided in terms of the tone used; the study presented the results in percentage of coincidence. Verdugo, however, performed no further acoustic analysis to quantify the degree of the difference between the learners' and the native speakers' utterances.

As discussed above, researchers have demonstrated that the optimum practice in second learning acquisition includes the following: 1) have the learner listen to the target utterance (either by the experimenter or from a recorded model); 2) present features of the model utterance visually to the learner—pitch contour; and 3) provide visual feedback after the learner production. Presenting both the model's and the learner's utterances enables the learner to compare both pitch contours and encourages the learner to change his/her utterance on subsequent trials to approximate the model utterance (Leather, 1990; Spaai & Hermes, 1993; Stibbard, 1996; Weltens & De Bot, 1984).

For Mandarin Chinese tone production training, Chun (1989) recommended such visual feedback. For this visual feedback, the computer screen is divided into upper and lower halves. The upper half presents a visual feedback of the native speaker's utterance, and the lower half portrays the learner's utterance F0. Such a visual display of the pitch contour is supposed to help make the learner more conscious of differences in the pitch contour pattern between his/her own production and the model. Not only does visual feedback provide a level of detail, especially with the difficulty the learners face in tone perception, but it is purported to help the learners attend to acoustic features that need to be modified. Such visual feedback especially would benefit those individuals who did not have previous experience with a tonal language (Hermes, 1998; Leather, 1990; Molholt, 1988; Weltens & De Bot, 1984; Öster, 1998).

Most of the available literature on Mandarin Chinese acquisition as a second language has focused on perception or recruited participants with some experience in tonal languages. Leather (1990) conducted a series of studies that considered the acquisition of the tonal system of standard Chinese by Dutch speakers from both perceptual and production perspectives. In an attempt to study whether the perception and production learning of Chinese tones are interrelated, Leather utilized a counterbalanced design and randomly assigned the participants into one of two groups. One group underwent perceptual training to differentiate between tones. Following perceptual training, their ability to produce the tones was evaluated. The second group first practiced tone production—with visual feedback—and then was tested on tone perception.

The results from this study indicated that perceptual training helped the learners produce the tones, and tone production training was helpful for learners in the other group to perceptually identify tones. From these results, Leather concluded that training in one modality, perception or production, is shown to be sufficient to enable a learner to perform in the other modality.

These results also showed that visual feedback used in this study was an effective method in helping participants with no tone experience to produce the tones. Although the previously reviewed studies involved learners with normal speech or hearing, it is worth mentioning that the use of the visual feedback is well established and utilized by speech-language pathologists in research and in treatment dealing with speech and hearing pathologies.

In most of the above reviewed studies, teachers or native speakers perceptually judged the tone production of learners to evaluate the effectiveness in a second language learning setting. As the review in the following section indicates, another study has considered acoustic analyses of tone production. Wang, Jongman and Sereno (2003) focused on the effect of Mandarin tones perceptual training on production as will be reviewed in the following section. Although the Wang et al. study posed different questions, by comparing the acoustic analysis of tone production among speakers and between different stages during training, these researchers highlighted the acoustic analysis of tone production.

Sixteen native speakers of American English participated in Wang et al. study. All the participants had spent either one or two semesters studying Mandarin Chinese. In a pre/post-test design, the researchers randomly assigned the participants to one of two groups (eight participants in each group): a control group and a perceptual training group. The control group performed only in the pre-test and the post-test, which were given two weeks apart. In addition to the pre-test and post-test, the perceptual training group participated in a perceptual training program for two weeks. For both the pre-test and the post-test, the researchers required the participants of both groups to read the same set of 80 monosyllabic words written with tonal diacritics (20 monosyllabic words for each of the four tones). The participants' tonal production was recorded for further analysis. Although the participants in the perceptual training group received tonal perceptual training on 40 of the 80 monosyllabic words stimuli, which they produced during the pre-test (researchers labeled these trained stimuli as "old stimuli"), they did not receive any perceptual training on the remaining 40 monosyllabic words (researchers labeled these untrained stimuli as "new stimuli").

The researchers then asked 80 Native Mandarin speakers to perceptually judge the participants' production. The perceptual judgment results were measured as percentages of correct tone production. The results indicated a significant group by test interaction. Although both groups were comparable on their pre-test scores, the results showed that the perceptual training group outperformed the control group on the post-test not only in the old trained stimuli but also in the new untrained stimuli. The control group performance in the post-test did not indicate any improvement and did not significantly differ from the pre-test.

The researchers further analyzed the perceptual judgment of “the old stimuli” for each participant and for each tone. These results indicated that, although the improvement was consistent across all tones, Tone 3 was significantly worse than the other three tones; this showed that the participants had difficulty in correctly producing this tone. In addition to using perceptual judgment, the researchers also acoustically analyzed the participants’ production in the perceptual training group. Wang et al. utilized the WAVES+ESPS software to derive the pitch contours for all the productions. For the acoustic analysis, they compared the extracted pitch contours between pre/post productions in relation to the native speakers’ norms. The researchers recorded the speech of four native speakers while producing the same 40 monosyllabic “old stimuli” to serve as the native speakers’ norms.

Major sources of differences between the participants’ pitch contour in acoustical output resulted from the individual differences in the F0 or speech rate differences; to eliminate such variations in the acoustic signal, the researchers normalized the pitch contours for all the participants. Normalization, a mathematical analogue which aims to extract the invariant acoustic correlates, utilizes a formula commonly employed for this purpose; such normalization allows direct comparison of normalized F0 values among each other (Rose, 1987). The F0 normalization equation yields F0 values that range from one to five which correspond to the five point scale in Mandarin Chinese. Moreover, the researchers normalized the pitch contour duration to overcome any differences in speech rate. The native speakers’ normalized pitch contours were averaged to present the norms for this study. Similarly, for the participants’ normalized pitch contours, two averages were calculated: a pre-test pitch contour average and a post-test pitch contour average. This signal preparation yielded three contours averaged across the participants (pre-test, post-test, and native speakers) for each of the four tones. The average pitch contours for the pre-test, post-test, and native speakers were compared based upon the following characteristics: 1) pitch values at 0% (onset), 25%, 50%, 75%, and 100% (offset) of the pitch contour; 2) F0 range, the difference between the highest F0 value and the lowest F0 values; 3) falling pitch range—from onset to valley; and 4) rising pitch range—from valley to offset.

In order to quantify the resemblance of the pitch contour between the pre-test and post-test relative to the pitch contour of the native speakers' norms. Wang et al. measured a deviation score as the dependent variable. The deviation score was calculated by measuring the difference between the normalized F0 values between two pitch contours at specific points of the pitch contour (at 0%, 25%, 50%, 75%, and 100% of the syllable duration). The results showed that the deviation score of the post-test pitch contour was significantly smaller than the deviation score of the pre-test pitch contour in relation to the pitch contour of the native speakers' norm. This finding indicates a closer resemblance between the post-test pitch contour and the native speakers' pitch contour after the perceptual training. The acoustic analysis results were consistent with the perceptual data; Tone 1, Tone 2, and Tone 4 demonstrated a greater similarity to the native speakers' pitch contour and also received a higher percentage of correctness when perceptually identified by the native speakers. Although the participants improved their Tone 3 production in the post-test, they demonstrated difficulty with this tone in terms of the pitch height; this difficulty was also consistent with a low score when perceptually judged.

Wang and her colleagues interpreted the results of their study as indicating that perceptual training of Mandarin tones improved the participants' production without further production training. A review of this article highlights the acoustic analysis. In the Wang et al. study, the perceptual judgment evaluated the tone production; the acoustic analysis added more details as to the original motor consequences that led to the acoustic output (the F0 changes in time as a result of vibration of the vocal tract). Having considered both the perceptual and acoustic analyses as complementing each other, this study provided a more complete picture of the nature of the production of tonal speech task by non-native speakers.

The preceding studies utilized the tone production task from a language perspective in order to facilitate or monitor the learners' progress in a second language program. Nonetheless, tone production is a complex speech task for native English speakers. On the one hand, English speakers utilize intonation to express emotions, intention, and grammatical form at the sentence level during which the F0 might increase or decrease at a slower rate. On the other hand, native Mandarin Chinese speakers vary their F0 at the lexical level in much faster ways.

From a motor perspective, the four tones—produced for the same syllable—require different control at the level of the vocal folds and coordination of the laryngeal and supralaryngeal movements.

As described earlier, each of the tones in the Mandarin Chinese has a unique pitch contour. Each pitch contour, from an articulation perspective, requires changes in the rate of the vocal fold vibration as well as the contraction and relaxation of laryngeal muscles (Lindqvist 1972, Ohala & Ewen 1973, Ohala 1978). Therefore, except Tone 1, the pitch contours for the other tones demonstrate more pitch range which also changes direction (turning point) for some tones; such change in the pitch contour entails complicated control at the level of the vocal folds.

Moreover, the duration and turning point temporal properties differ among the Mandarin Chinese tones. For example, a pitch rise usually requires a longer duration than a pitch fall (Sundberg, 1979). As suggested by Xu (2004), the dynamical limits inherent to the articulatory system can explain the time it takes for the pitch to change. Accordingly, Tone 2 and Tone 3 are longer in duration than Tone 4 (falling pitch). Although Tone 2 and Tone 3 seems to share the characteristic of being both longer in duration, they tend to differ in terms of their turning point—the point at which the pitch changes from falling to raising. The turning point for Tone 2 tends occur earlier than that of Tone 3. This temporal event plays an essential role in perception to differentiate Tone 2 and Tone 3. Interestingly, this temporal event is robust for each Tone and is independent of speech rate (Liu, Kuhl, & Tsao, 2007). The robustness of the turning point accord with the notion of the generalized motor program (GMP) the invariant characteristics, as discussed in the motor learning section.

According to Schmidt, GMP is an abstract representation for a class of movements. All movements within that class share invariant characteristics among them such as the relative time and force. On the other hand, this class of movement can vary in their absolute timing and overall force. Although it is not clear what is to consider as a GMP in speech production, the Mandarin Chinese tones differ among each other in their relative timing in which the pitch changes in direction (turning point). According to Schema theory, tasks with different relative timing are proposed to be governed by different GMP. Following the same logic, is seems reasonable to assume that the four tones might be governed by different GMP. This assumption raises an interesting issue, unlike the prediction of Maas et al (2008) in their review paper, that motor programs for speech would be generalized in the sense that all productions of a particular syllable are governed by the same GMP; the assumption that each tone might be governed by different GMP expand the dimension of speech GMP by including the tones required to produce the syllables in tonal languages.

The tonal task has many characteristics that render it a good candidate for research purposes in the speech domain: 1) it is a novel task for monolingual speakers of American

English; 2) it is a challenging speech task; 3) it has a motor component and an effect of movement component.

This section reviewed the tone learning literature from the second language acquisition as well as from the motor perspectives. From the motor perspective, the differences among the tones in terms of the rate of pitch change and its timing would appear to impose different challenges to the learners.

5.0 RESEARCH QUESTIONS

The proposed study is motivated by the robust findings in the literature regarding the beneficial effect of adopting EFOA on learning in a variety of motor tasks. Moreover, although instruction and feedback are integral parts in any clinical speech learning setting, their effects in terms of FOA have not been investigated. As such, the proposed study is the first that tests the role of FOA in the speech domain. The purpose is to examine the effects of FOA on learning a novel tonal task. The specific research questions for this purpose are:

5.1 PRIMARY RESEARCH QUESTIONS

- Are there significant differences in the slope for the root mean square error (RMSE) scores across the acquisition phase of the experiment among the three groups: EFOA, IFOA, Control?
- Are there significant differences in the overall RMSE scores during the retention phase of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the overall RMSE scores during the transfer phase of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the slope for the percentage of the correctly perceived words across the acquisition phase of the experiment among the three groups: EFOA, IFO, Control?
- Are there significant differences in the percentage of correctly perceived words during the retention phases of the experiment among the three groups: EFOA, IFO, and Control?
- Are there significant differences in the percentage of correctly perceived words during the transfer phase of the experiment among the three groups: EFOA, IFO, and Control?

5.2 SECONDARY RESEARCH QUESTION

A secondary research question is posed to determine whether learners would acquire each of the tones at the same level of accuracy or would show some differential difficulty with some tones as demonstrated in the second language learning literature.

- Are there significant differences in the percentage of correct productions among the four tones for each group, during the acquisition phase: Tone 1, Tone 2, Tone 3, and Tone 4?

6.0 METHODS AND PROCEDURES

6.1 PARTICIPANTS

There is no available literature on the effect of FOA on the learning of a speech task with which to estimate the sample size. Therefore, the sample size was estimated based on non-speech tasks by a power analysis. The power analysis revealed that a sample size of 42 subjects (14 per group) is required to detect a large effect size (f = .4) (Cohen, 1988) with a power = .80, at α =.05 (Cohen, 1988). Forty-two females participated in this study. Although the literature offers no evidence of gender differences in a second language speech sound learning, this study excluded males to decrease any variability due to the physiological and structural differences in the vocal folds that produce gender differences in the fundamental frequency (Kent & Read, 2002; Klatt & Klatt, 1990; Titze, 1989).

The participants were recruited through both the Department of Psychology-Research Participation-at University of Pittsburgh, and the Clinical Translational Science Institute (CTSI) Research Participants' Registry. All participants met the following inclusion criteria:

females between the ages of 18 - 24 (M = 18.9, range = 18-24); passed a pure-tone hearing screening test at 500, 1000, 2000, and 3000 Hz, and presented at 25 dB HL; passed a vision screening using the reduced Snellen Chart with aided or unaided vision of 20/40 or better; demonstrated the ability to discriminate between two pure tones—either presented at the same frequency or at different frequencies (see appendix A, for the tone-pair frequencies)—with an accuracy score of 75% or more (Bradshaw & McHenry, 2005); and had adequate vocal function: screened by the Computerized Speech Laboratory (CSL) software using the Multi-Dimensional Voice Profile (MDVP) and Real-Time Pitch (participants scored within one SD of normative data on: F0, Shimmer, Jitter, Noise to-Harmonic ratio).

As self-report verified, all the participants 1) were monolingual speakers of American English without experience with tonal language (such as: any dialect of Chinese, Cantonese, Thai, Lao, Vietnamese, Kru , Khoisan); 2) had no vocal training, no private speaking/singing lessons or fewer than five private voice training lessons; 3) lacked any advanced musical training (less than one year of music training or no training at all); and 4) had no learning deficit (neither diagnosed with a learning deficit nor registered at the office of Disability Resources and Services at the University of Pittsburgh). The participants were screened for the aforementioned inclusion criteria by answering a short questionnaire (included in Appendix B).

All participants followed the approved IRB protocol and received an explanation of the study procedures and the assurance that their participation was voluntary and that they could withdraw from the study at any time. Rather than informing the participants about the specific purpose of the experiment, the researcher only emphasized that the study dealt with learning a novel speech sound. All participants signed the informed consent form prior to the initiation of any screening or data collection.

After completion of the study, all participants recruited through the Psychology Research Pool received course credits; participants recruited through CTSI earned $25 as a compensation for their time.

6.2 INSTRUMENTATION

Adobe Audition Version 1 (Adobe Systems, Inc.) was utilized to record the instructions and the training stimuli. To ensure that the instructions were standardized within each condition, the instructions were digitally recorded by a speaker whose native language is English.

PRAAT, a comprehensive signal analysis, software (Praat5108), was used to extract the fundamental frequency (F0) at specific points on the pitch contour for the acoustic analysis. A "*Stimulate*" program designed specifically for this study, was utilized to control the timing of the experiment's events as follows: 1) presenting the auditory model; 2) recording the participants produced monosyllabic words; 3) initiating a PRAAT script to extract the F0 of the recorded auditory model; 4) presenting the visual feedback (pitch contours of both the learner and the model) during the acquisition phase; and 5) storing the sound files for further analysis.

6.3 STIMULI AND EXPERIMENTAL TASK

This study used four Mandarin Chinese monosyllabic words during the acquisition phase. All four words were practiced with the same Ma consonant-vowel (CV), but the syllable was presented with four different tones to differentiate the word meaning in Mandarin Chinese. In addition, two other syllables were utilized to assess transfer, Me and Na. Each of these syllables was produced with four tones, which yielded eight monosyllabic words in Mandarin Chinese (see Table 2 for citation of these monosyllabic words and their meaning in English). The tonal task was a good task for the current study because it is a novel task for monolingual speakers of

American English and because it is a complex motor speech task. The tonal task production requires the participants to control the rate of their vocal folds vibration in a specific pattern to produce the target tone; the pattern of change of F0 at the syllable level is novel to monolingual English speakers. For this study, the participants repeated the Chinese monosyllabic words after hearing an auditory model presented to them.

The auditory model for the 12 Chinese monosyllabic words was generated by digitally recording a female native speaker of Mandarin Chinese. For the auditory model recording, the speaker produced each word three times in a random order; the best of these productions, as judged by a native Mandarin speaker, served as the model for the current study. The stimuli were recorded in a sound booth and were digitalized at a sampling rate of 44.1 kHz using Adobe Audition Version 1 (Adobe Systems, Inc.). The stimuli were then edited by the same software to obtain a suitable loudness; a 50-millisecond silence was inserted at the beginning and at the end of each recorded syllable. Before presenting these words to the participants, the monosyllabic words were randomized within experimental phases. The stimuli were then transferred to the computer to be presented to the participants. The recorded words from the native speaker were used as auditory models, and the model's pitch contour was used as a visual feedback during the acquisition phase.

Table 2   Citation of the monosyllabic Mandarin Chinese words and their meaning in English

Monosyllabic Word	Tone 1	Tone 2	Tone 3	Tone 4
Ma	Mother	Hemp	Horse	To Scold
Na	OK[1]	Take	Where	That
Me	To Squint	Mystery	Rice	Secret

Note. [1] Na produced with tone 1 does not have a corresponding word in Mandarin Chinese. For this study, a meaning was assigned to this word by indicating that Na produced with a high tone in Thai language (another tonal language), means OK in English. Although this stimulus does not have a meaning in Mandarin Chinese, it can be easily judged by Mandarin Chinese native speakers as Na produced with tone 1 based on the assigned tone.

6.4 INDEPENDENT VARIABLE - DEFINITIONS OF TERMS

As discussed in the literature review, the benefits of external focus of attention (EFOA) on the performance and learning of motor skills not only extended to a wide range of motor skills but also generated research in other fields. Consequently, the purpose of this study was to test the generalizability of the EFOA effects in the speech domain.

This study is the first experiment to investigate the role of FOA in the speech domain. Therefore, this section first discusses how studies on FOA in motor skills defined what constitutes an external or internal focus of attention based upon the employed tasks and then establishes which aspect of the speech task utilized in this study was considered either EFOA or IFOA.

From the reviewed studies (see Appendix C - Tables 1-5 for a summary of the instructions utilized in some of the reviewed studies and see the literature review for a more detailed description), the EFOA instructions can be characterized as follows: 1) focus on the effect of the movement on the environment or on an implement (instrument) and 2) define the focus point as external to the body, with a task-related EFOA. The IFOA instructions can be characterized as follows: 1) focus on the body part performing the movement or the motor task and 2) accentuate the conscious control of the body part performing the movement. Although the EFOA and IFOA instructions seem to differ on the surface, they share many basic characteristics: 1) the instructions in both conditions directed the participant to achieve the same goal; 2) the wording was comparable except for some key words used to direct the participants' attention to either the effect of the participants' movement on an instrument or to the participants' body movements, respectively, and 3) FOA researchers did not require the participants to look at their FOA locus points, but instead asked the participants to look straight and to perform the task as instructed.

Most of the reviewed tasks in the FOA literature required the participants to interact with an instrument during practice. This made the effect of the movement (such as maintaining balance on a stabilometer) obvious to determine. Nonetheless, performance in other motor skills, such as swimming, dancing, and speech, does not require such an interaction with an instrument, which makes the definition of EFOA less obvious. Wulf (2007) proposed possible EFOA and IFOA focus points for motor tasks that have not yet been studied and, therefore, might have a less clear definition (see Wulf, 2007; p 62- 65).

Formulating instructions for this proposed study was unique in that the speech task, unlike most motor skills studied in the FOA literature, does not involve an interaction with an instrument. However, Wulf (2007) suggested that for public speaking, focusing on the projection of the voice might represent an EFOA, while focusing on the vibratory sensation from the vocal folds might constitute an IFOA. Moreover, because speech is considered as “a set of movements made audible” (Stetson, 1951, p.33), in this study, it seemed logical to focus on the produced sound or on the vocal folds vibration in EFOA and IFOA, respectively, for the task utilized in the current study.

This study, therefore, defined EFOA and IFOA as follows:

- EFOA refers to the sound that is the effect of the vocal folds vibration (movement) on the environment; and
- IFOA refers to the vibration of the vocal folds in the larynx.

The wording of the instructions in this study had much of the same content-as those in the FOA literature- but differed in some key words that help direct the participants’ FOA either to the produced sound in the EFOA or the movements within the larynx (vibration of the vocal folds) in the IFOA. Following the recommendation of Wulf (personal communication, August 4, 2010), the EFOA instructions of this study also excluded the word YOUR in order to avoid any possible internal focus effect for the EFOA instructions. After providing the general instructions about the task goal, the following instructions were presented to the EFOA and IFOA:

6.4.1 Instructions for the EFOA group

- Visually look at the computer screen. I want you **to focus on the sound.** When feedback is presented, and the sound is off target, think about how to correct this by changing **the sound**. During each trial, **focus on the produced sound**.

6.4.2 Instructions for the IFOA group

- Visually look at the computer screen. I want you **to focus on the vibration of your voice box.** When feedback is presented and you are off target, think about how to correct this by changing **from within your voice box**. During each trial, **focus on the vibration in your voice box**.

The researcher adapted the wording from Lohse et al. study (2010) that investigated the effect of FOA on a dart throwing task. These adapted instructions are comparable to those in the FOA literature; only subtle differences in word choices appear. Furthermore, the researcher formulated general instructions for this study. These general instructions explained the goal of the practice:

6.4.3 The general instructions

- In this experiment, you will practice producing words in Mandarin Chinese. I want you to listen to each word and repeat it as close to the recording as possible. Although the words might sound the same, they are four different words conveying different meaning according to how they are produced. Listen to the word and repeat it when prompted.

6.5 PROCEDURES

The participants were randomly assigned to one of three experimental groups: 1) external focus of attention group (EFOA), 2) internal focus of attention group (IFOA), or 3) control group (C). The three groups underwent the same practice; only the instructions differed for each group. Participants engaged in the experiment individually, not as part of a group.

The experimental set up was first explained to the participants. The participants were familiarized with the feedback that they would receive during the acquisition phase. An example of the visual feedback was described and presented on the computer monitor before the beginning of the practice: the researcher told the EFOA group the feedback showed how the produced sound changed over time, while the researcher told the IFOA group the feedback showed how the vibration in the voice box changes over time. All of the participants' questions were answered prior to the start of the experiment. Participants were seated comfortably in a sound-attenuated booth. A condenser microphone (omnidirectional) was positioned 30 cm away from the participants' mouth, and the auditory model was presented through speakers in the sound booth. The participants' task in this study was to repeat the Mandarin Chinese monosyllabic words after hearing the auditory model for each word. In accordance with the literature that demonstrated that lower feedback frequency enhances learning (Adams & Page, 2000; Steinhauer & Grayhack, 2000), participants in all conditions received feedback on 60% of their trials during the acquisition phase. The feedback was presented on a 12 x 9 inch computer screen that was placed in front of the participants. The *"Stimulate"* program controlled the protocol of the experiment by: 1) presenting the instructions both in written form and as a playback of the recorded instructions; 2) presenting the auditory model; 3) recording the participants' responses; 4) providing the visual feedback; and 5) storing the data for additional analyses.

The experiment consisted of four phases: baseline phase, acquisition phase, retention tests (immediate and delayed), and transfer test. The experiment was completed in two sessions on two consecutive days (see Figure 2). On the first day, the participants signed the consent form, performed the screening tests, and completed the baseline recordings, the acquisition phase, and the immediate retention test. Twenty-four to 48 hours after the first session, the participants returned to complete the delayed retention and the transfer tests. At the end of the experiment, the participants answered a short questionnaire about the performed task and the experiment. The number of trials was predetermined for each phase (as will be discussed later).

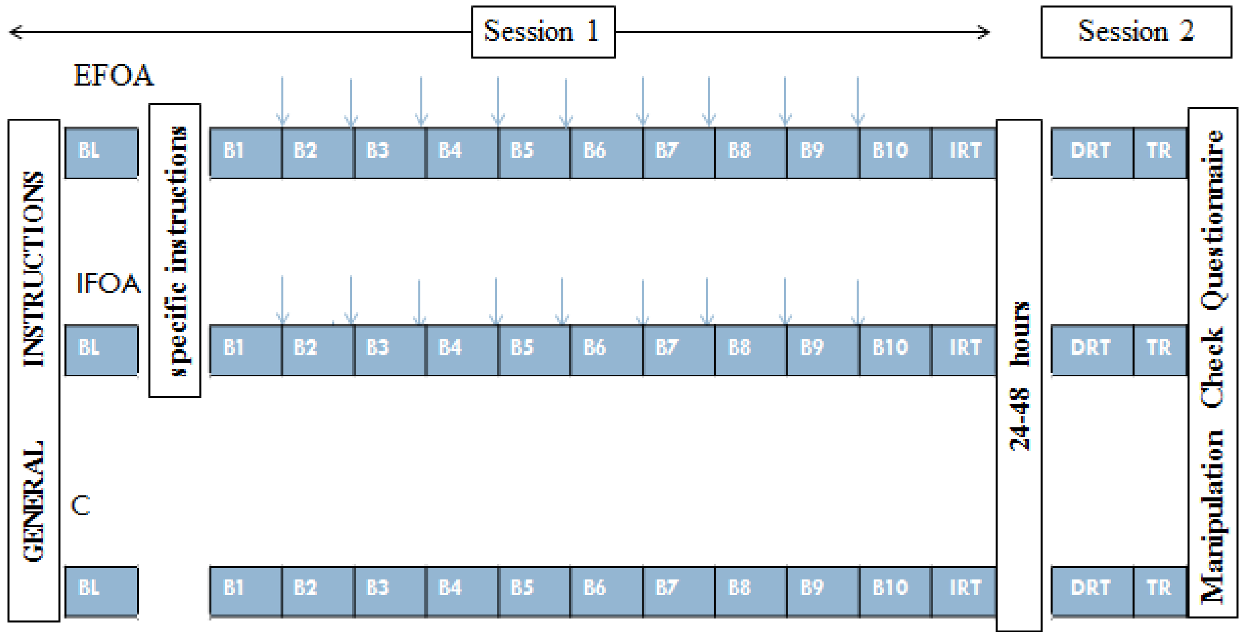


Figure 2 Experimental protocol for the three experimental groups; EFOA = external focus of attention, IFOA = internal focus of attention; BL = baseline; Bi=block (i= 1 to 10); IRT = immediate retention; DRT = delayed retention; and TR = transfer test. ↓ Each arrow indicates the repetition of the specific instructions for the EFOA and the IFOA groups.

6.5.1 Experiment phases

6.5.1.1 Baseline

Before the baseline recording, all participants received the general instructions (see the independent variable section for the general instructions); the general instructions appeared both in a written form on the computer monitor and as a verbal playback of the previously recorded instructions. During the baseline phase, the participants listened to each of the 12 tonal monosyllabic words and repeated each word after the auditory model. The 12 words, which were randomly ordered, were presented five times (12 words X 5 times each = 60 productions). The participants received no feedback on their performance during the baseline recording. All produced words were recorded and saved for further analysis.

6.5.1.2 Acquisition

Prior to the beginning of the acquisition phase and before the first trial, the participants in both the EFOA group and the IFOA group received additional specific instructions, according to their assigned condition, on what to focus on during the production task. However, participants in the control condition did not receive any additional instructions.

During the acquisition phase, the participants practiced repeating the four monosyllabic Mandarin Chinese words after hearing the auditory model; the four monosyllabic words utilized the same Ma syllable (see Table 1). Controlled by the “Stimulate” program, the computer first presented the auditory model of the target tonal monosyllabic words. The participants’ responses were recorded. After production, the participants received a visual feedback on the computer screen; the upper half of the screen presented the model’s pitch contour with the English meaning of the produced word printed next to the contour, and the lower half of the screen presented the learner’s pitch contour displaying the changes in F0 over time (See figure 3). Researchers have demonstrated the effectiveness of this type of visual feedback in studies that deal with learning Mandarin Chinese as a second language; such feedback allows the participants to monitor their success in achieving the tonal task goal (Albertson, 1982; Chun (1989; Leather, 1990; Stibbard, 1996; Weltens & De Bot, 1984). Moreover, according to the Schema theory, the availability of such augmented feedback is linked with strengthening both the recall and recognition schema (Schmidt, 1975). The presentation of the feedback was delayed for three seconds after the participants' response; such a feedback delay encourages the learners to actively engage in evaluating their response and enhance error detection and learning (Schmidt & Lee, 1999; Swinnen et al., 1990). Based upon the recommendation of researchers in the motor learning literature, the visual feedback was provided on 60% of the practice trials. Feedback was provided on 12 of the 20 trials; the participants received feedback on three of each five word repetitions.

The acquisition phase consisted of 10 blocks of 20 trials each, containing a total of 200 trials (50 trials X 4 words) of practice (see Table 2 for a description of the number of trials for each phase). Participants in both the EFOA and the IFOA groups received the FOA instructions on the screen and also heard them before every block during the acquisition phase.

The four words were presented to the participants in a quasi-random order within each block of practice, with no word being consecutively repeated more than twice. This approach follows the limb and speech motor learning literature that suggest that random practice might benefit learning (Adams & Page, 2000, Hall & Magill, 1995; Lee, Wulf, Schmidt, 1992), especially if the practiced tones are governed by different GMPs. The participants completed the acquisition trials in one session which lasted approximately one hour. Although the participants were told that they can request a break on their convenient, none of them requested a break in addition to the assigned breaks. The experimenter included a one-minute break between the blocks of practice and a longer break up to five minutes after the fifth practice block and. The number of practice trials in the current study was based on previous studies that required young adults to produce novel speech tasks in a motor learning paradigm (Adams & Page, 2000; Kim, 2007; Steinhauer & Grayhack, 2007).

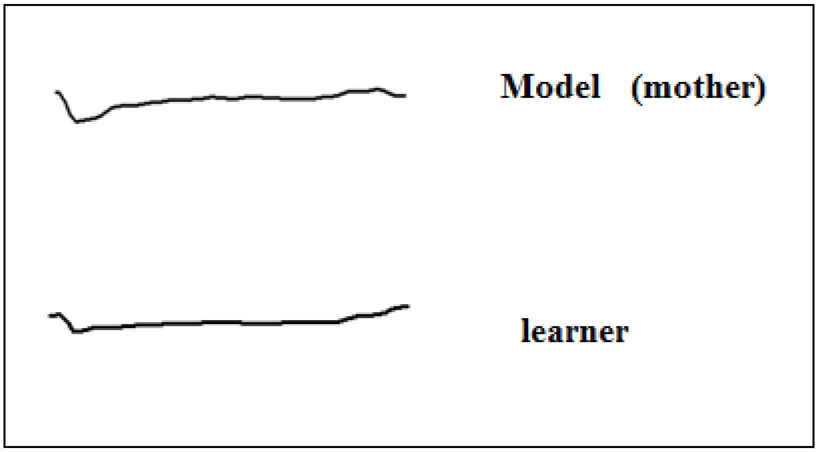


Figure 3 A snapshot of the visual feedback; the upper half of the screen has the model pitch contour with the English meaning of the produced monosyllabic Mandarin Chinese word; the lower half of the screen contains the learner pitch contour.

6.5.1.3 Learning tests (retention and transfer tests)

Before the beginning of each learning test, the participants only received the general instructions as a reminder of the speech task goal. During the performance of the retention tests (immediate and delayed retention tests) and the transfer test, the participants did not receive feedback on their performance.

At the end of the acquisition phase, which occurred in the first session, the participants performed an immediate retention test. During the retention testing, the participants repeated the same practiced monosyllabic words after hearing the model. The retention test entailed five repetitions for each practiced monosyllabic word (see Table 3). On the second session, given 24-48 hours after the first session, the participants performed a delayed retention test and a transfer test. The delayed retention test entailed the same productions as the immediate retention test; the only difference was that this delayed retention test was performed after 24-48 hours from the immediate retention test.

During the transfer test, the participants produced the practiced tones with unpracticed syllables after hearing the auditory model: Me and Na, produced the practiced tone with a different vowel /e/ and a different consonant /n/, respectively. Similar to the other learning tests, participants received only general instructions before the start of the transfer test; they did not receive feedback on their performance. The two stimuli used for the transfer test were based on the Ballard et al. (2007) assumption that Ma and Na might be governed by the same GMP but might require different parameters. Following this logic, the vowels /a/ and /e/ might also be governed by the same GMP. As such, generalization to these similar but untrained stimuli can be expected. All the monosyllabic words produced by the participants were saved for further analyses.

6.5.1.4 Probe testing of transfer test words

To assess the participants' ability to generalize the trained tone to a similar but untrained word, the transfer test words were probed throughout the acquisition phase at specific times: 1) after block-1 at the beginning of the practice; 2) after block-4 towards the middle of the practice; and 3) after block-10 towards the end of the practice. Given that the instructions were presented before every block, the probes were presented towards the end of the blocks to reduce the effect of the specific instructions on the probe production. Moreover, the participants did not receive feedback on their performance on the probe productions.

During the baseline phase, the transfer tests words, produced with Me and Na syllables, were presented in random order with the training stimuli (Ma monosyllabic words produced with the four tones). Each of the eight transfer words was presented five times. The participants repeated all words during the baseline phase after the auditory model; they did not receive feedback on their productions.

Instead of probing all the transfer test words during acquisition, the transfer test words were probed as follows in each experimental group: 1) half of the participants were probed on half of the transfer words (words produced with Tones 1 and 2) The other half of the participants were probed on the other half of the transfer words (words produced with Tones 3 and 4). To decrease any possible practice advantage on the production of the transfer words, not all of the words were probed during the acquisition phase. The rationale for pairing the tones in such a way (words produced with Tones 1 and 2 versus words produced with Tones 3 and 4) was to consider the hierarchy of tone acquisition reported in the literature. Therefore, in this study, the participants' probe testing was performed by matching one easy tone (either Tone 1 or Tone 4) with one more challenging tone (either Tone 2 or Tone 3).

As such, the participants produced the training stimuli and the probes during the acquisition phase. This yielded a total of 260 productions for the acquisition phase. At the end of the experiment, each participant had produced a total of 400 productions of the monosyllabic words. All these productions were saved for further analysis (See Table 3).

Table 3 Number of productions for each participant during each experimental phase

	Baseline	Acquisition	Immediate Retention	Delayed Retention	Transfer Test	Total Productions
Stimuli	20 X (Ma)	200	20	20		260
Transfer words Probes	20 X (Me) 20 X (Na)	20 trials X 3 probe testing times = 60	-----	-----	40	140
Total productions	60	260	20	20	40	400

6.6 MANIPULATION CHECK

Considering the conceptual nature of the independent variable—instructions to direct the participants' attention to specific aspects of the task—a manipulation check was utilized in order to determine whether the participants perceived and followed the instructions. After the completion of the learning tests at the end of the second session, all participants answered questions about different aspects of the experiment. However, only asking the participants about what they focused on during their practice might bias the participants to choose the focus given in the instructions. As an alternative, the questionnaire included questions about different aspects of the experiment: feedback; the task itself; amount of practice; the adopted FOA; and the participants' perceived task difficulty. The questionnaire utilized in this study was adapted from Porter, et.al., 2010 and Fasoli, et.al., 2002; see Appendix D.

6.7 DEPENDENT VARIABLES

For this study, two dependent variables were measured: 1) acoustic measure—the root-mean square error (RMSE) and 2) perceptual measure—the percentage of the correctly perceived production (%C). The next section provides details on how these dependent variables were measured.

6.8 DATA ANALYSIS

6.8.1 Acoustic analysis

The similarity between the learners' and the native speaker's pitch contours was quantified by calculating the RMSE as a measure of how much the learners' pitch contours differed from those of the model's pitch contours. RMSE is a frequently-used measure to quantify the differences between sets of values and to measure the similarity between two curves**.** The higher the resemblance between the two pitch contours, the smaller the RMSE value. One RMSE value was obtained for each production (a total of 400 RMSE for each participant).

Because the vowel is the tone baring part of the syllable (Xu, 2004), the RMSE measurement was performed on the vowel portion of the produced monosyllabic words. This acoustic analysis of the vowel required several steps. First, the beginning and end of the vowel were identified. Using the PRAAT software, the onset of the vowel was determined from the waveform and the spectrogram of each production. Because the consonant in the monosyllabic words produced in this study were the nasals M and N, an abrupt change in energy occurred between the two: the nasal consonant and the vowel could be detected on the waveform and was marked as the onset of the vowel. This abrupt change in energy results from the difference in the nature of the nasal consonants and the vowels.

During the production of the voiced nasal consonants, most of the high frequency energy, produced by the vibration of the vocal folds, is absorbed by the nasal cavity; this phenomenon is manifested as the nasal murmur, which is characterized by low frequency (300 Hz) on the spectrogram. On the waveform, the nasals are characterized by a relatively simple waveform, almost like a sine wave with low amplitude, due to the energy absorbance in the nasal cavity. When the oral cavity opens for the vowel production, the waveform demonstrates an abrupt change in its shape; the waveform changes from the simple waveform to a more complex wave. This complex wave indicates the vowel production. Specifically for this study, the point that was marked as the vowel onset was considered at the zero crossing of the first complex waveform (Ferrand, 2007; Fry, 2001). Figure 4 shows this point of the vowel onset on a waveform and on a spectrogram. The offset of the vowel was indicated as the point of abrupt attenuation of the signal amplitude. The vowel onset and offset were also checked by superimposing the formant tracks on a wideband spectrogram in the PRAAT software.

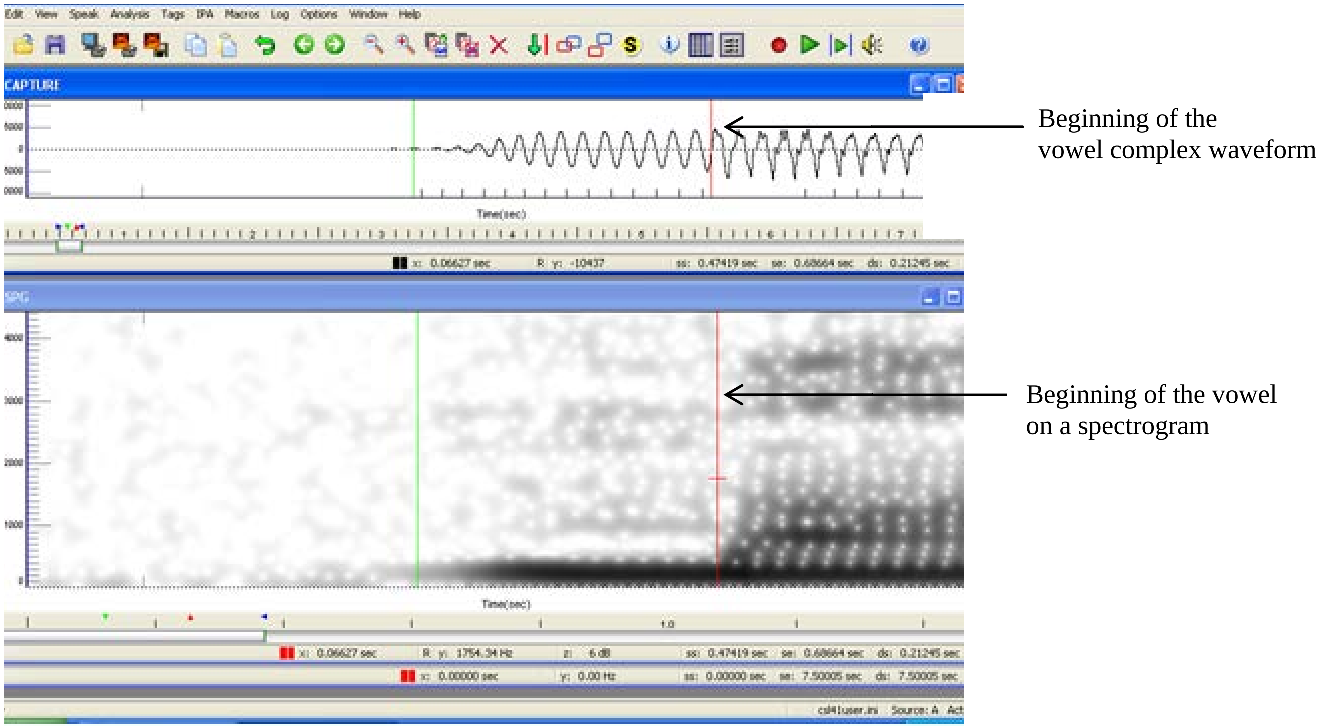


Figure 4 Snapshot of the CSL screen showing a waveform and a spectrogram for Ma production; the vertical line indicated the vowel onset.

Secondly, the pitch listing function in the PRAAT software was used to extract the fundamental frequency (F0) of the vowel. The extracted F0 values were exported to an Excel spreadsheet. Considering the vowel duration, the F0 was selected at the beginning of the vowel—at every 10% of the vowel duration—and at the end of the vowel. Thirdly, in order to directly compare the difference between the F0, both the participant's productions and the native speaker's auditory model were normalized. Such normalization accommodates any pitch range differences among speakers. To normalize F0, converting the F0 values to their logarithms, a commonly used formula for F0 normalization was applied (e.g., Rose, 1987; Wang, 2003):

$$(T = [(\lg X - \lg L)/\lg\mathrm{H} - \lg\mathrm{L}]\mathrm{X}5)$$

Where H and L represent the highest and the lowest F0, respectively, for a given speaker across the four tones, and X represents the value of F0 at any point of a pitch contour. The output of this formula is a T value that ranges from one to five. The T value accords with the five points pitch scale proposed by Chao (1984) to describe Mandarin Chinese tones. Lastly, the RMSE on the T values (normalized F0) was calculated using the following formula:

$$\sqrt{\frac{\sum\left(f(x_i)-y_i\right)^2}{n}}$$

Where χ and y represent the points compared from the two pitch contours, and n represents the number of all data points compared. Measuring F0 at specific points of the vowel is a method used by researchers interested in tonal languages (Liu, Tsao & Kuhl, 2007, Wang, Jongman, & Sereno, 2003; Xu, 2004).

This acoustic analysis, which yielded one value for every word production, indicates the similarity between each participant’s pitch contour and the native speaker’s pitch contour.

6.8.2 Perceptual analysis

Mandarin Chinese native speakers performed this analysis to obtain the percentage of correctly perceived words by the participants. The perceptual analysis was utilized to assess the effect of FOA instructions on the tone production. The percentage of correctly perceived production was calculated by dividing the number of correctly perceived monosyllabic words by the total number of monosyllabic words produced in that phase or block of practice. This part of the analysis was performed off-line after the completion of all data collection.

6.8.2.1 Evaluators

Three native speakers of Mandarin Chinese, all graduate students at the University of Pittsburgh, evaluated the participants' tone productions. The evaluators had normal hearing as self-reported.

6.8.2.2 Speech stimuli for the perception analysis

The speech stimuli for the perception analysis consisted of all the recorded monosyllabic words produced by the participants. The words, produced by each participant during all the experiment phases, were randomized. In addition, 20 stimuli (five percent) were selected randomly from each participant's production and again introduced to the listeners to check the intra-rater reliability. The total number of speech stimuli evaluated by the Mandarin native speakers was 17,640 (420 productions X 42 participants).

6.8.2.3 Procedures

Before the beginning of the perceptual judgment, all three Mandarin Chinese speakers listened—together—to the auditory model and a random sample of the participants' productions.

This session was performed to insure that all listeners agreed with the model and how to judge the words. After this group meeting with the researcher, the listeners then individually performed the perceptual judgments; every listener evaluated all the productions from the 42 participants. Each evaluator scheduled the sessions to listen to the productions at her convenience (Three-four sessions, approximately two hours each).

The evaluators understood that this was a forced choice rating task; they were also familiar with the words used in this study. The evaluators, who sat in front of a computer, listened to the words via earphones at a self-selected comfortable volume level. Each evaluator's task was to listen to each word and decide which tone she heard by pressing one of five specific keys on the keyboard for each trial: T1, T2, T3, T4, or "none".

The evaluators determined the pace of the stimulus presentation by pressing a "next" button on the keyboard to present the next word. They also had the option to repeat any word if they decided to listen again to any word before making their decision.

Agreement among two or three raters determined the response accuracy. A percentage score of correctly perceived productions was calculated by dividing the number of the correctly perceived production by the total number of trials in that block or phase.

7.0 STATISTICAL ANALYSIS

This study utilized a mixed model design. The between-subjects factor was the instructions—the independent variable—with three levels: EFOA, IFOA and C. The within-subject factor was the four tones. In some of the analyses, time was also considered a within-subject factor, as described in the following sections. For this study, two dependent variables were measured: 1) acoustic measure—the root mean square error (RMSE) and 2) perceptual measure—the percentage of correctly perceived production (%C). To assess the effect of the instructions on learning the monosyllabic Mandarin Chinese words, either a three-way mixed ANOVA, a two-way mixed ANOVA, or a non-parametric equivalent test was used according to the nature of the dependent variable. The results section provides more details and justification for the utilization of each analysis.

7.1 ACOUSTIC DATA STATISTICAL ANALYSIS

To determine that no significant difference on the RMSE appeared among the groups at the baseline, a three-way ANOVA was performed on the RMSE. The between-subjects factor was the group with three levels: EFOA, IFOA and C. The within-subject factors were the syllables (Ma, Me, and Na) and Tones (1, 2, 3, 4).

To determine the effect of the instructions on performance during the acquisition phase, a two-way mixed ANOVA was performed on the slopes as a function of the groups (EFOA, IFOA and C) and the four tones. For the retention test, a 3 X 4 X 2 mixed ANOVA was performed on the RMSE as a function of the groups (EFOA, IFOA and C), the four tones and the time (immediate retention and delayed retention). The group was the between-subjects factor, and the tones and time were the within-subject factors. For the transfer test, a 3 X 4 X 2 mixed ANOVA was performed on the RMSE as a function of the groups (EFOA, IFOA and C), the four tones and the syllables (Me and Na). The group was the between-subjects factor, and the tones and syllables were the within-subject factors.

This study, the first to assess the effect of FOA instructions on the speech domain, based its analysis on the Barlow and Hersen (1984) suggestion that a combination of both inferential and single subject analyses would be appropriate. Therefore, this study utilized also a single subject analysis (SSA) to assess the effect of the instructions on the performance. The SSA entails analyzing each participant's performance on the target behavior during different time points; because each participant serves as her own control, such detail provides a closer inspection of the effect of FOA on each individual performance. In this study, the SSA performed on the RMSE data. For the single subject analysis, each participant's RMSE, during baseline, acquisition, and retention was plotted separately for each tone. The qualitative analysis of the single subject examination involved the visual inspection of the data pattern, slopes, and variability for baseline, acquisition, and retention phases. To quantify the magnitude of change in the performance, an effect size was calculated using the Tau-U analysis, a new index for single subject analysis which combines the nonoverlap data points and trends in the data into its calculations (Parker, Vannest, Davis & Sauber, 2011).

7.2 PERCEPTUAL DATA STATISTICAL ANALYSIS

For this part of the analysis, non-parametric statistics were utilized when the data was not normally distributed. To determine whether the groups had significant difference on the percentage of correctly perceived productions—during baseline, acquisition, retention, and transfer, the statistical analysis were performed as follows: 1) to assess the between-subjects factor, the Kruskal-Wallis test was performed as a function of groups, and 2) to assess whether the tones were perceived with the same accuracy, the within-subject factor, the freedman-test was performed for each group as a function of the tones.

8.0 RESULTS

The purpose of this study was to examine the effects of FOA on learning a novel speech task: English native speakers learning Mandarin Chinese tones. Researchers have demonstrated that an EFOA benefited the performance and learning on several motor tasks. Based on the literature discussed in chapter Four, it was predicted that participants in the EFOA group would show less error than those in the IFOA and the C groups, relative to the Chinese native speaker's tones. When judged perceptually by native Chinese listeners, the monosyllabic words produced by the EFOA group would have a higher percentage of correctly perceived words than the IFOA and C group. In the single subject analysis, participants in the EFOA group would demonstrate a larger effect sizes in their performance during the acquisition phase compared to their baseline performance compared to participants in the IFOA and C groups.

In this study, two analyses determined the progress of the participants in learning the novel monosyllabic Chinese words: the acoustic analysis and the perceptual analysis. The results of these analyses are presented in separate sections.

For the acoustic analysis, the dependent variable is the root mean square error (RMSE), representing the difference between each participant's pitch contour and the pitch contour of the model—Chinese native speaker. The pitch contour represents the change in the speaker's fundamental frequency (F0) over the duration of the syllable production—specifically the syllable coda (the vowel). Each of the four tones has its unique pitch contour.

In order to assess whether the differences among the productions were perceptually detected, three Mandarin Chinese native speakers judged the tones productions of the English native speakers. The dependent variable for the perceptual analysis was the percentage of correct productions in each phase: baseline, acquisition phase, retention, and transfer.

8.1 THE ACOUSTIC ANALYSIS

The acoustic analysis is presented first to answer all three research questions regarding differences among the groups on the RMSE in the acquisition phase, retention tests, and transfer test. In addition, the data for each participant was examined.

The experimental task in this study included the production of the four tones of Mandarin Chinese under three instructional conditions. The literature of second language learning of Mandarin Chinese suggests that some of these tones (Tone-1 and Tone-4) are considered easy to produce and to perceive, while others are deemed difficult (Tone-2 and Tone-3). Therefore, the effect of FOA on the production of these tones was separately examined for each tone.

In order to test whether all groups were equivalent at baseline, the RMSE was compared among the groups. A three-way ANOVA was performed on the RMSE as a function of groups (EFOA, IFOA, C), tones (1, 2, 3, 4) and syllables (Ma, Me, Na). The group was the between-subject factor; the tones and syllable were the within subject factors. The dependent variable was the RMSE. The assumption of sphericity was not met for the tones, Mauchly's $W = 0.173$, χ^2

(5) = 66.27, $p < .001$; and for the syllables, Mauchly's $W = 0.712$, χ^2 (2) = 12.89, $p < .002$. The Hynh-Feldt test was used to assess potential main effects and interactions. All other assumptions were met. The means and standard deviations for each group productions at each tone are presented in Table 4. Results of the three-way ANOVA are presented in Table 5.

Table 4 The means and standard deviations of the RMSE in each monosyllabic production among the three groups during baseline

Syllables	Group	*Mean*	*SD*
Ma-1	EFOA	1.18	0.65
	IFOA	1.27	0.72
	C	1.37	0.94
Ma-2	EFOA	1.53	0.64
	IFOA	1.92	0.80
	C	1.75	0.71
Ma-3	EFOA	1.82	0.78
	IFOA	2.23	0.87
	C	2.04	0.77
Ma-4	EFOA	1.76	0.60
	IFOA	1.83	0.73
	C	2.07	0.80
Me-1	EFOA	1.26	0.68
	IFOA	1.21	0.70
	C	1.38	0.86

Table 4 (continued)

Me-2	EFOA	1.46	0.66
	IFOA	1.75	0.75
	C	1.64	0.80
Me-3	EFOA	2.09	0.79
	IFOA	2.39	0.95
	C	2.21	0.74
Me-4	EFOA	1.35	0.69
	IFOA	1.42	0.62
	C	1.74	0.72
Na-1	EFOA	1.20	0.68
	IFOA	1.26	0.72
	C	1.33	0.81
Na-2	EFOA	1.40	0.55
	IFOA	1.72	0.75
	C	1.62	0.72
Na-3	EFOA	1.98	0.83
	IFOA	2.47	0.97
	C	2.10	0.83
Na-4	EFOA	2.11	0.54
	IFOA	2.24	0.83
	C	2.55	1.02

Table 5 Results of the three-way ANOVA—3 (Groups) X 4 (tones) X 3 (syllables)—on the RMSE during baseline.

Interactions and main effects	Statistic	*P*	partial η^2
Group X Tone X Syllable	*F* (12, 234) = 0.51	0.82	0.03
Group X Tone	*F* (6, 117) = 1.36	0.26	0.06
Group X Syllable	*F* (4, 78) = 0.62	0.62	0.03
Tone X Syllable	*F* (3, 117) = 27.62	<.001	0.83
Group Effect	*F* (2, 39) = 0.53	0.60	0.03
Tone Effect	*F* (3, 117) = 33.91	<.001	0.46
Syllable Effect	*F* (2, 78) = 20.28	<.001	0.34

As summarized in Table 5, the results of the three-way ANOVA indicated that there was no significant difference on the RMSE among the groups. This finding suggests that the baseline performance of the three groups was comparable.

However, during baseline, the pattern of difference on the RMSE among the three syllables was significantly different among the four tones as indicated by the significant two-way (syllable X tone) interaction (Table 5). The main effect of tone and the main effect of syllable were both significant (Table 5). The mean RMSE for each syllable and tone, during baseline, is displayed in Figure 5.

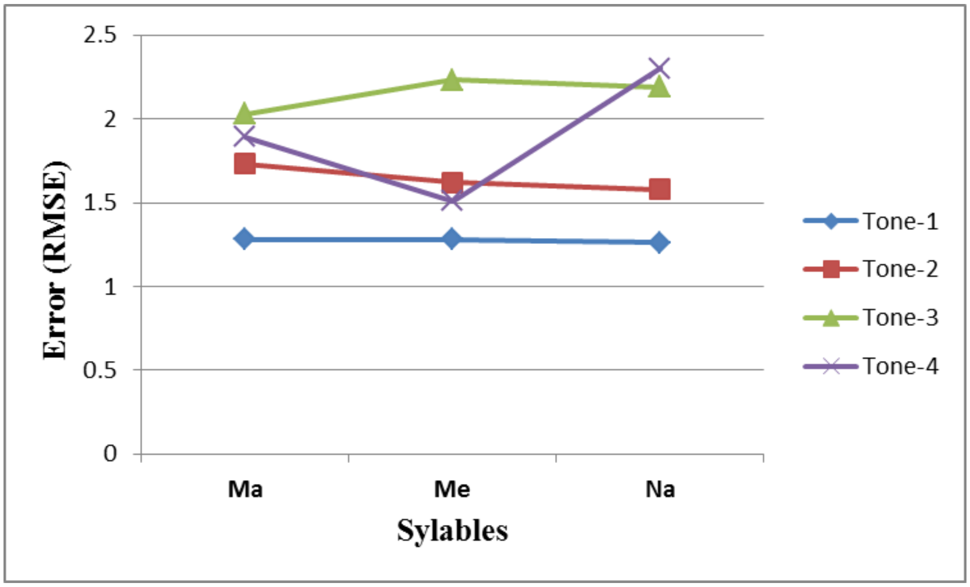


Figure 5 The RMSE of the produced words during baseline, by syllable and tone

In order to find the pattern of difference on the pitch contour RMSE among the tones, simple main effects analysis was computed on the RMSE for each syllable. There was a significant difference on the RMSE among the four tones for all Ma syllables, Me syllables, and Na syllables (F (3,117) = 26.00, p < .001, partial η^2 = 0.40, F (3,117) = 31.96, p < .001, partial η^2 = 0.45, F (3,117) = 47.96, p <.001, partial η^2 = 0.55 respectively).

In order to locate the pattern of difference on the RMSE among the four tones for each syllable, pairwise comparisons were performed among the tones for each syllable at α = 0.01. For the Ma syllable, means (with standard deviation in parentheses) for Tone 1 through Tone 4 were 1.28(0.76), 1.73(0.72), 2.03(0.18), 1.89(0.71), respectively. Participants produced Tone 1 with significantly lower RMSE than each of the other three tones: Tone 2, Tone 3, and Tone 4.

These data further show that participants produced Tone 2 with significantly lower RMSE than Tone 3. However, the RMSE for Tone 4 was not significantly different than either Tones 2 or 3. The results of all post hoc comparisons for the Ma syllables are presented in Table 6.

Table 6 Results of the post-hoc pairwise comparisons on the four tones RMSE of the Ma syllable productions during baseline, averaged across groups

Tone-pair	Mean Difference	*P*
Ma-1-Ma-2	-0.455	<.001
Ma-1-Ma-3	-0.756	<.001
Ma-1-Ma-4	-0.609	<.001
Ma-2-Ma-3	-0.301	<.001
Ma-2-Ma-4	-0.154	.069
Ma-3-Ma-4	0.148	.182

For the Me syllable, participants produced Tone 1 (M = 1.28, SD = 0.73) with significantly lower RMSE than each of the other three tones: Tone 2 (M = 1.62, SD = 0.73), Tone 3 (M = 2.23, SD = 0.82), and Tone 4 (M = 1.51, SD = 0.68). These data further show that participants produced Tone 2 with significantly lower RMSE than Tone 3. Moreover, participants produced Tone 4 with significantly lower RMSE than Tone 3. However, no significant difference on the RMSE between Tones 2 and 4 for the Me syllable emerged. The results of all post hoc comparisons for the Me syllables are presented in Table 7.

Table 7 Results of the post-hoc pairwise comparisons on the four tones RMSE of the Me syllable productions during baseline, averaged across groups

Tone-pair	Mean Difference	*P*
Me-1-Me-2	-0.333	<.001
Me-1-Me-3	-0.947	<.001
Me-1-Me-4	-0.225	.002
Me-2-Me-3	-0.614	<.001
Me-2-Me-4	0.108	.296
Me-3-Me-4	0.723	<.001

For the Na syllable, participants produced Tone 1 (M = 1.26, SD = 0.72) with significantly lower RMSE than each of the other three tones: Tone 2 (M = 1.58, SD = 0.68), Tone 3 (M = 2.19, SD = 0.88), and Tone 4 (M = 2.30, SD = 0.82). These data further show that participants produced Tone 2 with significantly lower RMSE than Tone 3. Moreover, participants produced Tone 2 with significantly lower RMSE than Tone 4. However, the RMSE between Tones 3 and 4 for the Na syllable showed no significant difference. The results of all post hoc comparisons for the Na syllables are presented in Table 8.

Table 8 Results of the post-hoc pairwise comparisons on the four tones RMSE of the Na syllable productions during baseline, averaged across groups

Tone-pair	Mean Difference	*P*
Na-1-Na-2	-0.319	.001
Na-1-Na-3	-0.923	<.001
Na-1-Na-4	-1.039	<.001
Na-2-Na-3	-0.605	<.001
Na-2-Na-4	-0.720	<.001
Na-3-Na-4	-0.115	.392

As the above data illustrate, the four tones were not produced with the same accuracy when measured by RMSE. There appeared to be a hierarchy for the error during the production of the three syllables: 1) for the Ma syllable, Tone 1 was produced with the lowest RMSE, followed by Tones 2 and 4, while Tone 3 scored highest on the RMSE; 2) for the Me syllable, Tone 1 was produced with the lowest RMSE, followed by Tones 4 and 2, while Tone 3 scored highest on the RMSE. 3) for Na syllable, Tone 1 was produced with the lowest RMSE, followed by Tones 2 and 3, while Tone 4 scored highest on the RMSE.

8.1.1 The acquisition phase

Research Question:

> Are there significant differences in the slope for the RMSE error scores across the acquisition phase of the experiment among the three groups (EFOA, IFOA, Control)?

During the acquisition phase, the participants practiced four monosyllabic Chinese words: Ma-1, Ma-2, Ma-3, and Ma-4. The English meaning of the four Chinese monosyllabic words, as typed on the feedback screen presented to each participant was: mother, hemp, horse, and to scold, respectively. The acquisition phase consisted of ten blocks of practice. Each block began with the presentation of the instructions for the EFOA and IFOA groups. To avoid examiner bias, the instructions were pre-recorded. The participants simultaneously read and listened to the instructions.

During each block of practice, the participants repeated the monosyllabic word after the recorded model and received visual feedback after their production. The feedback on the upper half of the computer screen represented the pitch contour of model; the feedback on the lower half of the screen showed the participant's produced pitch contour.

To examine the participants' performance during the acquisition phase, the slopes of the error measure (RMSE) were calculated from ten RMSE values for each participant. A 3 X 4 mixed ANOVA was performed on the average slopes as a function of the groups (EFOA, IFOA, C) and the four tones. The group was the between subject factor and the tones were the within subject factor. The assumption of sphericity was not met for the slopes, Mauchly's $W = 0.002$, χ^2

(5) = 66.27, p < .001. The Hynh-Feldt test was used to assess potential main effects and interactions. All other assumptions were met.

The means and standard deviations of the slopes are shown in Table 9. A negative slope indicates a decrease in the error (improvement), while a positive slope indicates an increase in the error during acquisition phase. Neither the main effects of group, F (2, 39) = 1.45, p = 0.25, partial ŋ2 = .023; of slopes F (3, 117) = 0.96, p = 0.41, partial ŋ2 = 0.02, nor the interaction, F (6, 117) = 0.68, p = 0.52, partial ŋ2 = 0.03, was significant.

Contrary to the prediction, the slopes for RMSE among the three groups for all the four tones showed no significant difference. From data in Table 9, it is apparent that the slope values are minimal. In order to assess whether these slopes were significantly different than zero, four separate t-tests were performed on the mean value of each tone slope averaged across the groups. As shown in Table 10, none of the slopes differed significantly from zero. Despite these non-significant results, the control group slopes showed a tendency towards error reduction during the acquisition phase for both Ma-3 and Ma-4 with slopes -0.018 and -0.017, respectively. On the other hand, the slopes of the EFOA of Ma-1 and IFOA of Ma-2 showed a smaller improvement trend with slopes -0.003 and -0.011, respectively.

For the main effect of groups during the acquisition phase, the observed effect size was .11 and the observed power was .21. Based on the observed effect size (.11), sample size determination showed that a total of 204 participants (68 per group) would be needed to detect an effect size of .11 with an alpha of .05 and a desired power of .80.

Table 9 The means and standard deviations of the error RMSE slopes in Ma syllables production in the three groups.

Slopes	Groups	Mean	SD
Slope-ma1	EFOA	-0.003	0.017
	IFOA	0.005	0.013
	C	0.006	0.012
	$\bar{x}$	0.002	0.014
Slope-ma2	EFOA	0.004	0.026
	IFOA	-0.011	0.026
	C	0.007	0.019
	$\bar{x}$	0.000	0.024
Slope-ma3	EFOA	0.011	0.040
	IFOA	0.008	0.029
	C	-0.018	0.057
	$\bar{x}$	0.000	0.044
Slope-ma4	EFOA	0.003	0.035
	IFOA	0.007	0.019
	C	-0.017	0.033
	$\bar{x}$	0.002	0.031

Table 10 Results of four t-tests, comparing the mean acquisition slopes—for each tone averaged across groups mean slope—to a test value of zero

Slope	t-test	*df*	*P*
Slope-ma1	1.066	14	.293
Slope-ma2	-0.994	14	.326
Slope-ma3	0.010	14	.992
Slope-ma4	-0.425	14	.673

8.1.2 Learning tests

To assess learning, the following two questions were proposed:

Research Questions:

Are there significant differences in the overall RMSE scores during the retention phase of the experiment among the three groups: EFOA, IFO, and Control?

Are there significant differences in the overall RMSE scores during the transfer phase of the experiment among the three groups: EFOA, IFO, and Control?

According to the motor learning literature, learning is defined as the ability of the individual to reproduce the practiced task either immediately after completing the practice sessions (immediate retention) or after some delay that could be hours or days (delayed retention). Learning is also tested by a transfer test. The retention test shows the performer's ability to maintain and produce the practiced behavior after the practice; this performance reflects the strength of the recall schema. Unlike the retention test, the transfer test requires that the individual to produce an untrained task similar to the trained task; the performance reflects the strength of the recall schema in choosing the most appropriate parameters in this novel situation to determine if the learned behavior was generalized to other related tasks.

8.1.2.1 Retention tests

During the retention test, the participants produced the practiced words without specific instructions or feedback. The study involved two retention tests: an immediate retention test (IRT) performed immediately after the completion of the practice and a delayed retention test (DRT) completed 24-48 hours after the IRT.

The means and standard deviations of the RMSE for the four monosyllabic words during IRT and DRT phases are shown in Table 11. A 3 X 4 X 2 mixed ANOVA was performed on the RMSE as a function of groups (EFOA, IFOA, C) and the four tones and time (immediate retention, delayed retention). The group was the between subject factor and the tones and time were the within-subject factors. The assumption of sphericity was not met for the tones, Mauchly's $W = 0.108$, $\chi^2 (5) = 83.883$, $p < .001$. The Hynh-Feldt test was used to assess potential main effects and interactions. All other assumptions were met. Results of the three-way ANOVA are presented in Table 12. For the retention tests, only the main effect of tone was significant.

Table 11 The RMSE means and standard deviations by the group and the phase for each tone

		IRT		DRT	
Word	Groups	*M*	*SD*	*M*	*SD*
Ma-1	EFOA	1.23	.758	1.26	.683
	IFOA	1.39	.730	1.39	.783
	C	1.50	.874	1.50	.912
	$\bar{x}$	1.37	.778	1.39	.785
Ma-2	EFOA	1.50	.632	1.52	.588
	IFOA	1.95	.834	1.90	.873
	C	1.82	.830	1.71	.802
	$\bar{x}$	1.76	.776	1.71	.763
Ma-3	EFOA	1.83	.702	1.92	.788
	IFOA	2.23	.910	2.19	.924
	C	1.98	.767	1.99	.899
	$\bar{x}$	2.01	.796	2.04	.830
Ma-4	EFOA	1.85	.650	1.94	.631
	IFOA	1.79	.726	1.87	.815
	C	2.08	.826	2.09	.711
	$\bar{x}$	1.91	.730	1.96	.711

Note. IRT = immediate retention test; DRT = delayed retention test.

Table 12 Results of the three-way ANOVA performed on the RMSE as a function of groups (EFOA, IFOA, C), tones (1, 2, 3, 4) and time (immediate retention and delayed retention)

Interactions and main effects	Statistic	*P*	partial η^2
Group X Tone X Time	F (6, 234) = 0.40	.88	0.02
Group X Time	F (2, 117) = 0.67	.52	0.03
Group X Tone	F (6, 117) = 1.70	.18	0.08
Tone X Time	F (3, 117) = 1.67	.18	0.04
Group Effect	F (2, 39) = 0.41	.67	0.02
Tone Effect	F (3, 117) = 24.04	<.001	0.38
Time Effect	F (2, 78) = 0.23	.63	0.01

During the retention phase, only the practiced Ma monosyllabic words were tested. Participants produced Tone 1 (M = 1.38, SD = 0.78) with significantly lower RMSE than each of the other three tones: Tone 2 (M = 1.73, SD = 0.76), Tone 3 (M = 2.03, SD = 0.81), and Tone 4 (M = 1.94, SD = 0.72). These data further showed that participants produced Tone 2 with significantly lower RMSE than Tones 3 and 4. However, no significant difference on the RMSE between Tones 4 and either Tone 2 or Tone 3 emerged (see Figure 6). The results of all post hoc comparisons for the Ma syllables during the retention tests are presented in Table 13.

Table 13 Results of the post-hoc pairwise comparisons on the four tones RMSE of the Ma syllable productions averaged across retention tests and groups

Tone-pair	Mean Difference	*P*
Ma-1-Ma-2	-0.345	<.001
Ma-1-Ma-3	-0.645	<.001
Ma-1-Ma-4	-0.556	<.001
Ma-2-Ma-3	-0.291	<.001
Ma-2-Ma-4	-0.202	.012
Ma-3-Ma-4	0.089	.396

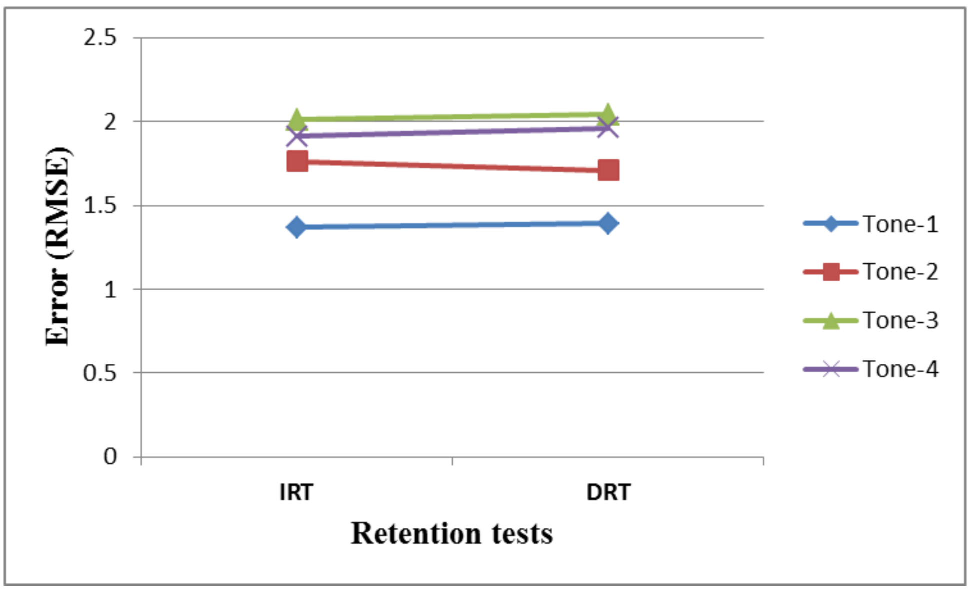


Figure 6 Mean RMSE for each tone produced during the IRT and DRT, averaged across groups

For the main effect of groups during the retention phase, the observed effect size was .1 and the observed power was .18. Based on the observed effect size (.1), sample size determination showed that a total of 246 participants (82 per group) would be needed to detect an effect size of .1 with an alpha of .05 and a desired power of .80.

8.1.2.2 The transfer test

While training utilized Ma syllable in four tones, transfer was tested on other monosyllabic words. During the transfer test, the participants produced the same practiced tones with either another phoneme N in the same vowel context (Na) or with the same phoneme M in another vowel context (Me). Each of these monosyllabic words was produced with the four tones, resulting in eight monosyllabic words during the transfer test. Participants received no specific instructions or feedback during the transfer test. The means and standard deviations of the RMSE for the words produced during the transfer test are shown in Table 14.

A 3 X 4 X 2 mixed ANOVA was performed on the RMSE as a function of groups (EFOA, IFOA, C) and the four tones and syllables (Me, Na). The group was the between subject factor and the tones and syllables were the between subject factor. The assumption of sphericity was not met for the tones, Mauchly's $W = 0.225$, χ^2 $(5) = 51.57$, $p < .001$. The Hynh-Feldt test was used to assess potential main effects and interactions. All other assumptions were met. Results of the three-way ANOVA are presented in Table 15.

Table 14 The means and standard deviations of the RMSE in each syllable production among the three groups during the transfer test

Word	Groups	*Mean*	*SD*
Me-1	EFOA	1.18	.80
	IFOA	1.26	.64
	C	1.46	.88
	$\bar{x}$	1.30	.77
Me-2	EFOA	1.34	.56
	IFOA	1.67	.70
	C	1.71	.82
	$\bar{x}$	1.58	.70
Me-3	EFOA	2.11	.85
	IFOA	2.20	.91
	C	1.92	.78
	$\bar{x}$	2.08	.84
Me-4	EFOA	1.56	.80
	IFOA	1.53	.62
	C	2.55	1.87
	$\bar{x}$	1.88	1.29
Na-1	EFOA	1.16	.72
	IFOA	1.32	.73
	C	1.46	.90
	$\bar{x}$	1.31	.78

Table 14 (continued)

Na-2	EFOA	1.31	.52
	IFOA	1.70	.71
	C	1.57	.77
	$\bar{x}$	1.53	.68
Na-3	EFOA	2.02	.78
	IFOA	2.39	1.04
	C	2.06	.86
	$\bar{x}$	2.16	.89
Na-4	EFOA	2.27	.62
	IFOA	2.39	.85
	C	2.69	.86
	$\bar{x}$	2.45	.79

Table 15 Results of the three-way ANOVA performed on the RMSE as a function of groups (EFOA, IFOA, C), tones (1, 2, 3, 4) and syllables (Me, Na), during the transfer phase

Interactions and main effects	Statistic	*P*	partial η^2
Group X Tone X Syllable	*F* (6, 234) = 1.43	0.25	0.07
Group X Tone	*F* (6, 117) = 3.19	0.01	0.14
Group X Syllable	*F* (3, 39) = 1.93	0.16	0.09
Tone X Syllable	*F* (3, 117) = 8.90	0.002	0.83
Group Effect	*F* (2, 39) = 0.72	0.49	0.04
Tone Effect	*F* (3, 117) = 35.34	<.001	0.47
Syllable Effect	*F* (1, 39) = 20.28	.005	0.19

In order to determine whether the pattern of RMSE values among the syllables was significant, a simple main effect of the tones was computed for each syllable. There was a significant difference on the RMSE among the four tones for both Me and Na syllable (F (3,117) = 11.78, $p < .001$, partial $\eta^2 = 0.23$, F (3,117) = 59.76, $p < .001$, partial $\eta^2 = 0.60$, respectively).

For the Me syllable, participants produced Tone 1 (M = 1.30, SD = 0.77) with significantly lower RMSE than each of the other three tones: Tone 2 (M = 1.58, SD = 0.71), Tone 3 (M = 2.10, SD = 0.84), and Tone 4 (M = 1.88, SD = 1.29). These data further show that participants produced Tone 2 with significantly lower RMSE than Tone 3. However, the RMSE of Tone 4 was not significantly different than Tone 2 or 3. The results of all post-hoc comparisons for the Me syllables during the transfer test are presented in Table 16.

Table 16 Results of the post-hoc pairwise comparisons on the four tones RMSE of the Me syllable productions during transfer test, averaged across groups

Tone-pair	Mean Difference	*P*
Me-1-Me-2	-0.275	<.001
Me-1-Me-3	-0.779	<.001
Me-1-Me-4	-0.582	.002
Me-2-Me-3	-0.504	<.001
Me-2-Me-4	-0.306	.07
Me-3-Me-4	0.197	.285

For the Na syllable, participants produced Tone 1 (M = 1.31, SD = 0.78) with significantly lower RMSE than Tone 3 (M = 2.16, SD = 0.89) and Tone 4 (M = 2.45, SD = 0.79). These data further show that participants produced Tone 2 (M = 1.53, SD = 0.68) with significantly lower RMSE than Tones 3 and 4. However, no significant difference between either Tones 1 and 2 or Tones 3 and 4 emerged. The results of all post hoc comparisons for the Na syllables during the transfer test are presented in Table 17.

Table 17 Results of the post-hoc pairwise comparisons on the four tones RMSE of the Na syllable productions during transfer test, averaged across groups

Tone-pair	Mean Difference	*P*
Na-1-Na-2	-0.213	.01
Na-1-Na-3	-0.843	<.001
Na-1-Na-4	-1.139	<.001
Na-2-Na-3	-0.630	<.001
Na-2-Na-4	-0.926	<.001
Na-3-Na-4	-0.296	.02

As the above data illustrate, the four tones were not produced with the same accuracy when measured by RMSE during the transfer phase. There appeared to be a hierarchy for the error during the production of the two syllables: 1) for the Me syllable, Tone 1 was produced with the lowest RMSE, followed by Tones 2 and 4, with Tone 3 scoring highest on the RMSE. 2) For the Na syllable, Tone 1 was produced with the lowest RMSE, followed by Tones 2 and 3, while Tone 4 scored highest on the RMSE.

For the main effect of groups during the transfer phase, the observed effect size was .11 and the observed power was .39. Based on the observed effect size (.11), sample size determination showed that a total of 105 participants (35 per group) would be needed to detect an effect size of .11 with an alpha of .05 and a desired power of .80.

As the result section of the acoustic analysis showed, when the average scores of the three groups were considered, the inferential statistics failed to detect significant difference among the groups on the RMSE in any of the experiment's phases. For the statistical analysis performed above, the RMSE was aggregated across subjects within each FOA group. Because this is the first study to investigate the role of FOA on speech motor learning, a closer inspection through single subject analysis of the data is warranted. Single subject analysis provides a more detailed examination of the effects of FOA on the individual's performance, as each participant serves as her own control. Detecting any change in performance during the acquisition phase would require closer inspection. The acquisition phase was only considered in the previous analysis in terms of averaged slopes—of the RMSE values-among the groups. Additional details would provide a better understanding of the acquisition blocks while the participants, in the EFOA and IFOA group, practiced the task under the specific instructions. Therefore, the next section addresses the single subject analysis during the acquisition and the retention phases in relation to the baseline phase.

8.1.3 Single subject analysis

This study is the first to investigate the four Mandarin Chinese tones from a speech motor learning perspective. The study's single subject analysis is beneficial because it provides a closer look at each participant's performance, which is essential for capturing any improvement for each participant. The RMSE values were plotted as a function of time as follows: 1) five data points representing baseline productions; 2) ten data points representing the acquisition phase (each point represents the average of five trials performed in one block); and 3) five data points representing retention phase productions. The data points were connected to show the data path; a separate line graph for each tone was constructed for each participant. The line of best fit was drawn for each phase, and its slope was calculated. Appendix E presents these figures for each participant. In order to judge whether the treatment had an effect, the figures were visually analyzed to indicate any change in level, trend, slope, or variability as a result of the intervention. The visual inspection of the data was supported by measuring an effect size index to quantify the magnitude of change and by measuring the variability of the data points.

The data of all participants are presented in Appendix E. It should be noted that the baseline data were not stable for most of the participants; instead, the data points indicated either an increasing or decreasing trend or increased variability. Ideally, a stable baseline is preferred to better interpret the data; any trends in baseline, specifically in the direction of improvement, challenge the interpretation of the results. To quantify the magnitude of change, the current study utilized a Tau-U effect size statistic. The Tau-U is a non-parametric method utilized to measure the effectiveness of treatment by measuring the percent of points during the intervention phase that do not overlap with the baseline data points (Parker, Vannest, Davis & Sauber, 2011); "judging data overlap between phases has been a part of visual analysis since at least the 1960s, along with judging the trend" (Parker et al. (2011), p. 285).

The Tau-U method is similar to the percentage of non-overlapping data (PND) analysis (Scruggs, Mastropieri, & Casto, 1987) in that both measures yield a percent score to indicate an improvement index. However, the Tau-U method is different from the PND analysis in that it considers both the nonoverlap data points and any trends in the phases in the effect size measurements. If the baseline trend is significant (Tau-U value $\geq$ 40%), the Tau-U analyses control for the baseline trend in a measured and conservative way that does not distort the original data. The Tau-U output score, usually expressed as percent (or proportion), it is interpreted, similar to the PND, as the percent of improvement in the intervention versus the baseline. The Tau-U analysis also provides a p-value and a confidence interval to determine the significance of the score.

To indicate whether the baseline had a trend, the current study utilized a web-based Tau-U calculator (Version 1.0) (Vannest, Parker, & Gonen, (2011) to calculate the baseline phase Tau-U score. Following the criteria of Parker, et al. (2011), a Tau-U value of 40% or more would indicate a trend in the baseline. To quantify the improvement in performance, the Tau-U scores between the phases (acquisition vs. baseline and retention vs. baseline) was calculated for each participant, using the web-based Tau-U calculator (Version 1.0) (Vannest, Parker, & Gonen, 2011). Tables 18-41 present the results of these calculations. It should be noted that data for each group were presented in two separate tables: one table presented the data of those participants whose baseline did not demonstrate a trend (as indicated in the second column, baseline Tau-U score); the other table presented the data of those participants who did demonstrate a baseline trend (as indicated in the second column, baseline Tau-U score). The Tau-U measurement controlled for the baseline trend in the second table. Because this study utilized an error measure in the acoustic analysis (RMSE), an improvement is indicated by calculating the percent nonoverlap data with values below the lowest RMSE in the baseline. The bolded Tau-U scores in the Tables 18-41 denote an error reduction.

8.1.3.1 EFOA group

EFOA group: Ma-1

As Tables 18 and 19 present, only three of the 14 participants demonstrated a baseline with no trend. The remaining 11 participants demonstrated a baseline trend as a Tau-U score of 40% or more. During the acquisition phase, only two participants (32 and 43) demonstrated a significant decrease in the Tau-U effect size relative to baseline. Their improvement was also evident on the retention test. Interestingly, although participant 20 demonstrated a non-significant decrease in her score during acquisition, her retention effect size indicated a significant 100% non-overlap. Although the non-overlapping Tau-U scores for participants 6, 10, 20, and 22 indicate an error reduction, these scores were not significant. No other significant effect sizes, indicating an improvement, emerged.

EFOA group: Ma-2

As Tables 20 and 21 present, half of the participants in EFOA demonstrated a baseline trend during Ma-2 productions. Only two participants (22 and 43) demonstrated an improvement during the acquisition phase relative to baseline as indicated by their Tau-U effect sizes. Only participant 43 demonstrated a significant improvement during the retention phase relative to baseline (-100%). No other significant improvement was demonstrated.

EFOA group: Ma-3

In Tables 22 and 23, 10 of the 14 participants in EFOA demonstrated a baseline trend in the direction of error reduction; only one participant (6) demonstrated a baseline trend which indicated an increase in error. During Ma-3 production, the Tau-U effect sizes of only two participants (22 and 43) indicated error reduction during acquisition. No other significant error reduction emerged.

EFOA group: Ma-4

As Tables 24 and 25 present, eight of the 14 participants in EFOA demonstrated an increasing trend during baseline; only one participant (46) demonstrated a baseline trend in the direction of error reduction. The Tau-U effect sizes of participants 10 and 43 showed a significant improvement during both the acquisition and the retention phase for Ma-4 productions. The Tau-U scores of the other participants did not demonstrate significant error reduction.

Table 18 Results of the Tau-U analyses effect sizes for the EFOA group Ma-1 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
29	**20%**	42%	.19	12%	.75
32	33%	**83%**	.03	**80%**	.07
46	0	18%	.58	0	.29

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 19 Results of the Tau-U analyses effect sizes for the EFOA group Ma-1 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	Acquisition vs. baseline *p*	Retention vs. Baseline Tau-U	Retention vs. Baseline *p*
6	40%	**10%**	.76	**4%**	.92
10	60%	**36%**	.27	**60%**	.11
13	40%	6%	.85	36%	.35
20	90%	**20%**	.54	**100%**	.01
22	40%	**20%**	.54	**28%**	.46
33	100%	33%	.16	20%	.46
38	40%	50%	.01	25%	.05
42	40%	26%	.43	**4%**	.92
43	100%	**86%**	.01	**88%**	.02
47	40%	**58%**	.08	64%	.09

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 20 Results of the Tau-U analyses effect sizes for the EFOA group Ma-2 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	Acquisition vs. baseline *p*	Retention vs. Baseline Tau-U	Retention vs. Baseline *p*
6	0	64%	.05	92%	.02
10	0	40%	.22	20%	.60
29	**20%**	**62%**	.06	**60%**	.12
32	0	70%	.03	72%	.06
38	33%	15%	.67	100%	.81
43	0	**96%**	.003	**100%**	.01

Table 20 (continued)

55	33%	52%	.14	50%	.22

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 21 Results of the Tau-U analyses effect sizes for the EFOA group Ma-2 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
13	60%	40%	.27	28%	.46
20	**40%**	**56%**	.08	4%	.92
22	40%	**88%**	.01	**70%**	.09
33	**70%**	74%	.02	**8%**	.83
42	40%	94%	.004	60%	.12
46	40%	10%	.76	**12%**	.75
47	60%	66%	.04	36%	.34

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 22 Results of the Tau-U analyses effect sizes for the EFOA group Ma-3 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	p	Retention vs. Baseline Tau-U	p
10	0	12%	.71	52%	.17
13	10%	16%	.62	32%	.40
32	10%	30%	.36	36%	.35

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 23 Results of the Tau-U analyses effect sizes for the EFOA group Ma-3 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	p	Retention vs. Baseline Tau-U	p
6	60%	**34%**	.30	**36%**	.35
20	**67%**	**67%**	.08	**60%**	.92
22	**40%**	**42%**	.05	**44%**	.25
29	**40%**	**26%**	.43	**44%**	.25
33	**40%**	38%	.24	4%	.92
38	**67%**	22%	.62	15%	.71
42	**40%**	100%	.001	100%	.003
43	**80%**	**80%**	.01	**60%**	.12
46	**80%**	**44%**	.18	**12%**	.75
47	**40%**	40%	.18	80%	.04
55	**100%**	35%	.45	50%	.33

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 24 Results of the Tau-U analyses effect sizes for the EFOA group Ma-4 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
22	20%	18%	.58	**60%**	.12
29	0	56%	.09	16%	.68
43	**30%**	**86%**	.01	**88%**	.02
47	**20%**	64%	.05	52%	.17
55	33%	7%	.86	**20%**	.65

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 25 Results of the Tau-U analyses effect sizes for the EFOA group Ma-4 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
6	40%	8%	.81	68%	.07
10	90%	**80%**	.01	**88%**	.02
13	80%	**26%**	.42	8%	.04
20	40%	76%	.02	84%	.03
32	60%	**28%**	.39	**68%**	.07
33	40%	16%	.62	**68%**	.07

Table 25 (continued)

38	100%	100%	.01	100%	.01
42	40%	**24%**	.46	**32%**	.40
46	**60%**	**32%**	.40	60%	.12

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

8.1.3.2 IFOA group

IFOA group: Ma-1

As Tables 26 and 27 present, half of the participants in IFOA demonstrated a baseline trend during Ma-2 productions. During Ma-1 productions, none of the participants demonstrated a significant error reduction. Although the Tau-U scores of some of the participants (2, 24, 41, 44) indicated an error reduction, these values were not significant.

IFOA group: Ma-2

As Tables 28 and 29 present, 10 of the 14 participants in the IFOA group demonstrated a baseline trend during Ma-2 productions. Except for participant 14, no other participant exhibited any significant improvement. The Tau-U effect sizes of participant 14 showed a significant improvement during both the acquisition and the retention phase for Ma-2 productions.

IFOA group: Ma-3

As Tables 30 and 31 present, eight of the 14 participants in the IFOA group demonstrated a baseline trend during Ma-3 productions. The Tau-U effect sizes indicated that during Ma-3 productions, error reduction was significant for only a few participants (2, 14, 24) during either the acquisition or the retention phase. It should be noted that, only one participant (50) demonstrated this significant reduction during both the acquisition and the retention phases.

IFOA group: Ma-4

As Tables 32 and 33 present, only four participants in the IFOA group demonstrated a baseline trend during Ma-4 productions. Although the effect sizes of four participants (7, 26, 27, 50, 41) showed a reduction in their RMSE during Ma-4, the reduction failed to reach significance except for participant 41. The effect sizes for participant 41 demonstrated a significant decrease in RMSE during both the acquisition and retention phases. No other significant error reduction emerged.

Table 26 Results of the Tau-U analyses effect sizes for the IFOA group Ma-1 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
2	**10%**	8%	.81	**20%**	.62
5	20%	100%	.002	100%	.01
11	20%	100%	.002	100%	.01
39	20%	4%	.9	36%	.34

Table 26 (continued)

55	33%	7%	.86	20%	.65
40	10%	70%	.03	60%	.12
44	**20%**	6%	.85	**8%**	.85
50	33%	55%	.11	40%	.32

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 27 Results of the Tau-U analyses effect sizes for the IFOA group Ma-1 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
7	**40%**	8%	.80	24%	.53
14	60%	86%	.01	76%	.04
24	**50%**	**40%**	.90	0	.83
26	**40%**	96%	.003	100%	.60
27	80%	84%	.01	60%	.12
34	80%	48%	.14	28%	.46
41	66%	**5%**	.88	**5%**	.90

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 28 Results of the Tau-U analyses effect sizes for the IFOA group Ma-2 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
7	**30%**	**24%**	.46	36%	.34
14	10%	**92%**	.004	**84%**	.02
39	0	62%	.07	55%	.17
50	0	**8%**	.8	**64%**	.09

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 29 Results of the Tau-U analyses effect sizes for the IFOA group Ma-2 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
2	60%	**2%**	.95	**12%**	.75
5	**40%**	20%	.54	48%	.21
11	40%	88%	.01	36%	.34
24	40%	30%	.35	**4%**	.91
26	60%	44%	.18	36%	.34
27	**40%**	80%	.01	16%	.67
34	40%	54%	.09	**100%**	.01

Table 29 (continued)

40	100%	14%	.66	24%	.55
41	**80%**	84%	.01	56%	.14
44	40%	**2%**	.95	20%	.60

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 30 Results of the Tau-U analyses effect sizes for the IFOA group Ma-3 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
5	0	**4%**	.9	**28%**	.46
11	20%	76%	.02	100%	.01
24	0	**50%**	.12	**80%**	.04
26	**33%**	35%	.32	40%	.32
39	20%	52%	.11	84%	.02
40	**20%**	**32%**	.32	**44%**	.25

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 31 Results of the Tau-U analyses effect sizes for the IFOA group Ma-3 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
2	60%	**84%**	.01	**56%**	.75
7	**40%**	14%	.66	36%	.34
14	**100%**	**75%**	.03	**50%**	.22
27	**100%**	88%	.01	100%	.01
34	**60%**	**20%**	.54	80%	.05
41	**40%**	**16%**	.62	**12%**	.75
44	80%	**30%**	.35	**44%**	.25
50	**40%**	**66%**	.04	**84%**	.02

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 32 Results of the Tau-U analyses effect sizes for the IFOA group Ma-4 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
2	**20%**	40%	.22	100%	.01
5	0	6%	.85	4%	.92
7	0	**64%**	.05	**28%**	.46
26	0	14%	.66	**12%**	.75
27	**20%**	22%	.50	**28%**	.46
34	0	22%	.50	20%	.60
39	0	28%	.39	20%	.60

Table 32 (continued)

41	20%	**80%**	.01	**92%**	.01
44	20%	56%	.08	32%	.40
50	0	**6%**	.85	4%	.9

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 33 Results of the Tau-U analyses effect sizes for the IFOA group Ma-4 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
11	**80%**	54%	.09	52%	.17
14	60%	84%	.01	52%	.17
24	**70%**	16%	.62	92%	.02
40	**60%**	0	1	0	.75

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).

The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

8.1.3.3 Control group

Control group: Ma-1

As Tables 34 and 35 present, six participants in the C group demonstrated a baseline trend during Ma-1 productions. Although some of the participants (3, 30, 31, 35, 48) showed a reduction in their error, only participants 30 and 35 demonstrated a significant error reduction in Ma-1 productions during the acquisition and the retention phase, respectively.

Control group: Ma-2

As Tables 36 and 37 present, eight participants in the C group demonstrated a baseline trend during Ma-2 productions. As shown in Table 36 and 37, six participants (4, 12, 18, 31, 35, 37) demonstrated error reduction during both the acquisition and retention phases; only participant 4 in both phases of Ma-2 productions and participant 12 during retention phase of Ma-2 reached significant error reduction.

Control group: Ma-3

As Tables 38 and 39 present, 11 participants in the C group demonstrated a baseline trend during Ma-3 productions. Although seven participants (4, 16, 18, 19, 30, 35, 37) demonstrated error reduction Ma-3 productions, only participant 19 during the retention phase and participant 35 during the acquisition phase reached significance error reduction in Ma-2 productions.

Control group: Ma-4

Five of the 14 participants demonstrated a baseline trend (Table 41). The Tau-U effect sizes demonstrated a significant error reduction for only three participants (4, 19, 23) during the acquisition phase and only for participant 16 during retention phase of Ma-4 production (Tables 40 and 41).

Table 34 Results of the Tau-U analyses effect sizes for the control group Ma-1 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	p	Retention vs. Baseline Tau-U	p
12	20%	72%	.02	60%	.12
21	0	90%	.01	100%	.01
23	**10%**	20%	.54	100%	.01
28	**20%**	30%	.35	64%	.09
30	33%	**72%**	.04	**40%**	.32
31	20%	**12%**	.50	**12%**	.75
35	**20%**	**38%**	.24	**80%**	.03
48	33%	**7%**	.80	**35%**	.39

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.

Table 35 Results of the Tau-U analyses effect sizes for the control group Ma-1 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	p	Retention vs. Baseline Tau-U	p
3	67%	**20%**	.57	**30%**	.46
4	40%	90%	.01	44%	.25
16	**80%**	100%	.004	100%	.001
18	70%	82%	.01	72%	.06
19	60%	80%	.33	70%	.11
37	60%	56%	.08	56%	.14

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 36 Results of the Tau-U analyses effect sizes for the control group Ma-2 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	p	Retention vs. Baseline Tau-U	p
3	**10%**	70%	.03	100%	.01
4	**20%**	**100%**	.002	**100%**	.01
12	**30%**	**60%**	.06	**100%**	.01
16	0	95%	.01	80%	.05
23	**20%**	68%	.03	52%	.17
31	0	**54%**	.09	**24%**	.53

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 37 Results of the Tau-U analyses effect sizes for the control group Ma-2 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
18	40%	**20%**	.54	**52%**	.17
19	**60%**	72%	.03	40%	.29
21	60%	54%	.10	48%	.21
28	40%	84%	.01	68%	.07
30	40%	38%	.24	100%	.002
35	70%	**48%**	.14	**16%**	.67
37	**40%**	**56%**	.08	**60%**	.12
48	40%	72%	.03	68%	.07

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 38 Results of the Tau-U analyses effect sizes for the control group Ma-3 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
12	**20%**	60%	.66	60%	.12
19	**20%**	**20%**	.54	**80%**	.03
28	**10%**	24%	.46	50%	.46

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 39 Results of the Tau-U analyses effect sizes for the control group Ma-3 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
3	**80%**	20%	.54	**20%**	.60
4	**100%**	**65%**	.16	**50%**	.33
16	**50%**	**6%**	.85	**8%**	.83
18	50%	4%	.90	**75%**	.06
21	**50%**	54%	.09	88%	.02
23	**50%**	50%	.12	80%	.02
30	**40%**	**44%**	.17	**36%**	.34
31	40%	14%	.67	84%	.03
35	**80%**	**80%**	.01	0	1
37	**60%**	**63%**	.05	**60%**	.11
48	**100%**	36%	.27	52%	.17

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 40 Results of the Tau-U analyses effect sizes for the control group Ma-4 production, for participants with no baseline trend (Tau-U < 40%)

Participants	Baseline Tau-U	Acquisition vs. baseline Tau-U	*p*	Retention vs. Baseline Tau-U	*p*
3	0	20%	.54	4%	.92
16	20%	0	1	**80%**	.04
19	**20%**	**64%**	.05	**44%**	.25
21	10%	52%	.11	28%	.46
23	0	**90%**	.01	35%	.39
28	0	60%	.07	4%	.92

Table 40 (continued)

31	0	12%	.71	**36%**	.34
35	**20%**	74%	.02	76%	.04
37	0	44%	.18	52%	.17

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

Table 41 Results of the Tau-U analyses effect sizes for the control group Ma-4 production, corrected for baseline trend (Tau-U > 40%)

Participants	Baseline	Acquisition vs. baseline		Retention vs. Baseline	
	Tau-U	Tau-U	*p*	Tau-U	*p*
4	**40%**	**64%**	.05	**60%**	.12
12	**40%**	12%	.71	12%	.75
18	**60%**	68%	.03	100%	.002
30	60%	30%	.35	**36%**	.35
48	**80%**	36%	.27	28%	.46

Note. The boldface in the second column denotes a trend in the baseline in the direction of error reduction (improvement).
The boldface in the third and fifth columns denotes improvement; The Tau-U effect size indicates the percent of non-overlap data points relative to baseline data.
The Tau-U effect sizes in the third and fifth columns are corrected for baseline trends.

The results of the single subject analysis suggest that the focus of attention instructions did not show an immediate effect on performance during the acquisition phase for any of the participants. Nonetheless, it was noted that the magnitude of decrease in error was very small in most participants as indicated by the Tau-U scores. However, it was noted from the visual inspection that the fluctuation of the data point (variability) seemed to decrease during the acquisition phase relative to baseline performance.

To explore whether the variability decreased during the acquisition phase, the data series for each participant was further examined. For this analysis, a measure of dispersion was examined by calculating the Coefficient of Variation *(CoV)* of the RMSE. The Coefficient of Variation has been commonly used to compare the dispersion of one set of data points to the dispersion of another set of data points. The Coefficient of Variation for a data set is calculated using the following formula:

$$CoV = \frac{SD}{\overline{X}}$$

Where, *SD* is the standard deviation and $\bar{x}$ is the mean. For each participant's data series, two separate *CoV* were calculated, one for the baseline data points and the other for the acquisition data series. The *CoV* was calculated separately for each tone.

Tables 42 to 45 present the Coefficient of Variation *(CoV)* for each participant during baseline and the acquisition phase for the production of Ma-1, Ma-2, Ma-3, and Ma-4. Comparing the acquisition phase *CoV* value to the baseline *CoV* value for each participant yielded the following results. In Ma-1 production, a reduction in the variability during the acquisition phase relative to baseline variability was demonstrated as follows: 11 of the 14 participants in the EFOA group; 11 of the 14 participants in the IFOA group; 10 of the 14 participants in the C group (Table 42). The *CoV* values in Table 43 indicate that for Ma-2 production, the *CoV* decreased as follows: eight of the 14 participants in the EFOA group; seven of the 14 participants in the IFOA group; and 10 of the 14 participants in the C group. According to the *CoV* values of Ma-3 production, eight of the 14 participants in the EFOA group; four of the 14 participants in the IFOA group; and five of the 14 participants in the C group demonstrated a reduction in their *CoV* during the training period relative to their baseline *CoV* (table 44). As Table 45 shows, the reduction in *CoV* for Ma-4 production was as follows: 11 of the 14 participants in the EFOA group; and 12 of the 14 participants in the IFOA group; 12 of the 14 participants in the C group.

Table 42 Coefficient of Variation (CoV) values for each participant in Ma-1 production, during the baseline and the acquisition phase

Group	Participant	Ma1- CoV	
		BL	Acquisition
EFOA	**6**	**.14**	**.08**
	10	**.16**	**.10**
	13	**.09**	**.02**
	20	.11	.12
	22	**.04**	**.02**
	29	**.07**	**.02**
	32	**.63**	**.05**

Table 42 (continued)

	33	.10	.17
	38	**.09**	**.08**
	42	**.17**	**.05**
	43	.24	.53
	46	**.05**	**.02**
	47	**.13**	**.08**
	55	**.45**	**.25**
IFOA	**2**	**.09**	**.07**
	5	**.28**	**.08**
	7	**.04**	**.01**
	11	**.08**	**.05**
	14	.09	.04
	24	.14	.15
	26	**.09**	**.06**
	27	**.11**	**.04**
	34	**.27**	**.09**
	39	.05	.06
	40	**.12**	**.05**
	41	**.64**	**.20**
	44	**.14**	**.09**
	50	**.04**	**.04**
C	**3**	**.49**	**.04**
	4	.06	.07
	12	**.26**	**.13**
	16	.09	.10
	18	**.11**	**.03**
	19	.04	.04
	21	**.47**	**.13**
	23	**.05**	**.04**
	28	.06	.06
	30	**.27**	**.08**
	31	**.22**	**.14**
	35	**.17**	**.16**
	37	**.13**	**.02**
	48	**.32**	**.01**

Note. **Bolded** numbers denotes that *CoV* value decreased during acquisition, compared to baseline (BL).

Table 43 Coefficient of Variation (CoV) values for each participant in Ma-2 production, during the baseline and the acquisition phase

Group	Participant	Ma1- CoV	
		BL	Acquisition
EFOA	6	.05	.07
	10	**.17**	**.08**
	13	**.08**	**.03**
	20	**.17**	**.07**
	22	.07	.07
	29	**.08**	**.04**
	32	**.14**	**.07**
	33	.07	.11
	38	**.06**	**.05**
	42	**.15**	**.08**
	43	.10	.12
	46	.03	.10
	47	.14	.14
	55	**.30**	**.10**
IFOA	2	.08	.16
	5	.05	.09
	7	**.19**	**.12**
	11	**.09**	**.03**
	14	**.06**	**.03**
	24	**.33**	**.13**
	26	**.10**	**.05**
	27	.05	.07
	34	.03	.10
	39	**.10**	**.05**
	40	.09	.09
	41	.11	.18
	44	**.34**	**.08**
	50	.04	.05

Table 43 (continued)

C	3	.13	.18
	4	**.18**	**.09**
	12	**.26**	**.18**
	16	.06	.06
	18	.03	.06
	19	**.19**	**.04**
	21	**.51**	**.19**
	23	**.07**	**.03**
	28	**.12**	**.06**
	30	**.18**	**.04**
	31	**.14**	**.12**
	35	.10	.10
	37	**.09**	**.05**
	48	**.23**	**.06**

Note. **Bolded** numbers denotes that *CoV* value decreased during acquisition, compared to baseline (BL).

Table 44 Coefficient of Variation (CoV) values for each participant in Ma-3 production, during the baseline and the acquisition phase

Group	Participant	Ma1- CoV	
		BL	Acquisition
EFOA	6	.08	.17
	10	**.11**	**.10**
	13	**.18**	**.10**
	20	.11	.19
	22	**.09**	**.06**
	29	.03	.08
	32	.07	.07
	33	**.13**	**.09**
	38	.09	.19
	42	**.17**	**.08**
	43	**.12**	**.05**
	46	**.12**	**.03**
	47	**.53**	**.12**

Table 44 (continued)

	55	.08	.24
IFOA	2	.03	.12
	5	.12	.13
	7	**.24**	**.16**
	11	.05	.14
	14	.06	.08
	24	.18	.22
	26	.06	.09
	27	**.11**	**.09**
	34	.07	.13
	39	**.13**	**.07**
	40	**.11**	**.02**
	41	.29	.37
	44	.11	.11
	50	.04	.05
C	**3**	**.58**	**.41**
	4	.17	.23
	12	**.46**	**.23**
	16	.05	.14
	18	.02	.05
	19	**.08**	**.05**
	21	.11	.15
	23	.04	.12
	28	**.07**	**.06**
	30	**.21**	**.11**
	31	.10	.13
	35	.04	.16
	37	.12	.18
	48	.10	.17

Note. **Bolded** numbers denotes that *CoV* value decreased during acquisition, compared to baseline (BL).

Table 45 Coefficient of Variation (CoV) values for each participant in Ma-4 production, during the baseline and the acquisition phase

Group	Participant	Ma1- CoV	
		BL	Acquisition
EFOA	**6**	**.17**	**.10**
	10	**.20**	**.09**
	13	**.23**	**.08**
	20	**.41**	**.17**
	22	**.11**	**.04**
	29	**.18**	**.08**
	32	**.14**	**.08**
	33	**.12**	**.11**
	38	.04	.08
	42	.06	.06
	43	.14	.18
	46	**.49**	**.12**
	47	**.29**	**.09**
	55	**.22**	**.18**
IFOA	2	.06	.11
	5	**.17**	**.09**
	7	**.08**	**.05**
	11	**.15**	**.06**
	14	**.12**	**.08**
	24	.17	.27
	26	**.17**	**.07**
	27	**.09**	**.05**
	34	**.38**	**.16**
	39	**.22**	**.10**
	40	**.29**	**.12**
	41	**.35**	**.19**
	44	**.15**	**.10**
	50	**.08**	**.05**
C	**3**	**.41**	**.13**
	4	**.19**	**.13**
	12	**.57**	**.15**
	16	**.15**	**.08**
	18	**.16**	**.07**

Table 45 (continued)

19	**.11**	**.06**
21	**.35**	**.13**
23	.09	.13
28	**.14**	**.08**
30	**.09**	**.05**
31	.10	.13
35	**.10**	**.10**
37	**.09**	**.05**
48	**.14**	**.08**

Note. **Bolded** numbers denotes that *CoV* value decreased during acquisition, compared to baseline (BL).

To assess whether there was a significant difference of the mean *CoV* between the baseline phase and the acquisition phase for each group, the mean *CoV* values for each group was compared. Because the *CoV* was not normally distributed, the Wilcoxon signed-rank test was used. As shown in Table 46, compared to their baseline variability, participants in the EFOA group showed a significant decrease in their variability during their production of Ma-4 in the acquisition phase. No other significant decrease in variability emerged across Ma conditions for the EFOA group. Participants in the IFOA group showed a significant decrease of *CoV* during their production of both Ma-1 and Ma-4. Participants in the C group demonstrate a significant reduction of their variability during the production of three of the four practiced words: Ma-1, Ma-2, and Ma-4.

Table 46 The Wilcoxon signed-rank test results for the CoV for each group—for the four practiced tones, during the acquisition phase

Group	Word	*Mean CoV*		*Z*	*p*
		BL	Acquisition		
EFOA	Ma-1	.18	.11	-1.79	.073
	Ma-2	.11	.08	-1.69	.091
	Ma-3	.14	.11	-0.21	.834
	Ma-4	.20	.10	-2.76	.006
IFOA	Ma-1	.16	.07	-2.98	.003
	Ma-2	.12	.09	-0.66	.506
	Ma-3	.12	.13	-0.83	.409
	Ma-4	.18	.11	-2.36	.018
C	Ma-1	.20	.07	-2.68	.007
	Ma-2	.16	.09	-2.51	.012
	Ma-3	.15	.16	-0.66	.509
	Ma-4	.20	.10	-3.05	.002

To summarize, the results of the single subject analyses suggest that some participants showed minute improvement in performance as indicated by the visual analysis and the Tau-U analysis. However, these changes were minimal and did not occur immediately. Nonetheless, the variability of some of the participants' performance decreased in the acquisition phase as compared to their baseline performance (as shown in tables 42-45). When compared to the baseline *CoV*, the mean *CoV* of the participants' productions was significantly lower for three of the monosyllabic words in the control group (Ma-1, Ma-2, and Ma-4); the variability of the participants' productions was significantly lower for two of the monosyllabic words (Ma-1 and Ma-4) in the IFOA group and only significant for Ma-4 production in the EFOA group.

8.2 THE PERCEPTUAL ANALYSIS

Three female Mandarin Chinese native speakers (mean age = 32 years old), judged the tone productions of the English native speakers to assess differences among the groups. The three judges were graduate students at the University of Pittsburgh with no reported speech or hearing impairments. The percentage of correct productions in the baseline, acquisition, retention, and transfer phases is the dependent variable for the perceptual analyses. The next section presents these results. Moreover, the last question regarding whether all four tones were acquired equally well was answered by analyzing the perceptual data.

For the perceptual analysis, the three Mandarin Chinese native speakers rated the monosyllabic words produced by the English native speakers. In a forced choice rating task, the listeners were instructed to press one of five buttons on a computer keyboard. Each button corresponding to one of the four tones and a fifth button indicated that the produced word did not correspond to any of the four tones.

The speech material consisted of 16,800 tokens produced by all 42 participants throughout the experiment: baseline (60) + training (200) + probes (60) + immediate retention (20) + delayed retention (20) + transfer test (40). The stimuli were pre-randomized within as well as between all of the phases: baseline, acquisition, retention, transfer; for each participant. The perceptual analysis was performed off-line after the completion of data collection from all the participants. Each listener judged all the stimuli individually. The listeners scheduled the sessions at their convenience. Before every session, the raters were required to listen to the Mandarin Chinese speaker model in order to assure that the listeners would rate the tones with consideration to the auditory model.

The percentage of correct trials per block was calculated by dividing the number of correctly perceived monosyllabic words by the total number of monosyllabic words. The acquisition phase consisted of ten blocks with five productions of each tone per block. The data from each block revealed one score-percentage of correctly perceived tones—for that block. Each rater judged all stimuli. The percentage of correct productions was calculated by dividing the number of correctly perceived monosyllabic words/tones by the total number of monosyllabic word in that tone (five monosyllabic words for each tone in every block). The percentage of correct productions was either 0, 20, 40, 60, or 100%. After the completion of the perceptual rating—by the native Mandarin Chinese speakers, agreement among two or three raters determined the response accuracy. This analysis was performed separately for each tone.

8.2.1 Intra-rater reliability

To assess raters' reliability, each rater responded twice to 5% of randomly selected stimuli (n = 840; 20 tokens from each participant). The reliability was established by computing the percentage of agreement. The raw point to point agreement was 87%, 95% and 88% for each rater.

8.2.2 Inter-rater reliability

Agreement between two of the three raters determined the accuracy of the response. The rating of the three raters was in agreement for 81% of the items.

The screening of the perceptual data revealed an outlier in the control group; the data of participant 31 deviated markedly, 3.5 SD below the group mean, and was excluded from all the perceptual analyses. Before addressing the research questions, the study first compared the three groups on the percentage of correct productions during the baseline phase. This was performed to determine whether all groups were equivalent on the dependent variable before beginning the acquisition phase. The means and standard deviations of the percentage of correct productions are presented in Table 47. Because the percentage of correct productions was not normally distributed, the non-parametric Kruskal-Wallis test was performed. As Table 48 shows, no significant difference on the percentage of correctly perceived tones emerged among the three groups (EFOA, IFOA, C) during baseline. This finding suggests that the groups' performance was similar before the acquisition phase.

Table 47 The means and standard deviations of the percentage of correctly perceived productions in each syllable production among the three groups during baseline

		% of correctly perceived productions	
Syllables	Groups	*Mean*	*SD*
Ma-1	EFOA	88	24
	IFOA	97	7
	C	97	7
Ma-2	EFOA	91	12
	IFOA	84	17
	C	85	14
Ma-3	EFOA	58	28
	IFOA	60	32
	C	61	29
Ma-4	EFOA	87	32
	IFOA	91	23
	C	100	_
Me-1	EFOA	87	32
	IFOA	100	_
	C	97	7
Me-2	EFOA	78	23
	IFOA	80	26
	C	88	23
Me-3	EFOA	66	30
	IFOA	60	44
	C	74	36
Me-4	EFOA	93	22
	IFOA	100	_
	C	98	5
Na-1	EFOA	81	38
	IFOA	98	5
	C	100	_

Table 47 (continued)

Na-2	EFOA	84	18
	IFOA	90	20
	C	91	13
Na-3	EFOA	66	28
	IFOA	78	35
	C	77	20
Na-4	EFOA	88	28
	IFOA	97	7
	C	100	_

Table 48 The results of the Kruskal-Wallis tests performed for each monosyllabic-word during baseline phase

Words	χ^2	df	P
Ma-1	0.608	2	.738
Ma-2	1.869	2	.393
Ma-3	0.304	2	.859
Ma-4	2.016	2	.365
Me-1	2.237	2	.327
Me-2	1.999	2	.368
Me-3	0.991	2	.609
Me-4	2.104	2	.349
Na-1	3.820	2	.148
Na-2	1.381	2	.501
Na-3	3.464	2	.177
Na-4	4.501	2	.105

In order to assess whether there was a significant difference on the correctly perceived productions among the tones, the Friedman test was performed on the percentage of correctly perceived productions for each syllable. There was a significant difference on the correctly perceived productions for each syllable among the four tones for all three syllables: Ma, Me, and Na (χ^2 = 45.90, *df* = 3; *p* < .001, χ^2 = 31.29, *df* = 3; *p* < .001, χ^2 = 31.19, *df* = 3; *p* < .001, respectively). In order to find the pattern of difference on the correctly perceived productions for each syllable among the tones for each syllable averaged across groups, the Wilcoxon matched pair signed rank test was performed.

The Ma syllable, Tone 1 (*M* = 92%, *SD* = 19%) was produced significantly more accurately than Tones 2 (*M* = 85%, *SD* = 18%), and 3 (*M* = 59%, *SD* = 30%). These data further show that Tone 2 scored significantly higher than Tone 3. Tone 4 (*M* = 91%, *SD* = 26) scored significantly higher than Tone 3. However, the percentage of correctly perceived productions between Tones 4 and 1 and between Tones 4 and 2 for the Ma syllable were not significantly different.

The Me syllable, Tone 1 (*M* = 93%, S*D* = 21%) was produced significantly more accurately than Tones 2 (M = 81%, SD = 24%), and 3 (M = 66%, SD = 36%). Moreover, Tone 4 scored significantly higher than Tones 2 and 3. No significant difference emerged between Tones 1 and 4 (*M* = 96%, *SD* = 14%) or between Tones 2 and 3.

The Na syllable, Tone 1 (*M* = 91%, *SD* = 27%) was produced significantly more accurately than Tones 3 (*M* = 73%, *SD* = 29%). However, difference between Tones 1 and 2 (*M* = 87%, *SD* = 17%) and between Tones 1 and 4 (*M* = 94%, *SD* = 19%) was not significant. These data further show that Tone 2 was produced significantly more accurately than Tones 3. Moreover, Tone 4 was produced significantly more accurately than Tones 3.

As the above data illustrate, the four tones were not produced with the same accuracy as was reflected by the percentage of correctly perceived productions. Moreover, the percentage of accurately perceived productions demonstrated a hierarchy for the four tones during the production of the three syllables: 1) for the Ma syllable, Tone 1 was produced with the highest accuracy, followed by Tones 4 and 2, while Tone 3 scored lowest; 2) For the Me syllable, Tone 4 was produced with the highest accuracy, followed by Tones 1 and 2, while Tone 3 scored lowest; 3) For the Na syllable, Tone 4 was produced with the highest accuracy, followed by Tones 1 and 2, while Tone 3 scored lowest. Tone 3 was perceived as the least accurately produced tone across the three syllable forms.

8.2.3 The acquisition phase

Research Question:

> Are there significant differences in the slope for the percentage of the correctly perceived words across the acquisition phase of the experiment among the three participant groups: EFOA, IFO, Control?

The percentage of the correctly perceived word/syllable data lacked variability in some of the tones. Table 31 shows the percentage of the data points that were judged to be a 100% correct during the acquisition phase. As Table 49 shows, 97% and 94% of Ma-1 and Ma-4 productions respectively were judged to be 100% accurate at ceiling.

The statistical analysis excluded those occurrences—Ma-1 and Ma-4—when the data were at ceiling. Ma-2 and Ma-3 scores were considerably more variable throughout the acquisition phase. Forty-seven percent of the Ma-2 scores and 46 % of the Ma-3 scores had the value of 100%. Therefore, for the acquisition phase, the slopes of only Ma-2 and Ma-3 were analyzed.

Table 49 The percentage of data points that had a value of 100% during the acquisition phase

Word/Syllable	Percentage of the data scored 100%
Ma-1	97%
Ma-2	47%
Ma-3	46%
Ma-4	94%

To examine the participants' performance during the acquisition phase, the slopes of the percentage of correctly perceived productions were calculated for each participant from the ten acquisition blocks. A 3 X 2 mixed ANOVA was performed on the slopes as a function of the groups (EFOA, IFOA, C) and two syllables (Ma-2 and Ma-3). The group was the between-subject factor and the tones were the within-subject factor. The assumption of normality was met for the slopes of the percentage of correctly perceived productions.

The means and the standard deviations of the slopes are shown in Table 50. For these perceptual data, a positive slope indicates an increase in the percentage of correctly perceived productions (improvement), while a negative slope indicates a decrease in the percentage of correctly perceived productions during the acquisition phase. Neither the two the mains effects (of group, $F(1, 39) = 1.57$, $p = 0.22$, partial $ŋ2 = 0.07$; of slopes $F(1, 38) = 2.91$, $p = .096$, partial $ŋ2 = 0.07$, nor the interaction, $F(2, 38) = 0.540$, $p = .587$, partial $ŋ2 = 0.028$, was significant.

Table 50 The means and SDs Ma-2 and Ma-3 slopes during the acquisition phase in the three groups

Slopes	Groups	*Mean*	*SD*
Slope-Ma2	EFOA	-1.19	2.76
	IFOA	-0.91	3.36
	C	0.11	2.53
	$\bar{x}$	-0.68	2.90
Slope-Ma3	EFOA	-0.22	3.08
	IFOA	1.49	3.83
	C	0.66	1.85
	$\bar{x}$	0.64	3.07

8.2.4 Learning tests

To assess learning, the following two questions were proposed:

Research Questions:

Are there significant differences in the percentage of correctly perceived words during the retention phases of the experiment among the three experimental groups: EFOA, IFO, and Control?

Are there significant differences in the percentage of correctly perceived words during the transfer phase of the experiment among the three participant groups: EFOA, IFO, and Control?

8.2.4.1 Retention tests

The data for some tones lacked variability (immediate retention test: Ma-1, Ma-4; delayed retention test: Ma-1) and were excluded from this analysis (Table 51). Means and standard deviations are presented in Table 52.The percentage of correct productions did not meet the assumption of normality. Therefore, the non-parametric Kruskal-Wallis test was performed on the percentage of correctly perceived productions for two monosyllabic words at the immediate retention test and for three monosyllabic words at the delayed retention test. One separate test was performed for each syllable. The percentage of the correctly perceived productions was not significantly different among the groups (Table 53).

Table 51 The percentage of data points with 100% score during immediate and delayed retention tests.

Retention test	Syllable-Tone	Percentage of the data scored 100%
IRT	Ma-1	100%
	Ma-2	36%
	Ma-3	36%
	Ma-4	95%
DRT	Ma-1	98%
	Ma-2	43%
	Ma-3	24%
	Ma-4	88%

Table 52 Means and SDs for the percentage of correctly perceived words during immediate and delayed retention tests

		IRT-%C		DRT-%C	
Word	Groups	*M*	*SD*	*M*	*SD*
Ma-1	EFOA	100	00	100	00
	IFOA	100	00	100	00
	C	100	00	100	00
Ma-2	EFOA	77	23	76	24
	IFOA	68	29	83	20
	C	90	10	85	20
Ma-3	EFOA	71	27	64	28
	IFOA	76	32	60	31
	C	78	21	74	25
Ma-4	EFOA	100	00	98	5
	IFOA	93	27	94	16
	C	100	00	98	5

Table 53 Kruskal-Wallis tests results on the percentage of correctly perceived words among the groups during immediate and delayed retention tests

Retention test	word/Syllable	χ^2	*df*	P
IRT	Ma-2	5.189	2	.075
	Ma-3	0.515	2	.773
DRT	Ma-2	1.359	2	.507
	Ma-3	1.478	2	.478
	Ma-4	0.562	2	.755

In order to assess the time factor (immediate vs. delayed retention test), the Wilcoxom matched-sign paired test was performed for each group. There was no significant difference between the immediate and the delayed retention tests in any of the groups (Table 54).

Table 54 Results of the Wilcoxon matched-sign paired test on the immediate and delayed retention tests percentages of correctly perceived words—for each group.

Word/Syllable	Groups	*z*	*p*
Ma-2	EFOA	-0.17	.863
	IFOA	-1.65	.098
	C	-1.24	.214
Ma-3	EFOA	-1.67	.096
	IFOA	-1.55	.121
	C	-1.00	.317
Ma-4	EFOA	-1.00	.317
	IFOA	-1.00	.655
	C	-1.00	.317

8.2.4.2 The transfer test

The data for some tones were at ceiling, lacking variability (Me-1, Me-4, Na-1, Na-4) and they were excluded from additional analysis (Table 55). Therefore, the percentage of correctly perceived productions was analyzed for four monosyllabic words at the transfer phase. Means and standard deviations are presented in Table 56. The percentage of correct productions was not normally distributed. Therefore, the Kruskal-Wallis test was performed. The test results yielded no significant difference on the percentage of correct productions among the groups. This was true for all of the analyzed monosyllabic words (Table 57).

Table 55 The percentage of data points demonstrating a ceiling effect with 100% score during the transfer test

Syllable-Tone	Percentage of the data scored 100%
Me-1	92%
Me-2	52%
Me-3	71%
Me-4	98%
Na-1	93%
Na-2	50%
Na-3	64%
Na-4	93%

Table 56 Means and SDs for the percentage of correctly perceived productions during the transfer test

		Percentage of correctly perceived productions %	
Word	Groups	*Mean*	*SD*
Me-1	EFOA	98	5
	IFOA	100	00
	C	98	5
	x̄	99	4
Me-2	EFOA	83	17
	IFOA	86	23
	C	81	28
	x̄	83	22

Table 56 (continued)

Me-3	EFOA	80	31
	IFOA	86	30
	C	96	7
	x̄	87	26
Me-4	EFOA	100	00
	IFOA	100	00
	C	100	00
	x̄	100	00
Na-1	EFOA	98	5
	IFOA	98	5
	C	100	00
	x̄	99	4
Na-2	EFOA	74	28
	IFOA	86	20
	C	89	19
	x̄	83	23
Na-3	EFOA	76	30
	IFOA	89	20
	C	91	17
	x̄	85	24
Na-4	EFOA	98	5
	IFOA	98	5
	C	100	00
	x̄	99	4

Table 57 Kruskal-Wallis analysis results for the percentage of correct production during the transfer test

Syllable-Tone	χ^2	*df*	*P*
Me-2	0.535	2	.765
Me-3	1.991	2	.370
Na-2	3.759	2	.153
Na-3	1.703	2	.427

To test whether the percentage of correctly perceived productions was significantly different between Tones 2 and 3, the Wilcoxon rank-sign test was performed for each group. There was no significant difference between the Tones 2 and 3 for both syllables (Table 58).

Table 58 Results of the Wilcoxon rank-sign test, comparing Tones 2 and 3 for Me and Na syllables for each group during the transfer test

Compared words	Group	*Z*	*P*
Me-2—Me-3	EFOA	-0.345	.730
	IFOA	1.000	--
	C	-1.890	.059
Na-2—Na-3	EFOA	-0.199	.842
	IFOA	-0.240	.810
	C	-0.322	.748

8.2.5 Probes

To assess whether the participants generalized the practice tones to unpracticed but similar monosyllabic words (transfer words), probe testing of the unpracticed monosyllabic words was administered throughout the acquisition phase. First, the baseline phase tested the probes by presenting them in random order with the practice monosyllabic words. Then the acquisition phase tested the probes after block-1, after block-4, and after block-10. The participants did not receive feedback on their performance on the probe productions. Moreover, the probe testing was introduced after the practice block and was not preceded by any focus instructions. To decrease any possible practice advantage on the transfer words productions, the probe testing involved two of the four tones: one easy tone and one complex tone. In each group, half of the participants were probed on half of the transfer words (words produced with Tones 1 and 2), and the other half of the participants were probed on the other half of the transfer words (words produced with Tones 3 and 4).

In order to inspect the probes productions, the RMSE and the percentage of correctly produced values were plotted for each individual (See Appendix F). For the baseline phase and the transfer test, all trials were plotted in order to visually examine any patterns; for the acquisition phase, each point on the figure represents the mean of five trials,

As the figures in Appendix F demonstrates, the probe data do not provide any clear evidence of a performance change or a possible transfer of the practiced tones. The data points for the RMSE do not demonstrate any substantial difference across the experiment phases. The RMSE of the four tones paralleled the accuracy hierarchy of the four tones; the RMSE was higher for Tones 3 and 4 than for Tones 1 and 2. In almost all the participants, the percentage of correctly produced probed data showed a ceiling effect that prevented any noticeable differences. Therefore, the probe results in the current study did not yield any information concerning any possible transfer of the practiced tones to similar but unpracticed words.

8.3 SECONDARY RESEARCH QUESTION

A secondary research question was proposed to determine whether learners would acquire each of the tones at the same level of accuracy or whether they would show differential difficulty with some tones, as demonstrated in the second language learning literature.

Research Question:

Are there significant differences in the percentage of correct productions among the four tones for the Ma syllable, for each group, during the acquisition phase?

In order to evaluate whether the four tones were produced with the same level of accuracy during acquisition phase, an average score was calculated for each tone. For each participant, to obtain the percentage of correctly perceived tone, the number of accurately perceived tones during acquisition was divided by the total number of trials for that tone during acquisition (40 trials). The percentage of correctly perceived tone productions was calculated for each participant during the acquisition phase (Table 59) and an average score was calculated for each group (see Figure 7).

These data did not meet the normality assumption. Therefore, to test whether the four tones differed within each groups, the non-parametric Friedman test was performed on the percentage of correctly perceived productions for each group.

Table 59 The percentages of correctly perceived words (Ma syllables) during the acquisition phase for all participants and the three groups means and standard deviations for the four tones

Group	Subjects	T-1	T-2	T-3	T-4
EFOA	10	100	48	94	100
	20	100	80	96	100
	32	100	70	88	100
	33	100	90	84	100
	38	100	82	34	98
	43	100	62	90	98
	47	100	96	90	100
	6	100	66	64	100
	13	100	58	92	100
	22	100	66	46	100
	29	100	80	70	100
	46	100	96	70	100
	42	100	92	72	100
	55	98	100	18	100
	$\bar{x}$	100	78	72	100
	SD	0.5	16.2	24.2	0.7
IFOA	2	100	44	94	100
	5	100	94	34	22
	11	100	72	82	100
	27	100	74	52	98
	34	98	94	100	100
	39	100	86	56	100
	44	100	90	96	100
	7	100	38	94	100

	14	100	40	90	100
	24	100	98	34	100
	26	100	100	56	98
	40	96	100	86	100
	41	100	56	100	100
	50	100	80	54	100
	$\bar{x}$	100	76	73	94
	SD	1.2	22.8	24.5	20.8
C	3	98	74	86	100
	4	100	96	98	100
	28	100	88	100	100
	30	100	72	100	100
	37	100	98	96	100
	48	100	92	92	100
	12	100	64	56	98
	16	100	96	76	100
	18	100	100	98	100
	19	100	94	92	100
	21	100	76	60	100
	23	100	46	92	100
	35	100	88	64	100
	$\bar{x}$	100	83	85	100
	SD	0.5	16.0	16.0	0.5

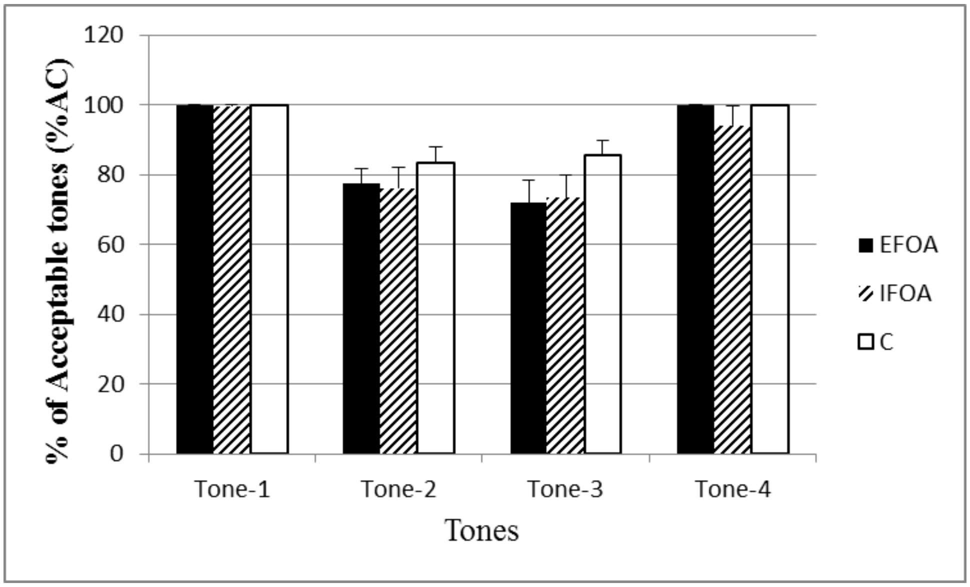


Figure 7 Percentage of the correctly produced tones in the acquisition phase averaged for the three groups: EFOA, IFOA and C

8.3.1 Analysis of the correctly perceived tones for each group

In order to assess whether the tones were produced with the same accuracy for each group, the non-parametric Friedman test was performed for each group.

8.3.1.1 EFOA group

The Friedman test results indicated a significant difference on the percentage of the correctly perceived productions among the four tones: $H(3) = 33.633$, $P < .01$. Table 59 presents the means and standard deviation of the percentage of correctly perceived tones for the EFOA group. In order to find the pattern of differences among the tones, bonferroni-adjusted post-hoc comparisons were computed using the Wilcoxon-matched pair test. As shown in Figure 7, for

EFOA group, Tones 1 and 4 were produced with the highest accuracy, followed by Tones 2 and Tone 3; Tone 3 scored lowest on the accuracy. The Wilcoxon-matched pair test identified a significant difference between the following: Tones 2 and 1; Tones 3 and 1; Tones 2 and 4; and Tones 3 and 4. The results for all comparisons are presented in Table 60.

Table 60 Wilcoxon-matched paired-test results (post-hoc comparisons) on the percentage correct productions for the EFOA group

Tone-pair	*z*	*P*
Ma-1—Ma-4	-0.577	.564
Ma-3—Ma-2	-0.534	.593
Ma-2—Ma-1	-3.235	.001
Ma-3—Ma-1	-3.297	.001
Ma-2—Ma-4	-3.183	.001
Ma-3—Ma-4	-3.297	.001

8.3.1.2 IFOA group

The Friedman test results indicated a significant difference on the percentage of the accurately perceived productions among the four tones: $H(3) = 20.328$, $P < .01$. Table 59 presents the means and standard deviation of the percentage of correctly produced tones for the IFOA group. In order to find the pattern of differences among the tones, bonferroni-adjusted post-hoc comparisons were computed using the Wilcoxon-matched pair test. As shown in Figure 7, for IFOA group, Tone 1 was produced with the highest accuracy, followed by Tones 4 and Tone 2; Tone 3 scored lowest. The Wilcoxon-matched pair-test identified a significant difference between the following: Tones 2 and 1; Tones 3 and 1. That is, Tones 2 and 3 were perceived significantly more poorly then tones 1 and 4, but not significantly different from each other. The results for all comparisons are presented in Table 61.

Table 61 Wilcoxon-matched paired-test results (post-hoc comparisons) on the percentage correct production for IFOA group

Tone-pair	*z*	*P*
Ma-1—Ma-4	-0.412	.680
Ma-3—Ma-2	-0.314	.753
Ma-2—Ma-1	-3.005	.003
Ma-3—Ma-1	-3.114	.002
Ma-2—Ma-4	-2.167	.03
Ma-3—Ma-4	-2.669	.008

8.3.1.3 Control group

The Friedman test results indicated a significant difference among the four tones: $H(3) = 29.147$, $P < .01$. Table 59 presents the means and standard deviations of the percentage of correctly perceived tones for the C group. In order to find the pattern of differences among tones, bonferroni-adjusted post-hoc comparisons were computed using the Wilcoxon-matched pair test. As shown in Figure 7, for C group, Tone 1 was produced with the highest accuracy, followed by Tones 4 and Tone 3; Tone 2 scored lowest. The Wilcoxon-matched pair test identified a significant difference between the following: Tones 2 and 1; Tones 3 and 1; Tones 2 and 4; and Tones 3 and 4. In summary, Tones 1 and 4 were significantly higher than Tones 2 and 3, and these two tone pairs were not significantly different from each other.

The results for all comparisons are presented in Table 62.

Table 62 Wilcoxon-matched paired-test results (post-hoc comparisons) on the percentage correct production for the control group

Tone-pair	*z*	*P*
Ma-1—Ma-4	-0.816	.414
Ma-3—Ma-2	-0.210	.833
Ma-2—Ma-1	-3.183	.001
Ma-3—Ma-1	-2.941	.003
Ma-2—Ma-4	-3.185	.001
Ma-3—Ma-4	-2.751	.006

As can be seen from Tables 59 and Figure 7, the percentage of correctly perceived tones for both Tone 2 and 3 were higher for the C group than for the EFOA group and the IFOA group. To test whether this difference among the groups was significant, the Kruskal-Wallis test was performed. The results indicated the percentage scores for Tones 2 and 3 were not significantly different among the groups ($\chi^2(2) = 0.917$, $p = .632$; $\chi^2(2) = 3.556$, $p = .169$; for Ma-2 and Ma-3, respectively).

8.3.1.4 Additional analysis

To assess if any learning occurred among the three groups, for each group, the retention percentage of correctly perceived tones was compared to the baseline percentage of correctly perceived tones using the Wilcoxon-matched paired test. The result indicated that the only significant difference between the pre- and post-test scores was in the control group Ma-3 production. The percentage of correctly perceived tones was significantly higher in the retention test (85%) than the baseline (61%). No other significant differences emerged. The results of the pre- and post-tests for the three groups are presented in Table 63.

Table 63 Results of the Wilcoxon-matched paired test; comparing baseline to retention %AC for each group

Syllable-tone	Groups	*z*	*p*
Ma-1	EFOA	-1.63	.102
	IFOA	-1.41	.157
	C	-0.58	.564
Ma-2	EFOA	-1.64	.101
	IFOA	-1.65	.098
	C	-0.90	.366
Ma-3	EFOA	-1.80	.072
	IFOA	-1.62	.096
	C	-2.16	.031
Ma-4	EFOA	-1.34	.180
	IFOA	-0.45	.655
	C	-1.00	.317

8.4 MANIPULATION CHECK QUESTIONNAIRE

Due to the conceptual nature of the independent variable in this study, all participants were required to answer a short questionnaire at the end of the experiment as a manipulation check to insure that the participants perceived and followed the instructions (modified from Porter, Nolan, Ostrowski, & Wulf, 2010; Fasoli, Trombly, Tickle-Degnen, & Verfaellie, 2002). Using a five point equal appearing interval scale, the participants rated the difficulty of the task, the usefulness of the instructions and feedback, and the ability to maintain the same focus of attention as instructed throughout the experiment (see appendix D for the manipulation check questionnaire). Choosing from multiple answers, the participants indicated their focus of attention during their tone production in the acquisition phase. In addition, the participants addressed whether they tried to focus as instructed and whether they thought they learned the four tones equally well. They ranked the figures of the four pitch contours, referring to the four tones, according to how they perceived their difficulty.

As shown in Figure 8, most of the participants ranked the task between easy and difficult (rating the task as 2, 3, or 4 on a 5-point scale, where 1 = extremely easy and 5 = extremely difficult). Nineteen of the 42 participants rated the task as “somewhat easy;” almost half of these responses came from the participants in the control group, and the other half came from a combination of the other two groups (EFOA = 4, IFOA = 7, C = 8). Fifteen of the 42 participants reported the task as “easy;” almost half of these responses came from the participants in the EFOA, group and the other half came from a combination of the other two groups (EFOA = 7, IFOA = 4, C = 4). Only six participants thought that the task was difficult (EFOA = 1, IFOA = 3, C = 2). One individual in the EFOA group stated that the task was very easy; no one rated the task as extremely difficult. Collectively, 85% of the participants perceived the task as ranging from “extremely easy” to “somewhat easy.”

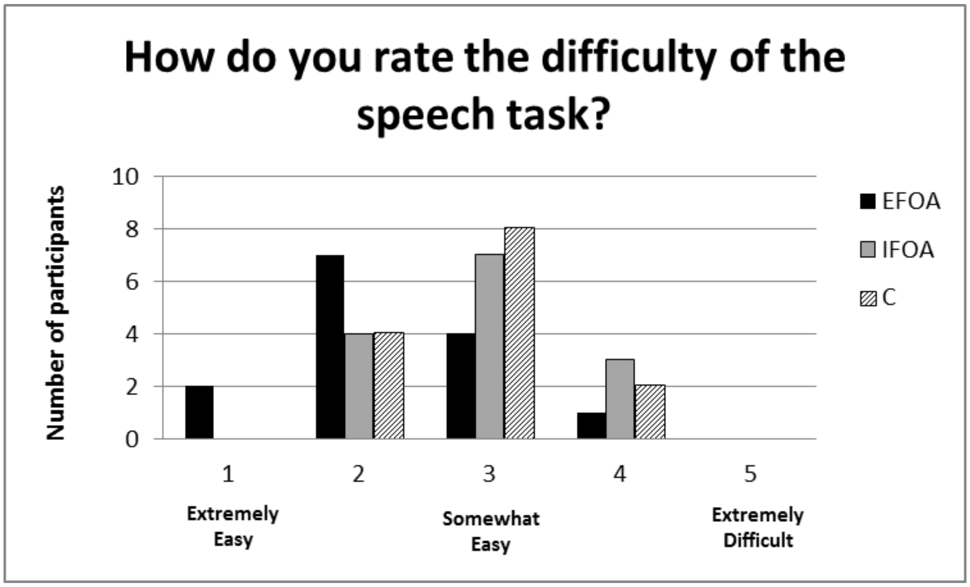


Figure 8 The number of participants responded in each of the categories for the question: How do you rate the difficulty of the speech task?

When the participants ranked the four tones by ordering the tones from the easiest to the most difficult, the overall response varied among the groups. As is apparent in Table 64, only nine participants perceived the easy tones (1 and 4) as easy and the difficult tones (2 and 3) as difficult (EFOA= 2 participants; IFOA = 4 participants; C = 3). On the other hand, only three participants in the EFOA group perceived the difficult tones (2 and 3) as easy and the easy tones (1 and 4) as difficult. The majority of participants (71%) ranked either Tone 1 or Tone 4 as easy followed by either Tone 2 or Tone 3 (EFOA= 9; IFOA = 10; C = 11). The participants' production of these tones also reflected their perception of the difficulty among the tones (See

Figure 7). Despite these differences among the tones in terms of perceived difficulty and in production, half of the participants thought that they did not learn the tones equally well.

Table 64 The participants' ranking of the four tones from the easiest to the most difficult tone in the three groups: EFOA, IFOA and C

The order of the tones (four-pitch contours) according to the participants' perceived difficulty (from the easiest to the most difficult tone)	Number of responses		
	EFOA	IFOA	C
1,3,2,4	1	5	3
1,3,4,2	2		
1,4,2,3		1	
1,4,3,2		2	
2,4,3,1	1		1
2,3,4,1	1		
3,1,2,4	1		4
3,1,4,2	1	1	
3,2,1,4	2		
3,4,1,2	2		
3,4,2,1			2
4,1,2,3	2	1	3
4,2,1,3		3	
4,2,3,1		1	1
4,3,1,2	1		

As shown in Figure 9, almost all participants (95%) reported that the feedback was helpful; however, the degree of perceived helpfulness varied among them. Only one participant in the EFOA group reported that the feedback was "not at all helpful". Another participant in the control group ranked the feedback between "somewhat helpful" and "not at all helpful". It is worth mentioning that the feedback was identical for all groups; the instructions were the only variable manipulated in this study.

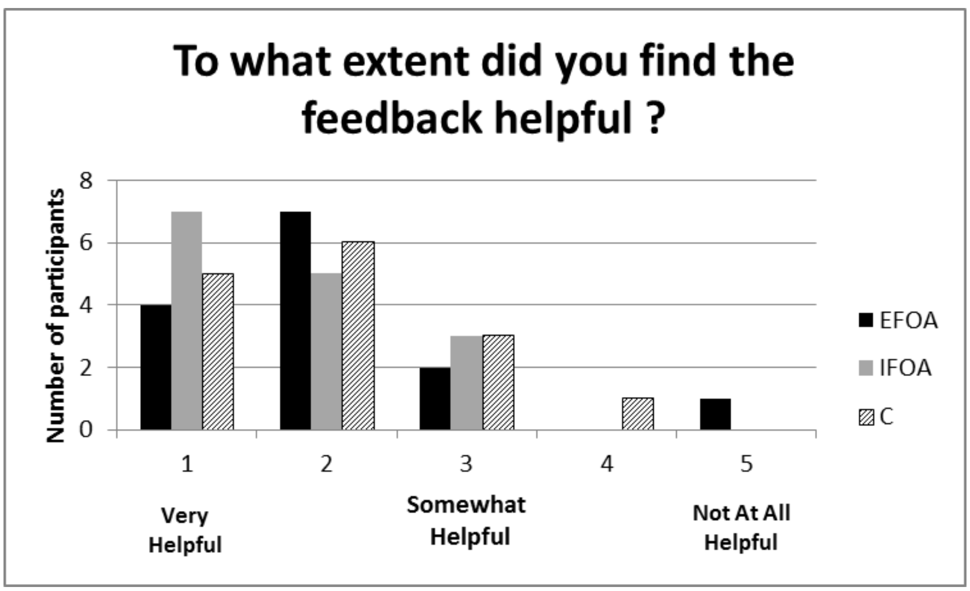


Figure 9 The number participants responded in each of the categories for the question: To what extent did you find the feedback helpful?

The questionnaire was primarily administered to assess the effectiveness of the FOA manipulation. The participants' responses as to the helpfulness of the instructions were wide-ranging. Figure 10 illustrates that 64% of the participants in the control group thought that the instructions were "very helpful," and the other 36% thought that the instructions were "helpful." Moreover Figure 10 illustrates that the extent to which the participants felt that the instructions were helpful varied between both the EFOA and IFOA groups. Only two participants, one in the EFOA group and the other in the IFOA group, ranked the instructions as between "somewhat helpful" and "not at all helpful." None of the participants thought that the instructions were "not at all helpful".

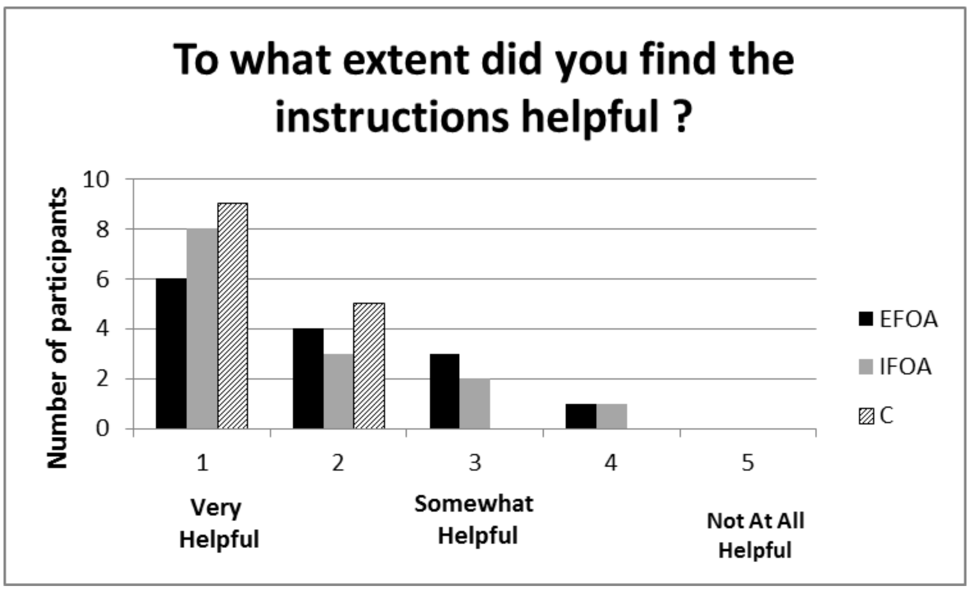


Figure 10 The number of participants responded in each of the categories for the question: To what extent did you find the instructions helpful?

Although all participants responded “yes” to whether they focused as instructed, their responses to the next question indicated that this was not the case. The Pie Chart for each group show how the participants responded to the question about which aspect of the task they focused on (Figure 11, 12, 13). The Pie Chart indicates that the focus of attention varied among the participants in the three groups. It is important to note that for this question, the participants could choose from five responses (see Appendix D). The only response solely considered an internal focus of attention dealt with “vibration in the larynx.” To stress “produced sound” for the external focus of attention group, and “vibration in the larynx” for the internal group, the instructions were bolded on the screen and emphasized in the recorded directions.

The responses on the pie chart for EFOA group indicate that only four participants actually focused as instructed by indicating that they focused on “sound I produced” (Figure 11). Another two participants in the EFOA group reported focusing on “sound I produced,” but one of these participants included that she had also focused on “the feedback,” while the other participant reported that she had focused on other aspects of the task; this second participant provided a written response stating that she focused on “matching tone by listening.” Six of the 14 participants in the EFOA group reported that they focused their attention on “sound I produced,” which indicated an external focus of attention, as instructed. Another six participants in the EFOA group reported that they focused on the feedback. The response of one participant in the EFOA group suggested that she utilized both EFOA and IFOA while performing the speech task (focused on “vibration in the larynx,” “feedback,” and on “sound she produced”). Interestingly, one participant reported attending to the “vibration in the larynx,” although the instructions did not mention this aspect of the task to the EFOA group.

It is apparent in the Pie chart of the IFOA group responses that only two participants focused as instructed (Figure 12). One participant indicated that she focused on “vibration in the larynx” and “feedback.” Another three participants in the IFOA group focused on a mixture of internal and external foci; “vibration in the larynx”, “sound I produced,” and on the “feedback.“ Three of the participants in the IFOA indicated that they only focused on “sound I produced”, while another three participants reported that they focused on both “sound I produced” and on “feedback”.

The control group only received general instructions without any emphasis on type of focus of attention. Responses on the pie chart (Figure 13) indicate that one participant reported that she “did not focus on anything in particular.” Half of the control group participants thought that they focused externally on the “sound I produced.” Another five participants in the control group focused on “feedback”. The remaining participant said that she focused on both “sound I produced” and “feedback.” The responses of 13 participants indicated that they focused on “the feedback” only (EFOA = 6 participants; IFOA = 2 participants; C = 5 participants). One-third of the participants (31%) indicated that they focused on “the feedback” only (EFOA = 6 participants; IFOA = 2 participants; C = 5 participants).

The question “Which aspect of the task did you focus on” garnered the following responses: 28% of the participants in the EFOA group, 16% of the participants in the IFOA group, and 7% of the participants in the control group focused exclusively as instructed.

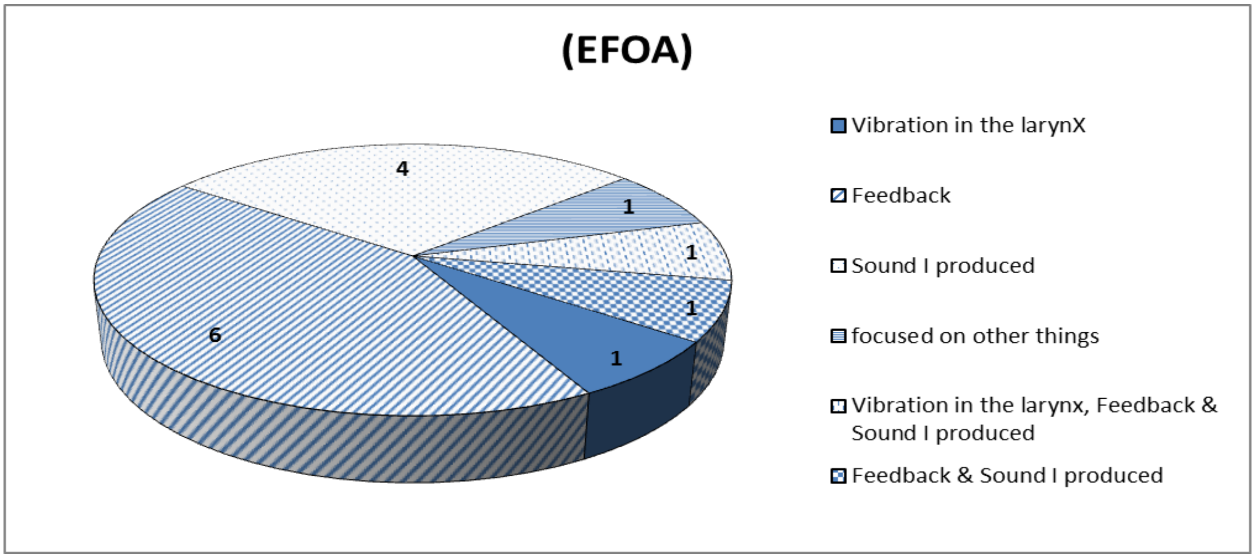


Figure 11 The responses of the EFOA group to the question: which aspect of the task did you focus on?

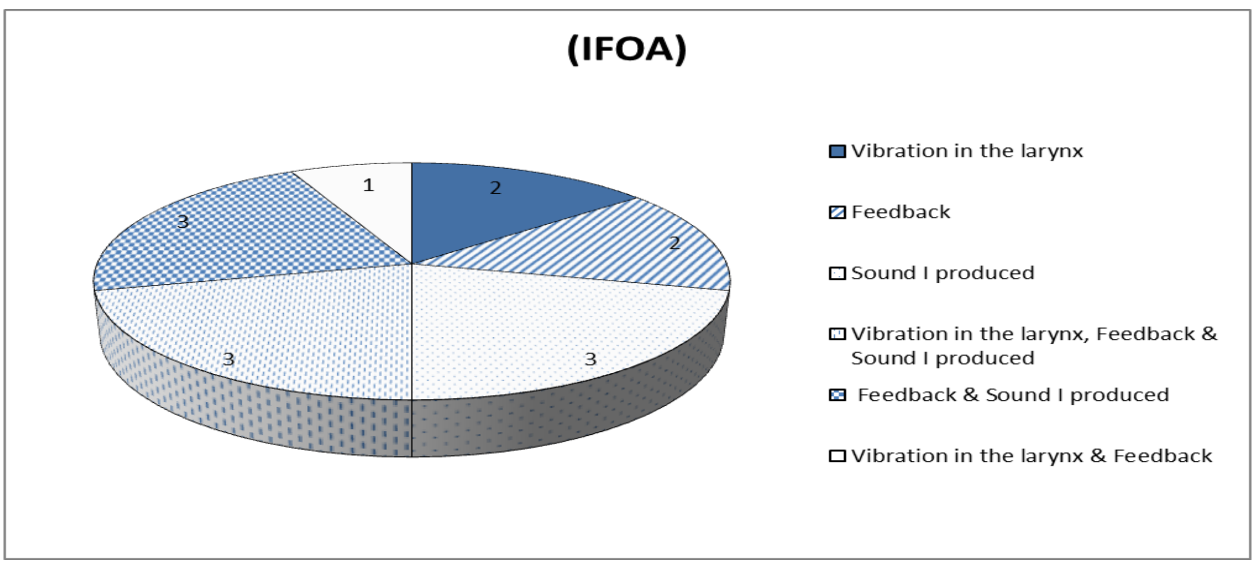


Figure 12 The responses of the IFOA group to the question: which aspect of the task did you focus on?

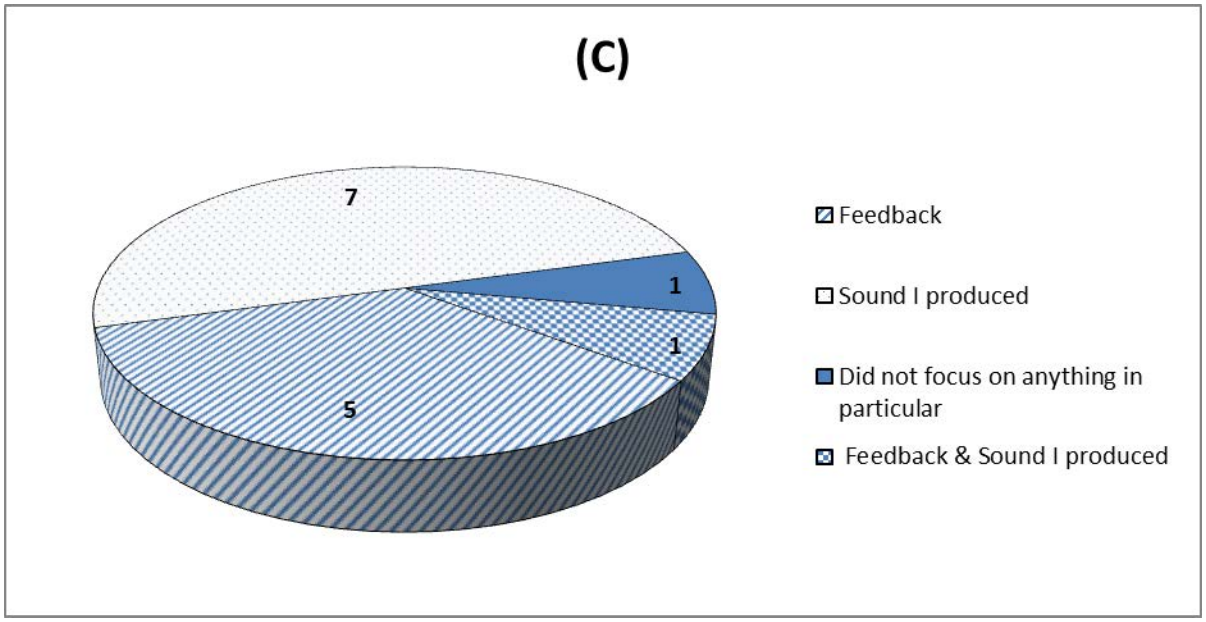


Figure 13 The responses of the Control group to the question: which aspect of the task did you focus on?

Figure 14 show the participants' response to whether they focused on the same aspect of the task throughout the acquisition, almost half of the participants stated, (23) "Most of the time" (EFOA = 4; IFOA = 11; C = 8). Twelve participants responded as "All of the time" (EFOA = 5; IFOA = 2; C = 5). Five participants responded "some of the time" (EFOA = 4; IFOA = 1). Few responded that they focused "little of the time" (EFOA = 1; C = 1). None of the participants reported that they shifted their focus or stopped to focus on the same aspect of the task.

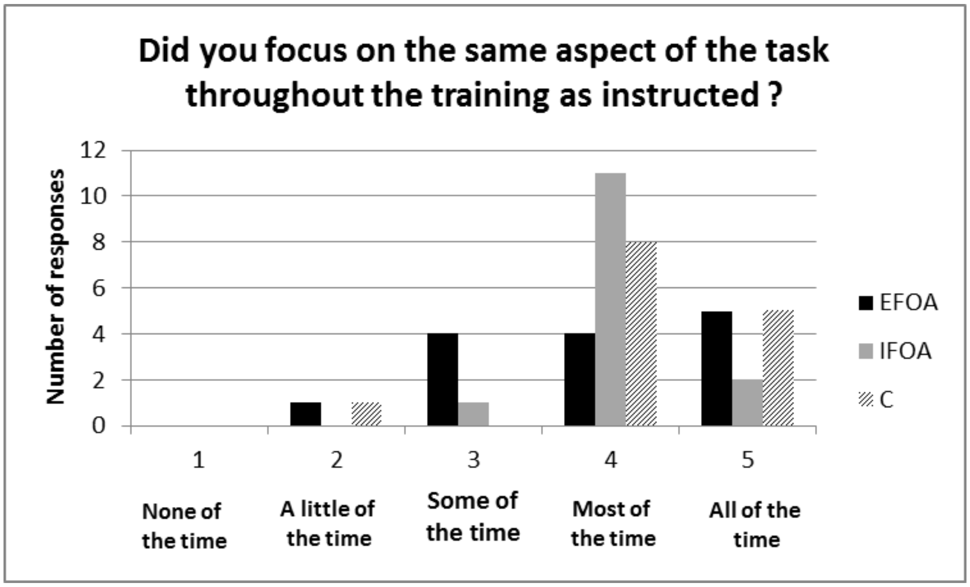


Figure 14 The participants' responses to the question: Did you focus on the same aspect of the task throughout the training as instructed?

To summarize, this questionnaire was administered to ensure that the participants considered the instructions and focused as directed. The FOA varied in the responses. Of the 14 EFOA participants, only four reported they focused solely on "sound I produced". Two other participants reported focusing on "sound I produced" in addition to focusing on other aspects of the task, such as "vibration in the larynx" or "feedback". One participant in the EFOA group reported that she focused on "other things," specifically by matching the tone by listening. Though this response indicates an EFOA, it is not clear whether the participant meant the sound of the model or her sound; and if that was the case, why the participant did not simply choose "sound I produced".

Of the participants who were instructed to focus internally, only two reported they focused solely on “vibration in the larynx.” Four other participants reported focusing on “vibration in the larynx” in addition to other aspects of the task, such as “feedback” and “sound I produced.”

Participants in the control group received general instructions without any specific focus. Only one participant reported she “did not focus on anything in particular.” Half of the participants reported focusing on “sound I produced.” Five participants thought they focused on “feedback,” while one participant reported that she focused on both “sound I produced” and “feedback.”

9.0 DISCUSSION

According to Speech-Language pathologists and researchers interested in speech motor control, the treatment of motor speech disorders can benefit from the implementation of the principles of motor learning derived from the limb literature (Adams & Page, 2000; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000; Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008; McNeil, Katz, Fossett, Garst, Szuminsky, Carter, & Lim, 2010). Evidence from the limb motor literature also indicates that FOA is another factor that affects limb motor learning. Researchers have demonstrated that adopting an external focus of attention improved the performance and learning on several motor tasks (Wulf, 2007a; Wulf, McConnel, Gartner, & Schwarz, 2002; Wulf & Su; 2007; Wulf, Landers, Lewthwaite, and Töllner, 2009; Wulf & Prinz, 2001; Zachry, Wulf, Mercer, & Bezodis, 2005). Moreover, researchers interested in speech motor learning have shown that EFOA improved the performance on a non-speech task that involved the oral motor control system (Freedman, Maas, Caligiuri, Wulf, & Robin; 2007). Although researchers in the speech domain have not yet tested the construct of focus of attention, speech emerges as a viable motor task to test it based on the following premises: 1) speech motor control utilizes a common anatomy that is shared with the oral motor control (e.g., Ballard, Robin, & Folkins, 2003; Robin, Solomon, Moon, & Folkins, 1997) and 2) empirical evidence from the motor learning literature shows the successful transfer of principles of motor learning to the speech domain (Adams & Page, 2000; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000; Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008; McNeil, Katz, Fossett, Garst, Szuminsky, Carter, & Lim, 2010). In addition, FOA has the potential to be directly incorporated during any speech therapy, either through the instructions or the feedback that are considered integral elements in any treatment setting. If the FOA were demonstrated to influence the performance and learning in the speech domain, such knowledge would help in the construction and refinement of more effective treatments that would benefit the learners.

The purpose of the current study was to extend the notion of FOA from the limb motor literature to the speech domain by testing the effects of FOA on the learning of non-native (Mandarin Chinese tones/words) speech by young American English monolingual adults. Forty-two young-adult females [18-24 years] were randomly assigned to one of three learning groups: 1) EFOA group, 2) IFOA group, and 3) control (C) group.

The experimental paradigm involved the following: baseline phase, acquisition phase, retention tests [immediate and delayed], and transfer test. The participants completed the baseline phase, acquisition phase, and immediate retention test during the first session. During the second session—which occurred 24-48 hours after the completion of the first session—the participants completed the delayed retention test and the transfer test. After the completion of the transfer test, the participants answered a short questionnaire as a manipulation check for the independent variable (FOA). The experimental task consisted of repeating the Mandarin Chinese word after an auditory model that had been recorded by a native Chinese speaker. All participants received feedback with a frequency of 60% on their production during the acquisition phase. Although all groups engaged in the same experimental procedures, the groups differed in terms of which instructions they received before each block of practice, during their acquisition phase. The instructions represented two levels of FOA: EFOA and IFOA. The third group was a control group.

To avoid examiner bias, the instructions were pre-recorded for each group. The participants simultaneously read and listened to the instructions. In order to test the differences among the groups, the current study used two dependent variables to measure the participants' productions: 1) an acoustic measure of error (RMSE) that indexed how the produced pitch contour deviated from the model and 2) a perceptual measure of accuracy; the percentage of correctly perceived productions judged by native Mandarin speakers.

Based on the FOA literature with its accumulating evidence of the benefit of adopting an EFOA in learning limb/body motor tasks, this study predicted that the group that received the EFOA instructions would outperform the other two experimental groups during all phases of the experiment subsequent to the baseline: the acquisition phase, IRT, DRT, and TR. Specifically, this study predicted that participants in the EFOA group would outperform those in the IFOA and C groups. That is, the pitch contours of the tones produced by participants in the EFOA condition would have less RMSE compared to the model's pitch contour. When judged perceptually by native Chinese listeners, the monosyllabic words produced by the EFOA group would have a higher percentage correct score (relative to the other groups: IFOA and C). Moreover, in the single subject analysis, each participant in the EFOA group would demonstrate a larger effect size in her performance accompanied by a decrease in RMSE level during the acquisition phase, compared to her baseline performance.

Contrary to the predictions, there was no significant difference among the three FOA groups on the dependent variables. In other words, the results were not affected by the instructions on either the RMSE or the percentage of correctly perceived productions for any of the experiment phases. Moreover, the single subject analysis did not yield clear evidence of an immediate effect of the instructions for any of the participants. A post experimental questionnaire suggested a wide range of interpretation of the experimental tasks in all groups.

The findings of this study do not support previous research reporting the benefits of EFOA instructions on the performance and learning of a motor task (Wulf, 2007a,b; Wulf, McConnel, Gartner, and Schwarz, 2002; Wulf & Su; 2007; Wulf, Landers, Lewthwaite, and Töllner, 2009; Wulf & Prinz, 2001; Zachry, Wulf, Mercer, & Bezodis, 2005) as applied to this speech task. However, in the additional analysis, the C group demonstrated a significantly higher percentage of correctly perceived productions for the most difficult tone (Ma-3) in the retention test relative to its baseline. Moreover, during the acquisition phase, the mean *CoV* of the C group was significantly lower than its baseline CoV for Ma-2 (an equally difficult tone) productions.

A secondary interest and internal experimental control of the current study was to assess whether the participants would produce the four tones equally well during acquisition. The percentages of correctly perceived productions during the acquisition phase were analyzed. Based on the second language learning literature, it was predicted that the four tones would not be produced equally well; rather, the tones' percentages of accurately perceived productions would demonstrate a hierarchy. The results accord with the second language learning literature. Tones 1 and 4, had significantly higher accuracy scores than Tones 2 and 3. There was no significant different between Tones 1 and 4 or between Tones 2 and 3. This hierarchy was demonstrated in all groups irrespective of the instructions received.

Several limb motor studies have indicated that FOA benefits both performance and learning (e.g., Wulf & Su; 2007; Wulf, Landers, Lewthwaite, and Töllner, 2009; Wulf & Prinz, 2001; Zachry, Wulf, Mercer, & Bezodis, 2005). However, the results of the current study are not consistent with these findings. This study failed to demonstrate benefits from EFOA on the learning of this tonal task.

The primary cause for the non-significant results in the current study may be due to the underpowered nature of this study. The current study used a sample size of 42 participants (14 participants per group), which was based on a priori power analyses that indicated that this sample size was adequate to detect a large effect size (f = .4), with $\alpha = .05$ and an adequate power of .8. Although this study was set with an adequate power to detect large effect sizes, the observed effect sizes were small, which rendered the current study to be underpowered to detect such small effect sizes. The effect sizes observed in this study, during acquisition, retention and the transfer phases, were $f = .1$, $f = .1$ and $f = .11$, respectively. To calculate the power of the current study, these obtained effect sizes were fed in a reverse power calculator (G-power 3.1 software). The current study had a power of only .21, .18, and .39 to detect such small observed effect sizes during the acquisition, retention, and transfer tests, respectively. Therefore, instead of 42 participants (14 participants per group), a sample size of 246 participants (82 participants per group) would be needed to detect an effect size of .1 with an alpha of .05 and a desired power of .80.

This is not to argue that by merely increasing the sample size, the power of the current study would have increased. Nonetheless, it is important to consider other factors that might have contributed to the small effect size observed in the current study. As Baguley (2004) states, "…there is a danger that routine use of statistical power leads to neglect of other important factors in the design of a study" (p.77). Therefore, in addition to the small sample size in the current study, other limitations might have contributed to the small effect size and affected the power.

The current study is the first to test the FOA construct in the speech domain. Although the groups did not demonstrate statistical differences based upon the instructions they received, it would be premature to make strong claims about the role of FOA on speech learning. Instead, it is critical to gain additional insight into the potential factors that might have accounted for the current results.

Moreover, FOA is only one among other factors that might have affected performance and resulted in the small observed effect size and low power. It can be speculated that this discrepancy between the results of the current study and the results of the literature of FOA could be due to 1) differences in the motor system targeted, 2) differences pertaining to the task utilized, 3) differences in the instructions provided, and 4) the dependent variables utilized.

A discussion of these factors is warranted in order to provide a context for the findings, to highlight some limitations of the current study, and to provide suggestions for future studies. The following sections address these factors.

Methodological issues

Before considering the lack of the FOA effect, some methodological problems in designing and conducting the experiment are discussed. In the current study, the Schema Theory was used as the theoretical framework upon which to formulate the speech motor learning. Proponents of Schema Theory argue that many variables influence motor learning; designated as "principles of motor learning" (Schmidt and Lee, 2005; Shea and Wulf, 2005). The practice and feedback variables are among the most influential factors of motor learning. Schema Theory makes explicit predictions about how manipulation of these principles should affect motor learning.

These principles of motor learning, demonstrated in both the limb and speech motor literature to enhance learning, were inherent in the design of the current study. Practice is the most significant principle proposed to influence motor learning (Schmidt and Lee, 2005). Amount of practice refers to the amount of trials a performer executes on a task (number of trials before retention or transfer are tested). The number of practice trials in the current study was based on previous research that required young adults to produce novel speech tasks in a motor learning paradigm (Adams & Page, 2000; Kim, 2007; Steinhauer & Grayhack, 2007). In the current study, the acquisition phase consisted of 200 trials (50 practice trials for each tone). Considering the importance of the variability of practice in the motor learning literature (Adams & Page, 2000, Hall & Magill, 1995; Lee, Wulf, Schmidt, 1992), trials of different tones were practiced in random order. In the current study, the four tones produced with the Ma syllable were practiced in random order within each block, with no word being consecutively repeated more than twice. Random practice, which has been shown to enhance the retention and transfer of motor skills relative to blocked practice, may facilitate learning by increasing the difficulty of the learning environment and by approximating a natural context (Knock, Ballard, Robin, & Schmidt, 2000). Although this study based the amount of practice employed on previous studies in which it has been sufficient to effect change in performance and to result in learning, the design in the current study appears to have been too short to produce a detectable and reliable change in performance.

In the current study, the variability in the baseline data challenged the interpretation in the single subject analysis. As this study had a defined set of baseline recordings to limit any possible learning from exposure, future studies might extend the baseline recording to ensure stability before initiating the acquisition phase. Moreover, as evident from the single subject analysis showed that the effects of FOA were not immediately evident from the shallow slopes and the unchanged levels during the acquisition phase, more practice might be needed to demonstrate these effects. Moreover, insufficient practice for this specific tonal task might explain the lack of acquisition in the current study. One discrepancy between the previous studies on tones acquisition and the current study relates to the participants' experience with tonal languages. Unlike Guo and Tao (2008), who studied the percentage of acceptable tones by American students who were already involved in Chinese language classes, or Wang, Jongman, and Sereno, (2003) who involved students who had enrolled in one or two semesters of Chinese courses, the current study was designed as a novel learning study. One of the inclusion criteria for participating in the current study was that the participants had no previous experience with tonal languages. These inexperienced participants may have required a longer acquisition phase before they were able to demonstrate changes in performance.

Although Schema Theory is based on the assumption that practice strengthens schema, it should be noted that the effect of practice in the motor learning literature is not always obvious. Amount of practice can interact with practice schedule: constant, blocked, or serial (Giuffrida, Shea, and Fairbrother, 2002; Shea & Kohl, 1991) as well as with different feedback conditions, which might result in obscuring its effect in some studies.

In the current study, the participants received a visual feedback on 60% of the trials. The visual feedback was presented after a three second delay with the intent to enhance learning and error detection (Swinnen, Schmidt, Nicholson, & Shapiro, 1990). In addition, according to Schema theory, the availability of such augmented feedback is linked with strengthening both the recall and recognition schema (Schmidt, 1975). Researchers in both the limb and speech motor learning literature acknowledge the importance of such augmented feedback in motor learning (e.g., Adams & Page, 2000; Adams, Page, & Jog, 2002; Ballard et al., 2007; Salmoni, Schmidt, & Walter 1984; McNeil, Katz, Fossett, Garst, Szuminsky, Carter., et al., 2007; Steinhauer & Grayhack, 2000). Additionally, based on the guidance hypothesis, researchers have demonstrated that more frequent feedback, such as that provided on every trial, enhances acquisition due to the guidance it provides, but also diminishes performance during the retention tests (Salmoni, Schmidt, & Walter, 1984). Although both the amount of practice and the feedback frequency employed in the current study were selected because of the empirical evidence of their effectiveness in motor learning, they might not have been sufficient to enhance the learning of the tonal task. One difference between the current study and studies that focused on learning Mandarin Chinese as a second language (Albertson, 1982; Chun, 1989; Leather, 1990; Stibbard, 1996; Weltens & De Bot, 1984) is that those studies utilized visual feedback with 100% frequency, and reported its effectiveness only on the participants' acquisition of the behavior not with the feedback removed.

Even though the results of the current study do not provide clear evidence that learning occurred, the effect of FOA was expected to emerge during the acquisition phase regardless of learning (maintenance). This study failed to demonstrate significant differences among the experimental groups during the acquisition. This unexpected finding is inconsistent with the results of those studies that demonstrate the beneficial effect of EFOA; in some studies, the effect of FOA instructions was apparent following the first few trials (e.g., McNevin, & Wulf, 2002), while in other studies, the effect did not emerge during practice but was only obvious during the retention test (e.g., McNevin, Shea & Wulf, 2003).

Novelty and complexity of the speech task

The task utilized may be another possible factor that might have contributed to an absence of an EFOA effects. As highlighted by researchers in the FOA literature (Wulf, Toller, and Shea, 2007), in order for the EFOA effects to manifest themselves, the task should be complex and pose a challenge to the performer. The task employed in this study was novel and its complex characteristics appeared to render it suitable to assess the FOA effect. It required finite control over the articulators to change the fundamental frequency during the vowel production. This tonal task required the participants to change the pitch of their voice (by changing the vibration of the vocal folds rate and upper airway configuration) over short durations at the syllable level.

Indeed, the results confirmed the prediction that the participants would not produce the four tones at the same accuracy. These results were consistent with those of Guo and Tao (2008), when judged by Mandarin Chinese speakers, that the percentage of correctly perceived tones produced by English speakers was higher for both Tones 1 and 4 than for both Tones 2 and 3. Moreover, no significant difference occurred between Tones 1 and 4 or between Tones 2 and 3. However, each of the Tones 1 and 4 differed significantly from both Tones 2 and 3. Higher error scores (RMSE) for Tones 2 and 3 than Tones 1 and 4 reflected this hierarchy of difficulty among the tones. Although Tone 4 was frequently shown to be easy, in the current study the error in Tone 4 production, when measured by RMSE, was higher than the error in either Tones 2 or 3 for some words. This high error rate of Tone 4 might be attributed to the limited falling pitch range in English speakers when compared to the pitch range in Mandarin Chinese speakers (Chen, 1974; White, 1981). Chen (1974) and White (1981) noted that Mandarin Chinese speakers utilize a-one-and-a-half time larger pitch range than the pitch range of the non-native Mandarin Chinese speakers. The source of error on Tone 1 production, although the lowest among the tones, might be attributed to the fact that Tone 1 is usually produced at a different pitch register by non-native Mandarin Chinese speakers (Miracle, 1989).

Tones 2 are 3 are difficult tones, both in terms of production and perception, because they require a specific change in the fundamental frequency (F0) at a specific relative timing of the vowel (Wang, Jongman, & Sereno, 2003). Because tones in the English language are not utilized to differentiate words, English speakers pay less attention to the important perceptual cues such as pitch contour (Gandour, 1983). This decreased sensitivity to the important perceptual cues (specifically F0 changes) of the tones might have reduced the participants' awareness of how to achieve the task goal and to perceive the difficulty of the task.

In order to overcome any disadvantage resulting from the lack of sensitivity or awareness towards specific tonal changes, future studies might include some perceptual training before the acquisition phase. The rationale for not including such perceptual training in the current study was to decrease the participants' exposure to the tones until baseline recording to avoid learning by simple exposure.

Although the current study failed to demonstrate an FOA effect in the speech domain, the hierarchy of difficulty among the four tones is consistent with the literature. The difference among the tones was demonstrated on both dependent variables: RMSE and percentage of correctly perceived productions. This robust finding may be inconsistent with a hypothesis that the task was not complex enough.

The word frequency of the English words presented on the computer screen that accompanied the pitch contour feedback might offer an alternative explanation for the hierarchy demonstrated among the tones. Typing the meaning of the words on the feedback screen may have contributed to the hierarchy of the tones. For example, according to the SUBTLEX word frequency database, the word “mother”—the meaning of Tone 1—occurs more frequently than the word “horse”—the meaning of Tone 3. Similarly, the verb “to scold”—the meaning of Tone 4—occurs more frequently than the word “hemp” —the meaning of Tone 2.

However, the current data do not support this alternative explanation of the hierarchy because a similar hierarchy of tones was also demonstrated during the baseline phase and learning tests productions during which no feedback was provided.

Although the hierarchy of tones was consistent with the hierarchy in previous studies in the second language literature, the acoustic analysis revealed a significant interaction between the syllables and the tones. The RMSE for Tone 4 was significantly higher than for Tone 2 during the Na syllables productions but not significant in the other two syllables (Ma and Me). This interaction between the syllables and tones was not expected, because the only difference between the Na and Ma syllables resides in the manner of the consonants’ articulation. Therefore, further exploration is required in order to more clearly interpret this interaction.

As discussed above, the results indicate that, despite the complexity of the task, no change in performance was demonstrated as an effect of learning or practice. Therefore, the question to tackle next is why the participants did not demonstrate the expected performance change, even with this limited amount of practice.

The Challenge Point Framework proposed by Guadagnoli and Lee (2004) might offer a possible answer to this question. According to this framework, learning does not depend only on the task complexity but also on the learners themselves. That is, effective motor learning requires that learners face sufficient challenges with the task. Therefore, one might argue that the participants in the current study might not have reached their challenge point when producing the tones. In the current study, the participants produced the complex tones with over-practiced syllables (Ma, Me, and Na) which might have caused the participants to be less challenged or less motivated to learn such a simple over-practiced syllable; in Duffy's words, "…under most circumstances, speech is produced with an ease that belies the complexity of the operations underlying it" (Duffy, 2005; p. 3).

As indicated by the responses in the questionnaire, 35 of the 42 participants thought that the task was either easy or somewhat easy. The non-significant difference among the groups in the current study accords with the results of Wulf et al. (2007) and Lander et al. (2005), which showed a non-significant difference among the EFOA, IFOA, and Control conditions when the participants performed a relatively easy task. In the Wulf et al. study, young adults maintained balance on a solid surface; in the Lander et al. study, participants with Parkinson's disease maintained balance while standing on a stable surface quietly with their eyes either opened or closed. Although the results of the current study contradicted the relatively consistent evidence for the superiority of EFOA under these conditions, they were consistent with the Wulf et al. (2007) Experiment 1 that hypothesized that if the task was not challenging enough for the performer, the FOA effect would not occur. Although participants did not produce each of the four tones with the same accuracy in the current study, the possibility that the participants did not perceive even the two more difficult tones as challenging enough cannot be ruled out in this study. By not perceiving the task as challenging, the young-adult participants in the current study, for whom speech is an over-practiced and automatic task, may not have been motivated to engage during the practice and to challenge themselves to achieve the goal. Researchers have indicated that not only do principles of motor learning employed during practice affect the performance and learning of the motor task, but the motivation of the performers is also an essential pre-practice factor (Schmidt and Lee, 2005).

Because six of the 42 participants reported that the task was challenging, their data can be used to explore the difficulty hypothesis. The lack of improvement in these participants may be due to the participants' trying too hard to achieve the task goal. That is, previous findings suggest that trying too hard while performing a motor task, such as a bimanual coordination task, can be detrimental to performance (Lee, 1998). Unfortunately, the performance of the six participants who reported that the task was difficult (participant 43 from the EFOA group; participants 11, 24, and 27 from the IFOA group; and participants 19 and 28 from the C group) did not differ from the other 36 participants in any relevant way. While these data do not support this post-hoc hypothesis, the current data do not allow an adequate exploration of the relationship between perceived task difficulty and performance, the possibility of task difficulty perception on the performance cannot be ruled out.

Given the current results, another question that emerges is whether FOA would enhance the performance of the speech task in individuals with no speech motor control problems. Although any conclusion is highly suspect, some findings in the current study raised this question.

The results demonstrated by the C group participants might also provide some insights into the "less is more" concept when it comes to instructions. When the current study compared the average *CoV* during acquisition to the baseline *CoV* for each group, it showed the following: 1) participants in all groups significantly decreased their variability during the acquisition of Ma-4; 2) participants in both the IFOA group and C group significantly decreased their variability during the acquisition of Ma-1; and 3) participants in the C group were the only ones to significantly decrease their variability during the production of Ma-2 (complex tone) in addition to the other two easy tones. Moreover, when compared to their percentage of correctly perceived productions at baseline, the C group demonstrated a higher percentage of correctly perceived productions of Ma-3 during the retention phase. It should be remembered that both Ma-2 and Ma-3 are considered difficult tones. Although the C group findings were not predicted based on the FOA literature, they were not surprising as they accorded with the results of some studies in the FOA literature (Cohen, 2010; Vuilleme & Nafati, 2007). The Cohen study utilized a supra-postural task to examine the effect of FOA manipulation on gait parameters. In the Cohen study, the task was to walk for a determined distance, while holding a cup filled with water; the participants were required to complete the task without spilling the water. Results of the Cohen study demonstrated that the participants in the control condition, in which no specific focus instructions were provided, performed better than in the other two conditions; EFOA and IFOA. Moreover, the Vuilleme and Nafati study examined the effect of FOA on body sway while standing still. The results demonstrated that providing additional instructions to actively control posture decreased the participants' automatic processes and the efficiency of their posture control relative to the control condition. The similarity between the current study and the studies of Cohen (2010) and Vuillime and Nafati (2007) may reside in the nature of the task utilized. The nature of the walking task is in some respects comparable to that used in the current study. Both walking and speaking are considered automatic tasks, especially when performed by individuals with no disorders that might compromise these activities. Any attempt to break down the task's components might be expected to disrupt the task's automaticity. The results of the current study, noticed in the C group, speak to the possibility that additional instructions might have affected performance in the two FOA groups. Participants in the C group, who practiced with no specific focus of attention, managed to decrease their *CoV* on a complex tone (Ma-2) while the EFOA and IFOA groups did not. Speech motor skill development, which results from refining the relation between the laryngeal (vocal folds) and the upper airway articulators as a result of practice, is consistent with the decrease in variability during the acquisition phase. Grigos (2009) linked the decrease in variability with the acquisition of more stable movements of the articulators during speech production. Although the perceptual analysis yielded no significant difference among the groups on the percentages of correctly perceived productions based on the instructions provided, it was noted that, although the difference was not significant (See Figure 7), the percentages of correctly perceived productions of Tone 2 and Tone 3 were higher for the C group than the other two groups (EFOA, IFOA). In addition, the C group demonstrated a significant increase in its percentage of correctly perceived Ma-3 productions in the retention test compared to its baseline (table 45).

Interestingly, the decreased variability in the C group emerged during the acquisition phase when the FOA instructions were absent for this group but were repeated before each block for the other two groups. Because this study showed no significant difference among the groups, whether the control group might benefit from less instruction when performing an automatic activity compared to both instruction groups is of interest. This point, however, raises questions for further research as to whether less instruction would enable young healthy adults to more effectively learn a novel speech task. No clear conclusions can be drawn from the above discussion because the differences among the groups did not reach statistical significance.

FOA Instructions

Given the conceptual nature of the independent variable in the current study, the focus of attention manipulation might not have been successful. Because the wording of the instructions was employed from a previous study that had successfully manipulated the participants' FOA, it was assumed that similar influences could be expected in the current study (Lohse et al., 2010). Regarding the construction and wording of the instructions, every attempt was made to include all of the characteristics from the FOA literature and to maintain the similarities between both the EFOA and IFOA instruction, except for the key words (See section 6.4 independent variable definitions of terms). Defining what might constitute an EFOA for speech was not a straight-forward task given the fact that while speaking we do not usually interact with an instrument. Even in the study by Freedman et al. (2007), which attempted to assess the FOA on the oral-motor system, they employed a task which required an interaction with oral-motor pressure instrument. For the current study, the consequence of the movement (the sound produced) was considered as an EFOA. The sound was also chosen because it is what we focus on naturally when we speak; we often state "that does not sound right," to evaluate what we say.

The non-significant difference among the groups in the current study might also be attributed to the possibility that both the EFOA and IFOA instructions may have induced an IFOA. If this were the case, then the participants might have focused internally in both the EFOA and IFOA groups. As a result, they would have interrupted the normal motor control processes. If both instructions in this study are considered as inducing an IFOA, then the results of this study are consistent with the findings of Landers et al. (2005), McNevin and Wulf (2002), Wulf et al. (1998), Wulf and McNevin (2003, Experiment 1), and Wulf, Weigelt, Poulter, and McNevin (2003) that show similar performance between the IFOA and C conditions. These finding suggest that when participants do not receive specific instructions, they tend to adopt an IFOA (Wulf, Weigelt, Poulter, & McNevin, 2003). The questionnaire responses from the participants in the C group indicated that half of the C group in the current study focused on the sound they produced, which was considered EFOA, in the current study. Therefore, it is not clear whether the sound produced induced IFOA in the current study.

Another reason why the EFOA and IFOA instructions did not differentially affect performance might be due to the notion of externality and distance of FOA. McNevin, Shea, and Wulf (2003) proposed that increasing the distance between the body and the external focus of attention reinforces the externalizing effect and further improves performance. A study by McNevin, Shea, and Wulf (2003) supported this hypothesis. The stability on a stabilometer was measured in four different groups according to the distance of the EFOA locus from the body: 1) the internal group was instructed to focus on their feet; 2) the near external group was instructed to focus on a marker positioned in front of their feet; 3) the far-inside external group was instructed to focus on markers positioned between their feet; and 4) the far-outside external group were asked to focus on markers positioned outside their feet. The McNevin et al. (2003) study results supported the previous finding that on the retention test, the internal focus group had significantly more errors than the external focus groups. Furthermore, the results also showed that both the far external focus groups had significantly higher movement frequency adjustments on the platform and fewer errors when compared to both the near external group and the internal group during retention. Relating the current study instructions to this notion, the proximity of the EFOA locus to the body in the current study might also explain why the groups did not demonstrate differences in their performance.

Yet another possible explanation for the current study results is that the impact of the instructions may have been unintentionally weakened in this study. Rather than informing the participants about the specific purpose of the experiment, the researcher only emphasized that the study dealt with learning a novel speech sound. The rationale for not discussing the purpose of the study was to decrease any sensitivity or bias to the treatment. This possibility that the instructions were not powerful seems less reasonable because the instructions were introduced to the participants before every practice block. Another possible reason for not following the instructions might be because the participants perceived the task to be unchallenging, as was discussed earlier; therefore they might not have paid much attention while listening to and reading the instructions to achieve the task goal. As there is no way to ascertain where the participants focused their attention during the experiment, in the current study, participants answered the questionnaire so as to attempt to assess the manipulation of the conceptual independent variable. Although all participants responded "yes" to whether they focused as instructed, their responses to the question "which aspect of the task did you focus on" indicated that this was not the case. Only four of the EFOA group, two in the IFOA group, and one in the C group indicated a focus of attention that complied with the instructions. These findings suggest that the other participants' preference as to what to focus on might have also affected the results of this study. Wulf and Prinz (2001) noted that "……individual differences do not play a significant role in the relative effectiveness of an external versus internal focus of attention. Rather the benefits of an external focus appear to be more general in nature", this perspective might not have been the case in the current study.

The following section discusses how the performance was quantitatively measured in the current study.

Levels of the dependent variables

The constrained Action Hypothesis emphasizes that adopting EFOA enhances learning through achieving automaticity; however, it does not explicitly specify the mechanism. Therefore, we are not sure at which level of the performance this automaticity can be captured. The level at which the FOA effects can be determined may be task dependent. If the task involves a movement that extends over many joints such as that required for maintaining balance on a stabilometer, which was the task employed in most of the FOA studies, the effect of adopting EFOA might be captured with an error measure relative to the task goal. However, this might not have been the case in speech. The current study employed an error measure based on the acoustic data and a correctness measure based on a perceptual analysis with the meaning of the word/tone as the target for correctness, from a closed set of four tones. These two dependent variables were most appropriate to examine the data. The acoustic analysis provided an error measure consistent with what most of the researchers who studied FOA have reported as their dependent variable. The error measure was selected in the current study to capture the changes of fundamental frequency over time and to relate the findings of the current study to the results in the literature. Moreover, because speech has a sound transmission (communicative) goal, perceptual analysis is an appropriate analysis. However, perceptual measures might not have been sensitive enough to capture any subtle differences between the groups in the current study because listeners tend to perceive the tones categorically. The Mandarin Chinese speakers have a highly proficient perceptual system that is sensitive to changes in the fundamental frequency, given its importance in tonal language. Therefore, instead of critically indicating whether the produced tones accurately represent a specific Mandarin Chinese tone, Mandarin Chinese judges in the current study might have been somewhat lenient by relying only on their categorical perception of the four tones to evaluate the productions of the participants. This may explain the ceiling effects in the perceptual analysis, which do not match the error values of the acoustic analysis that were calculated by the more sensitive RMSE. This is not to argue that perceptional measures should not to be used; they should be used when supplemented by more detailed measures in a similar study to the current one.

Another possible reason for not finding differences among the groups might involve the dependent variable employed. That is, the effects of FOA on a speech task might not be manifested at the produced acoustic signal but instead might be registered at a more proximal level in the speech motor system. It is not clear at which level of the movement the FOA effects might be realized and effectively measured. The speculation here is that the dependent variables might not have been sensitive enough to capture these changes in performance.

10.0 LIMITATIONS AND FUTURE DIRECTIONS

The discussion section reviewed many assumptions that guided the design of the study that might have led to the current results. Although this study raised more questions than it answered, these questions provide a clearer direction for future research. This section 1) addresses the limitation of the current study and 2) suggests future research to either overcome the limitations of the current study or to explore some of the raised speculations.

First, as both the results and discussion sections addressed, the effect sizes observed in the current study were small. It cannot be confirmed from the results of the current study whether this was a true magnitude of the effect which was undetected because of low power, or whether other factors might have caused this small effect sizes in the current study. Therefore, it is highly recommended that future studies explore the effect of FOA in the speech domain to obtain better estimates of the effect size of FOA on the specific speech task.

Second, this study only included young adults with no known motor speech disorders. In young adults, who have intact nervous and motor systems who perform speech in an automatic mode, the complexity of the intact speech motor system might have obscured the FOA effects and made them difficult to isolate. Therefore, testing the effects of FOA in individuals with compromised motor speech system is warranted. As Teitelbaum (2012) advocates, assessing a phenomenon in a compromised system can be revealing because “the physically simplified remainder of the nervous system cannot produce more complexity than can the whole, intact, system." Future research geared toward assessing the effects of FOA on the speech of individuals with a compromised speech motor system is warranted to further our knowledge about the significance of FOA on speech tasks and to locate the effects of FOA on the speech motor system.

Third, future research might consider the limitation of the methods employed in this study. This study used many principles of motor learning in an attempt to optimize learning; yet, an interaction among these factors might have obscured the effect of FOA. One suggestion is that future research might replicate this study by using only one of the complex tones and by employing a blocked practice as an initial step to explore the effect of FOA on a speech task. Researchers in the FOA literature who have utilized such an approach have successfully demonstrated the effects of FOA (e.g., Brydges, Dubrowski and Carnahan, 2007; Freedman, Maas, Caligiuri, Wulf, & Robin; 2007; Mornell, 2007). Another suggestion for future research is that before assessing the role of FOA with different combinations of the principles of motor learning, researchers need to first estimate the effect of FOA on a speech task. For example, future research might design an experiment that is parallel to the experiments in the FOA motor literature. First, by providing instructions and feedback on every practice trial and after the acquisition effects of FOA are demonstrated, researchers can then manipulate one principle of motor learning and assess its interaction with FOA.

Based on the results of the current study, the participants would not be expected to acquire, maintain or generalize the tones within one session. Therefore, in order to overcome the limitation due to the small number of trials in the current study, a future study might determine how much practice is required to 1) permit a change in acquisition level and slope and 2) result in learning as measured by retention and transfer (generalization) tests. To further ascertain a change in performance, future studies might require participants to practice until they reach a pre-determined criterion. To optimize acquisition and placing less emphasis on maintenance and generalization, researchers could ask participants to practice under 100% feedback and blocked practice conditions.

However, it is important to note that the motor learning advocates low frequency feedback, while the second language learning literature endorses a high frequency feedback. Therefore, future studies might investigate the role of FOA on learning this speech tonal task under a fading feedback schedule: starting the practice with 100% feedback and then decreasing the frequency of feedback as the practice progresses.

Fourth, most of the studies in the FOA literature employed tasks in which the participants interacted with an instrument. To make a speech task more similar to tasks utilized in the FOA literature, future researchers might consider the following: assessing the role of FOA on speech production while the speaker interacts with an instrument. For example, the researchers might assess the effect of FOA instructions on participants using an artificial palate or tracking movements with a transducer. The instructions to participants in the EFOA group could be to focus on touching a specific area on the artificial palate, while the instructions to participants in the IFOA group could be to focus on the movement of the tongue to the place of articulation.

In a recent study, Ballard, Smith, Paramatmuni, McCabe, Theodoros, and Murdoch (2012) utilized electropalatography (EPG) to investigate the effect of kinematic feedback frequency on learning a non-native speech sound. Although the researchers did not make FOA central to their study, they discussed their task from an FOA perspective:

One might propose that an internal focus refers to controlling the direction, timing, and force of tongue movements to produce a given speech sound and an external focus might involve attending to the resulting acoustic signal that is perceived by the speaker and their listeners. (Ballard, et al., 2012, p. 109)

The researchers also speculated that the knowledge of performance (KP) utilized in their study might have caused participants in the KP group to focus internally. Nonetheless, the main purpose of their experiment was to study the feedback; they did not investigate the notion of FOA.

Other researchers might employ a speech task comparable to the motor tasks in the FOA literature that involve maintaining a dynamic balance on a stabilometer and a movement that extends over many parts of the body. Future studies could use a speech task that requires more modulation of fundamental frequency and more noticeable changes at the level of the respiratory system to increase the movement involved in the speech task. Such a task would require the participants to pay attention to their chest movement or their tidal volume, and would enable the researcher to include more measurements, such as kinematic measures, to test the effect of FOA at several levels of the speech system.

Fifth, due to the ceiling effect, the perceptual analysis measure might not have been sensitive enough to capture any differences among the groups. To rectify this, future studies might ask native speakers to identify the word meaning without knowledge of the intended target, thereby removing the closed set limitation of the current study, which might have contributed to the ceiling effects obtained.

Sixth, the instructions were provided before every block of practice, which consisted of 20 trials. This mode of presentation might have decreased the impact of the instructions on the participants. Future studies might increase the intensity of the independent variable by repeating the instructions after every other trial.

Seventh, because the possibility of task difficulty perception effect on the performance cannot be ruled out in the current study, a future study exploring this potential interaction is highly recommended. Future research might also 1) test the effect of FOA on participants with a challenged speech motor system for whom the current task might be more challenging or 2) increase the difficulty of this tonal task for young adults with no speech production disorders. The task difficulty might be increased by 1) shortening the syllable duration, 2) speeding the response (making it a reaction time task), or 3) presenting the target within a phrasal close technique; that is a phrase for the participants to fill in the missing word, which carries the right tone, to complete the meaning of the phrase.

Another venue for future research is to explore the possibility that any further instructions applied to young adults, for whom speech is automatic, might disrupt the speech automaticity. If future studies replicated the results in the control group, they might provide more evidence that any additional instructions given to young adults without speech motor control disorders decrease rather than enhance the performance and learning of a novel speech task. If confirmed, these results would accord well with the alternative viewpoint of Beilock and colleagues (Beilock, Carr, MacMahon, & Starkes, 2002; Beilock, Wierenga, & Carr, 2002). According to this opposing perspective, any instructions that would direct the expert performer's focus of attention to the task, whether EFOA or IFOA, would cause the performance to suffer in an individual who usually carries out the task automatically. It would be interesting to test the opposing predictions of Beilock and colleagues and Wulf and colleagues in a speech task (for example, see Sims, 2010).

According to the externality and distance external focus perspective, one possible suggestion for future research is to test whether this notion applies to speech. The hypothesis of extending the externality of the FOA can be tested by including another group. The participants might be instructed to focus on conveying the meaning of the monosyllabic word to a listener siting across the room rather than simply producing the words and receiving pitch contour feedback.

Because the Constrain Action Hypothesis does not specify the level at which the effects of FOA occur, it is recommended that future studies include other measures, such as kinematic and muscle activity measures. The current study results cannot confirm any speculation about the level, at which FOA effects might be captured in the speech motor system. Therefore, to gain a better understanding about how and at which level EFOA might affect speech production, future studies should assess the effect of FOA within the confines of the muscle activity level or at the movement level. For example, studies that incorporate an EMG or kinematic measures might capture more subtle differences if present. Based upon the Constrained Action Hypothesis, the adoption of EFOA frees automatic processes to control the movements. However, participants who rely on an IFOA have a higher possibility of consciously interfering with these control processes. Researchers might expect to find a lower EMG activity and/or a higher stability index when participants focus externally. If it is confirmed that the effect of FOA might only appear at the muscular or kinematic level of the speech motor system but is not obvious on the produced speech then the question that might be raised is whether such a possible difference at a more proximal level of the speech motor system would be of practical importance.

To argue that the FOA of the speaker might not affect the speech task requires a further study that implements both a limb task and a speech task to assess whether FOA will affect both tasks. Such a study might utilize the visuomotor tracking task, a well-established task in the motor limb literature that is comparable to the speech tonal task utilized in the current study.

Finally, the hierarchy demonstrated among the four tones in terms of accuracy in the current study was consistent with the literature. The differential difficulty among the tones might inspire those researchers interested in defining the motor complexity for the single speech unit (Maas, Robin, Wright, Austermann Hula, et al., 2008; Wright, Robin, et al., 2009). These researchers employed a single syllable with varying durations, short versus long, to test whether the syllable duration constitutes a parameter of motor complexity; the result of their studies did not support syllable duration as a reliable parameter of motor complexity. This finding suggests that the shifting of the fundamental frequency (F0) over a short duration might be a fruitful place to explore complexity at the syllabic level. It seems plausible that future research could test whether these tones would be differentiated in terms of preparation time before production measured as reaction time or study time.

11.0 CONCLUDING REMARKS

The aim of the current study was to assess the role of FOA on the learning of a novel speech task. Gaining this knowledge might further our understanding about those overlooked factors that might optimize learning in the speech domain. The applicability of the FOA concept is feasible in any speech rehabilitation setting in which individuals for whom speech is compromised by pathology attempt to learn speech. The FOA has the potential to be directly incorporated during speech therapy, either through the instructions or the feedback, both of which are considered integral elements in any treatment setting.

As this is the first study attempting to test the role of FOA in the speech domain, every effort was made to enhance the participants' learning and to allow any effects of FOA to be measured. As previously addressed, there was no clear conclusion of FOA effects on speech in the current study. The primary question about the effects of FOA on learning a speech task was not answered, either due to construction of the experiment or due to the decreased power that affected the detection of any difference among the groups. However, the findings of this study replicated the findings in the literature regarding the hierarchy of the four tone productions in terms of accuracy (Guo & Tao. 2008; Wang, Jongman, & Sereno, 2003).

Despite the inconclusive results of this study on how FOA affects learning a speech task, the findings of this study can still benefit researchers. The strengths and limitations of the methods used in this study can provide a foundation for researchers to enrich their understanding of the role that FOA might play in the speech domain. Because instructions are ubiquitous in any clinical setting, researchers should test whether refining the words utilized in those instructions might affect the performance and learning. With additional research, clinicians may find ways to use FOA to improve treatment outcomes in speech rehabilitation.

With respect to the hierarchy of correctness of the four tones, researchers could investigate other potential dimensions at the syllable level in order to better understand the motor complexity in the speech unit. Because previous research indicated that the syllable duration might not be a reliable parameter for motor complexity, it would be interesting to compare the four tones in terms of the preparation time before production. Such knowledge might further our comprehension about speech complexity and programming.

APPENDIX A

TONE DISCRIMINATION TEST: TONE-PAIRS

Table 65 List of the tone pairs' frequencies in the tone discrimination test (Adapted from Bradshaw and McHenry, 2005)

Tone-pair	Same Frequency Pair (Hz)	Different Frequency Pair (Hz)
1		261-277
2		261-269
3		261-269
4		1046-1108
5		261-269
6		253-554
7		523-538
8	523-523	
9	1046-1046	
10		523-538
11		1046-1108
12	261-261	
13		523-538
14	523-523	
15		1046-1077
16	523-523	
17		1046-1108
18		1046-1077
19		253-554
20		261-277
21	1046-1046	
22	261-261	
23		1046-1077
24		261-277
25	261-261	
26		523-554
27	1046-1046	

APPENDIX B

SHORT PARTICIPANTS' HISTORY QUESTIONNAIRE

Subject #: ______________ Date:________________ Age: ____________

Race/Ethnicity (please circle one):

Caucasian African-American Hispanic Asian other ______________

1) Is American English your native language? No Yes

2) Do you speak or understand any of the following tonal languages : Chinese languages, Cantonese, Thai, Lao, Vietnamese language, Kru languages, Khoisan languages?
   No Yes
   If yes, please specify,____________________________

3) What is your highest level of education? _____________

4) Have you had any kind of vocal training?
   No Yes
   If Yes, how many sessions,______________________

5) Have you had any kind of musical training?
   No Yes
   If Yes, how many years of musical training,______________________

6) Have you been diagnosed with a learning deficit?

No Yes

If Yes, please specify,_____________________

7) Are you registered at the office of disability resources and services at the University of Pittsburgh?

No Yes

If Yes, please provide the reason,_____________________

APPENDIX C

SUMMARY OF SELECTED ARTICLES FOR FOA INSTRUCTIONS

Table 66 Summary of focus of attention instruction in selected studies

Authors	Task	Groups +instructions		
		IFOA	EFOA	Control
Wulf, Weigelt, Poulter & McNevin (2003, Experiment1) (both groups:keep ball in the center) Retention/ transfer (no bar)	Suprapostural task Stabilometer & Holding a wooden bar with ball in the bar	They were able to see the bar “focus their attention on their hands and try to keep them horizontal”	“Focus on the tube”	

Table 66 (continued)

Authors	Task	Groups + instructions		
		IFOA	EFOA	Control
Shea & Wulf (1999)- concurrent FB	Stabilometer	Try to keep your feet at the same height Lines on screen to be thought of as representing their feet	Try to keep the yellow lines in front of their feet at the same height Lines on screen to be thought of as representing the yellow lone in front of their feet	
Wulf, McConnel, Gartner, & Schwarz (2002, Exp. 1)	Volleyball – “tennis serve”	Feedback statements with reference to body parts which are performing the movement (e.g., shoulders, arms, legs)	Feedback statements with reference to movement effect (e.g., shift body weight)	
Wulf, McConnel, Gartner, & Schwarz (2002, Exp. 2) Effect of relative FB as a function of AF (4G) I vs E (100 % vs 33%)	Soccer task “lofted soccer pass”	Feedback statement referred participants to their body movements	Feedback statement worded to direct attention to the movement effect	
McNevin, & Wulf (2002) Within participant design	Suprapostural task Standing still Facing a hanged sheet	“Try to minimize the movement of the index finger over the duration of the trial”	“Try to minimize the movement of the sheet over the duration of the trial”	Stand still no sheet
Wulf, McNevin, and Shea (2001) Dual task	Stabilometer	Focus attention on feet and keep them horizontal	Focus on the markers attached to the platform (22 cm from participants feet).	

Table 66 (continued)

Authors	Task	Groups +instructions		
		IFOA	EFOA	Control
Wulf, Weigelt, Poulter & McNevin (2003, Experiment1) (both groups:keep ball in the center) Retention/ transfer (no bar)	Suprapostural task Stabilometer & Holding a wooden bar with ball in the bar	They were able to see the bar "focus their attention on their hands and try to keep them horizontal"	"Focus on the tube"	
Wulf, Weigelt, Poulter & McNevin (2003, Experiment2)	Suprapostural task Stabilometer & Holding a wooden bar without the ball in the bar+ instruction before second trial	Focused their attention on their hands and try to keep them horizontal	"Focus on the tube"	
Wulf, Mercer, McNevin & Guadagnoli (2004) Within participant design	Suprapostural task inflated disk holding a pole	Minimize the movement of their feet Holding their hands still	Minimize the movement of the disc Holding the pole still	
Vuillerme & Nafati (2007)	Quiet standing	Focus attention on their body sway and to increase their active intervention into postural control.		Participants stood upright without specific instruction concerning their attentional focus of attention.
Zachry, Wulf, Mercer,& Bezodis (2005) + EMG	Basketball task	Concentrate on the snapping motion of the wrist during the follow-through of the free throw shot.	Concentrate on the canter of the rear of the basketball hoop.	
McNevin, Shea, and Wulf, (2003) Distance of EFOA	Stabilometer	Focus attention on feet (as their control group)	Focus on the markers (near and inside/outside far)	

Table 66 (continued)

Authors	Task	Groups +instructions		
		IFOA	EFOA	Control
Wulf & Su (2007, Exp. 1)	Golf	Instruction directed at the swinging motion of their arms	The attention directed towards the pendulum-like motion of the club.	No attentional focus instructions
Wulf & Su (2007, Exp. 2) - Within participant design	Golf- experts	Focus on arm motion	Focus on club motion	What they normally focus on
Vance, Wulf, Tollner, McNevin, & Mercer (2004) +EMG	Biceps curls	Concentrate on biceps muscle	Concentrate on curl bar	
Wulf, Landers, Lewthwaite, & Töllner (2009) Within participant design	balance on inflated disc (Parkinson's disease)	To reduce movement of the feet	To reduce movement of the disc	
McAlister, 2006	Simulated occupation therapy rehabilitation setting (using prosthesis)	"While completing the following task, pay attention to your shoulder, elbow, and wrist as you move the prosthesis during each part of the task…."	"While completing the following task, pay attention to the cup, bowl, and plate during each part of the task…"	
Wulf, Mercer, McNevin & Guadagnoli (2004) Within participant design	Suprapostural task inflated disk holding a pole	Minimize the movement of their feet Holding their hands still	Minimize the movement of the disc Holding the pole still	

Table 66 (continued)

Authors	Task	Groups +instructions		
		IFOA	EFOA	Control
Freedman,Maas, Caligiuri, Wulf, & Robin (2007)	Exerting a rapid pressure on a rubber bulb by the hand /or by the tongue when the bulb was placed in mouth.	Keep focusing on your tongue/hand, focus on tongue/hand. Push with your tongue/ hand	Keep focusing on the bulb, focus on the bulb. Push on the bulb	
Lohse, Sherwood, and Healy (2010)	Dart throwing task	Focus on the movement of your arm	Focus on the flight of the dart	
Porter, Nolan, Ostrowski, and Wulf, *(2010)*	Agility task: running a course of two five-meter long paths "L" shape.	To focus on moving their legs as fast as possible and to focus on planting their feet with maximum effort.	To focus on running towards the cones as fast as possible and to focus on pushing the ground with maximum effort	Only general instruction to run the path with maximum speed and effort
Fasoli, Trombly, Tickle-Degnen, Verfaellie, (2002)	Three functional reaching tasks commonly used in occupation therapy: 1) "removing a can from a shelf and placing it on the table"; 2) " taking an apple off a shelf and putting it into a basket; and 3) moving an empty coffee mug from the table onto a saucer"	The researchers instructed the participants to pay attention to their arms and to think about how their elbow straightens and how their wrist and fingers move.	The researchers instructed the participants to pay attention to the can and it's position on the shelf and to think about how big and heavy the can is, (or the mug or the apple)	

APPENDIX D

MANIPULATION CHECK QUESTIONNAIRE

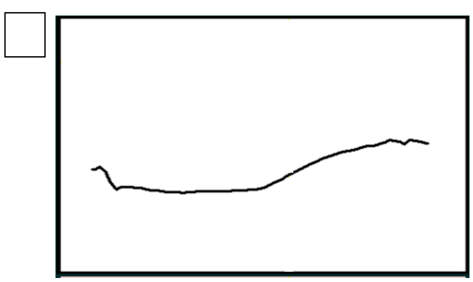
(Adapted from Porter, Nolan, Ostrowski, & Wulf, 2010; Fasoli, Trombly, Tickle-Degnen, & Verfaellie, 2002)

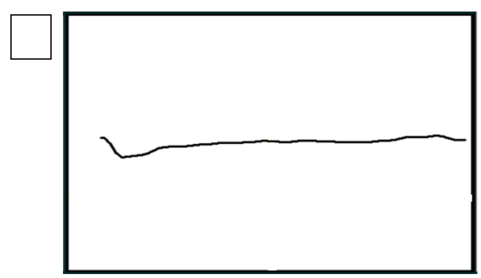
Please answer the following questions by using the scale:

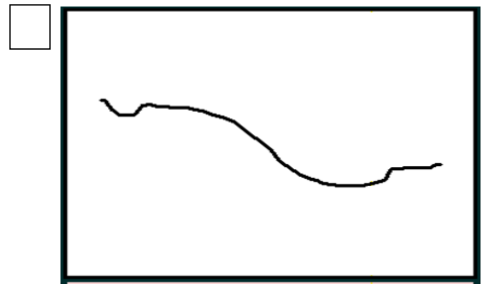
1. How do you rate the difficulty of the speech task?

1	2	3	4	5
Extremely Easy				Extremely difficult

2. Rank the four tone contours based on your perceived difficulty by assigning numbers (1, 2, 3, or 4) in the boxes next to the figure; 1 being the easiest and 4 being the most difficult and 2 or 3 to the next easiest and so on.







3. To what extent did you find the instructions helpful?

1	2	3	4	5
Very helpful		Somewhat helpful		Not at all helpful

4. Did you try to focus as instructed?

Yes No

5. Which aspect of the task did you focus on:

    1. Vibration in the larynx.
    2. Feedback.
    3. Sound I produced.
    4. Did not focus on anything in particular.
    5. Other things-please specify---------------------------------------

6. To what extent did you find the feedback helpful?

1	2	3	4	5
Very helpful		Somewhat helpful		Not at all helpful

7. Did you focus on the same aspect of the task throughout the training, as instructed?

1	2	3	4	5
None of the time	A little of the time	Some of the time	Most of the time	All of the time

6. Do you think you learned all four tones equally well?

Yes No

APPENDIX E

SINGLE SUBJECT DATA: RMSE FOR THE PRACTICE WORDS DURING BASELINE, ACQUISITION PHASE, AND RETENTION TEST

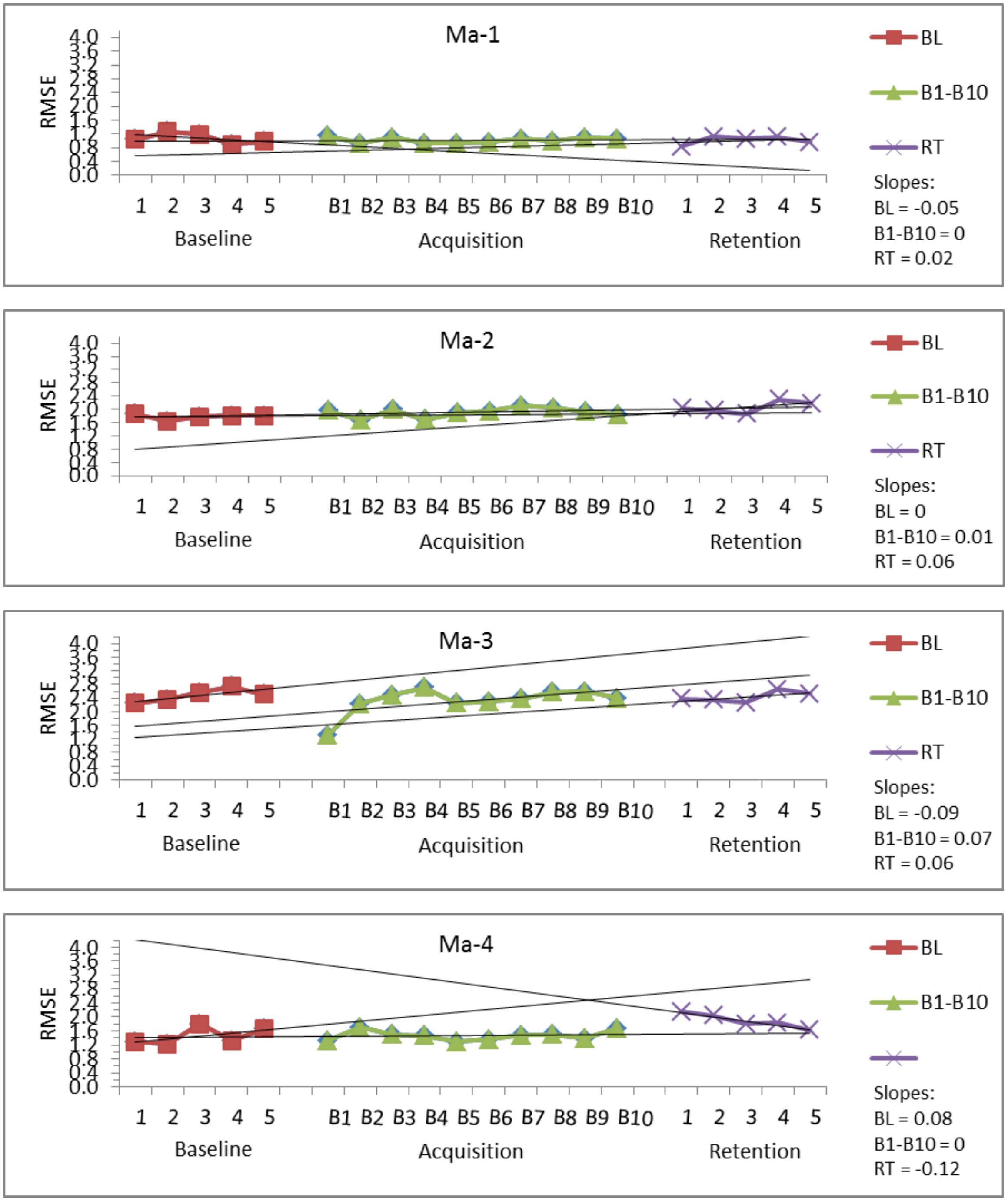


Figure 15 EFOA-Participant-6 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

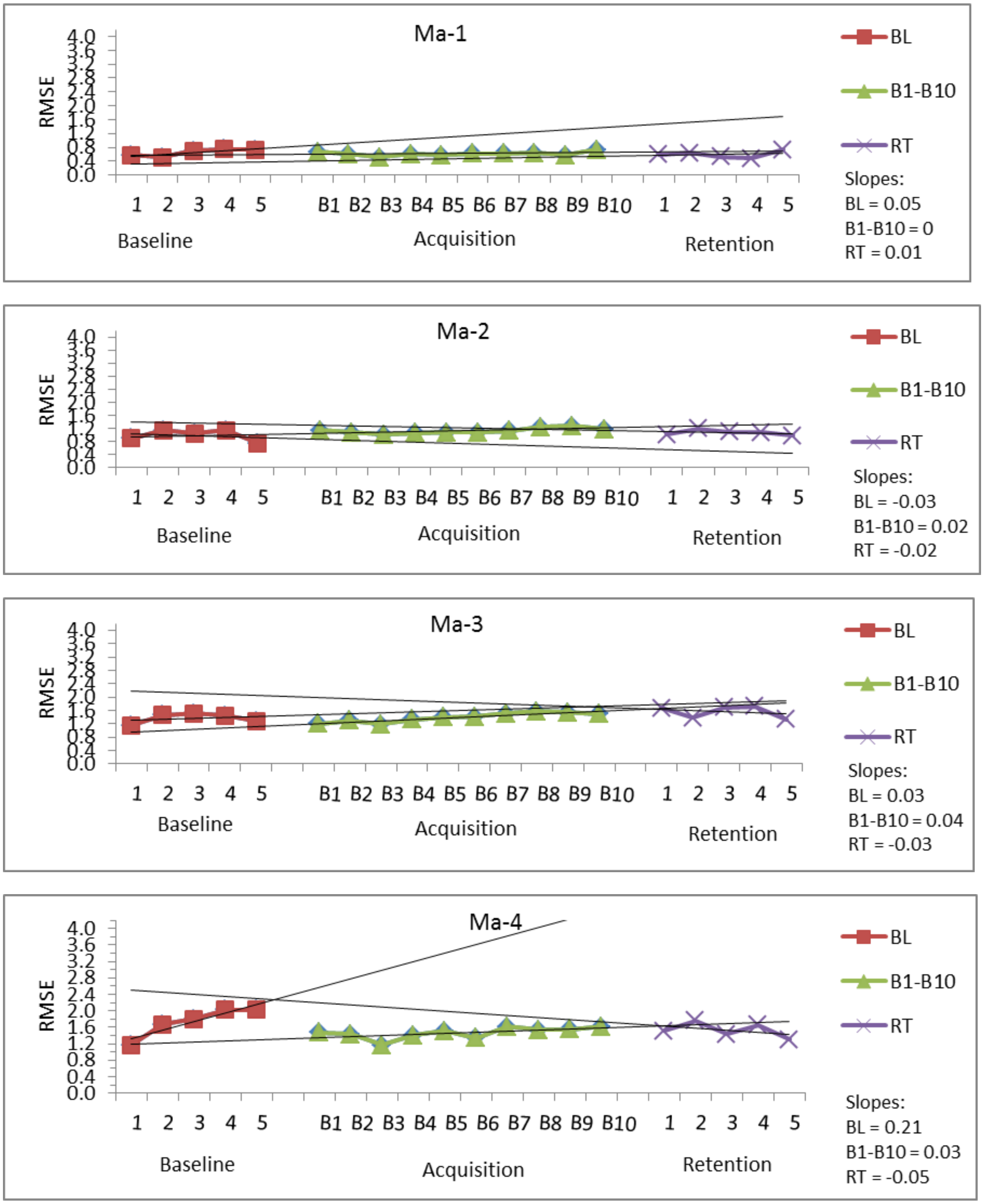


Figure 16 EFOA Participant-10 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

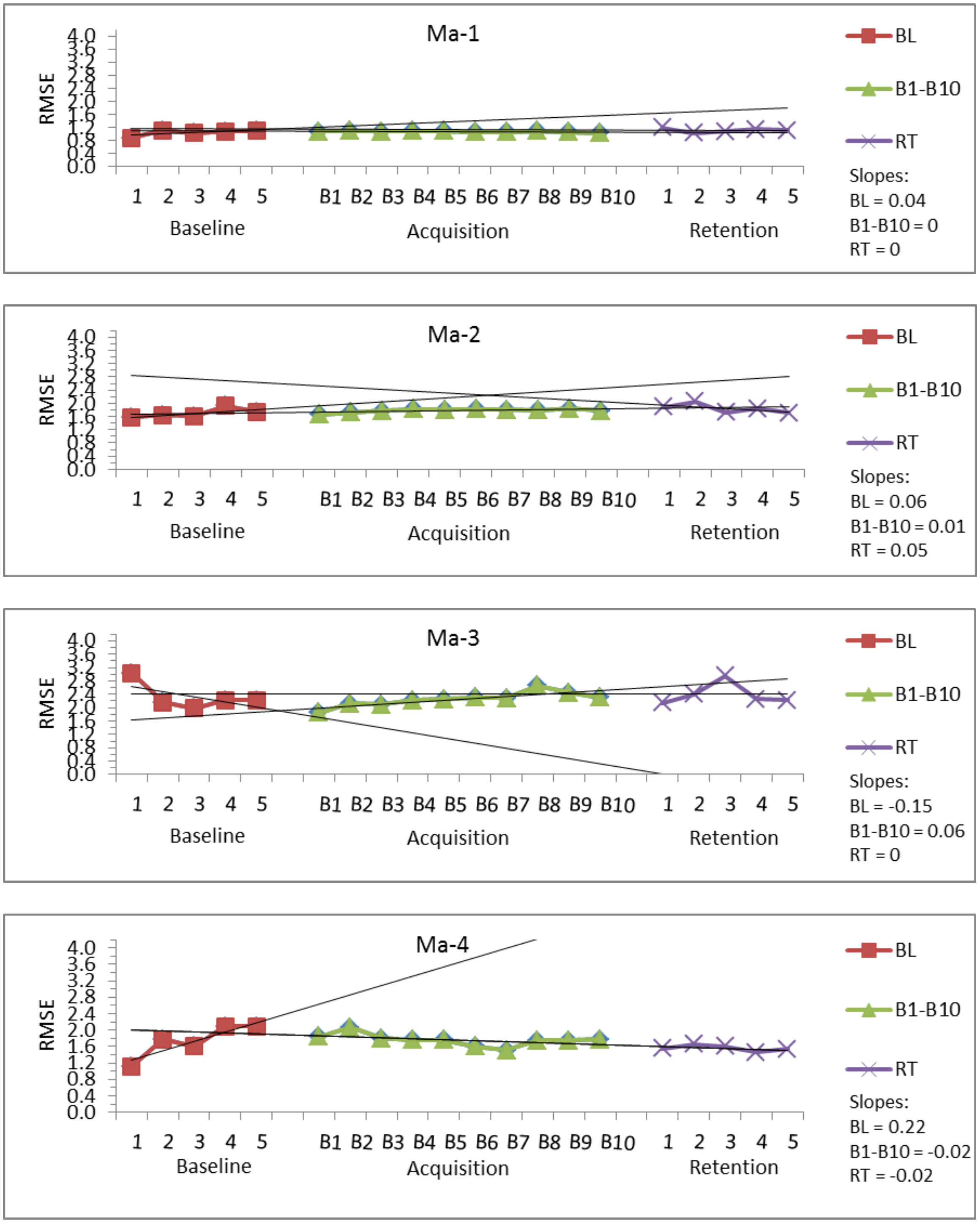


Figure 17 EFOA Participant-13 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

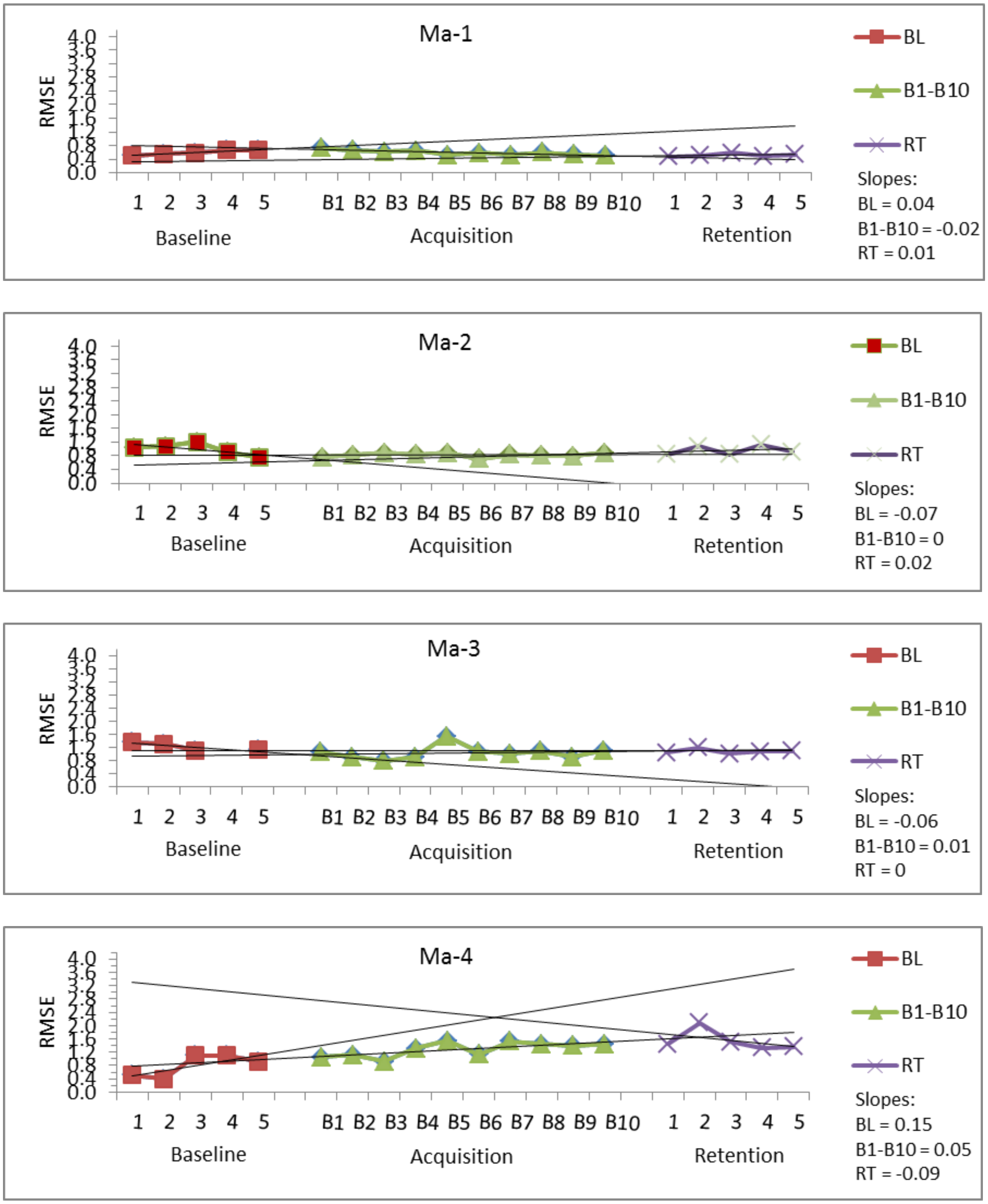


Figure 18 EFOA Participant-20 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

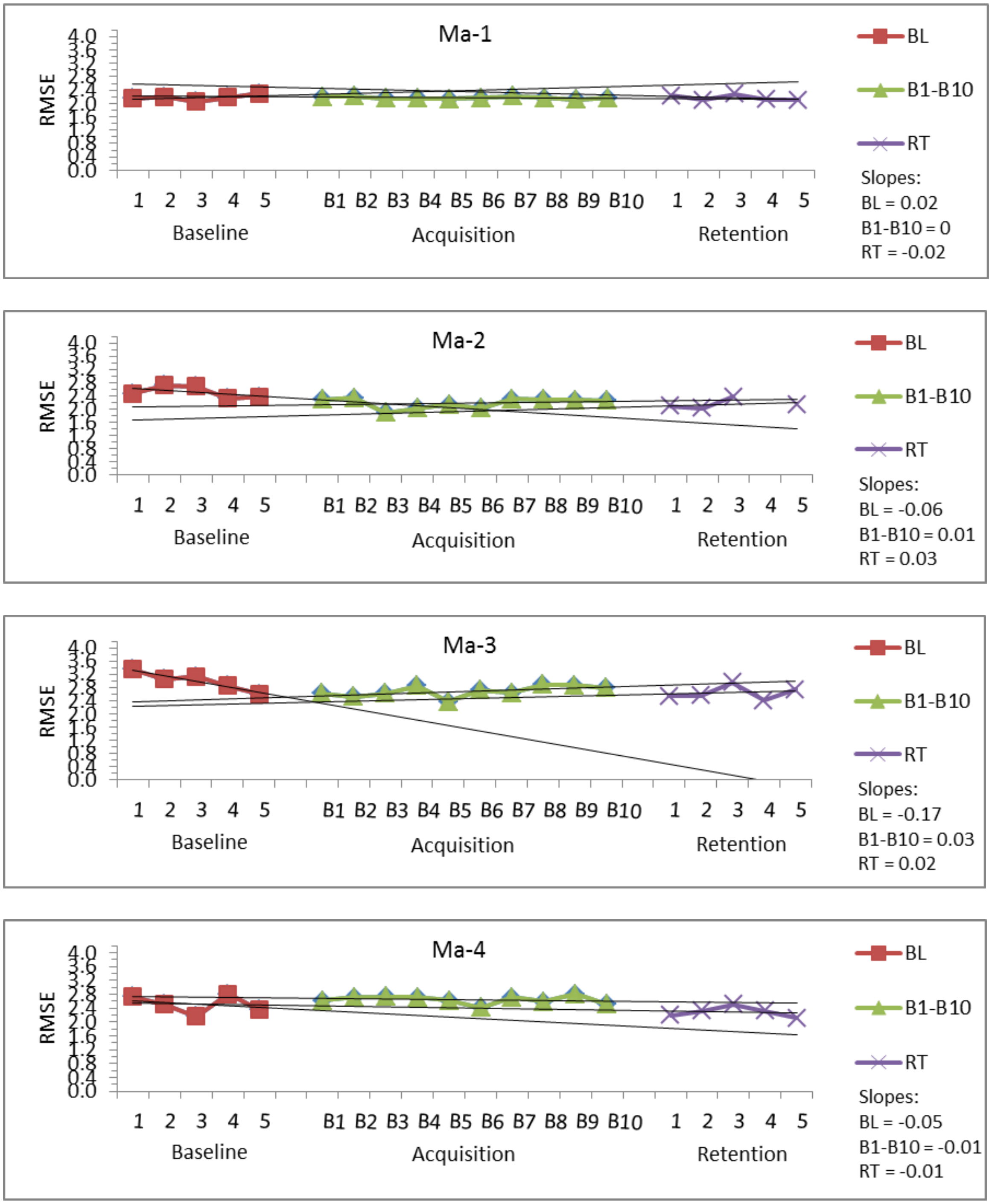


Figure 19 EFOA Participant-22 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

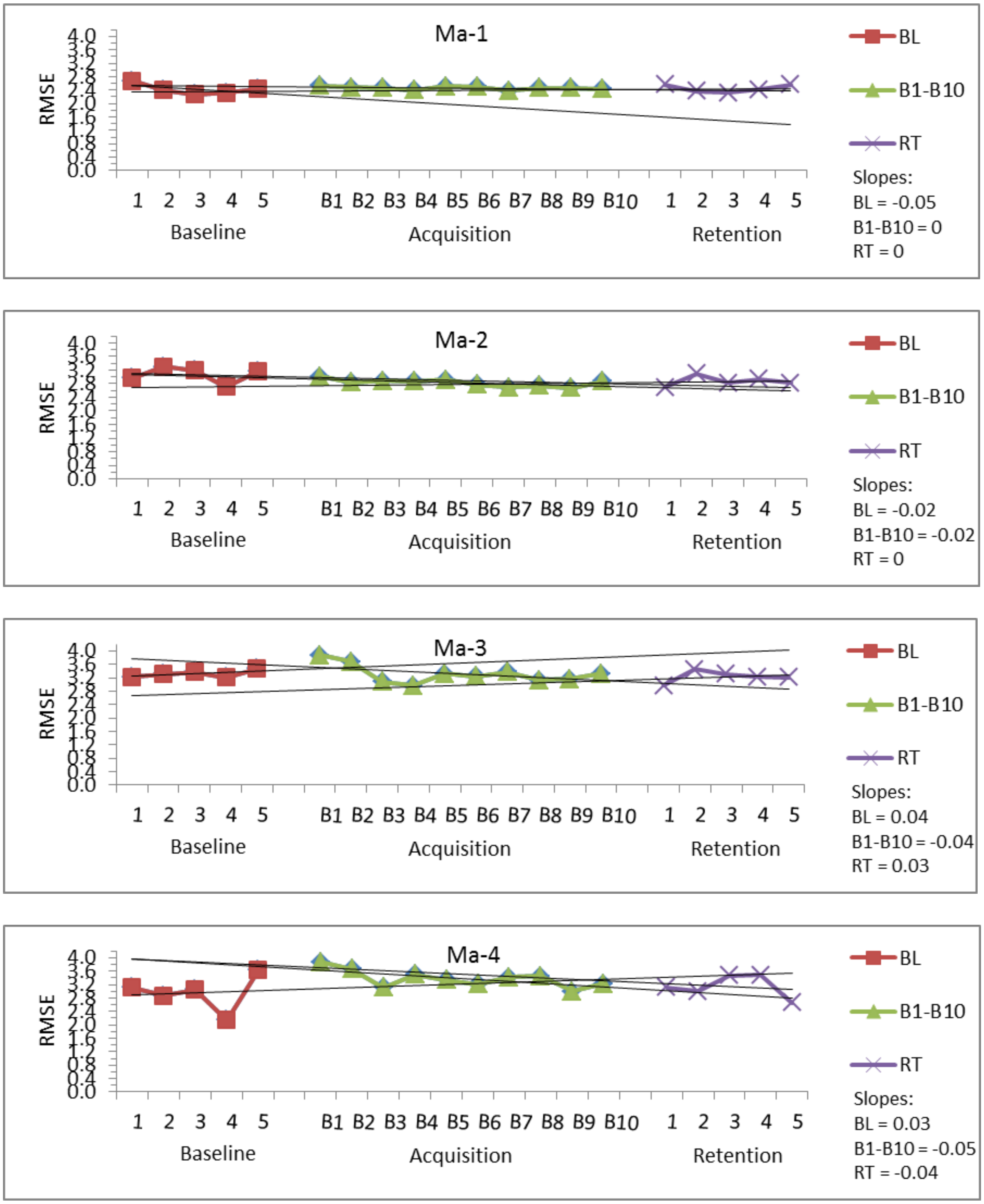


Figure 20 EFOA Participant-29 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

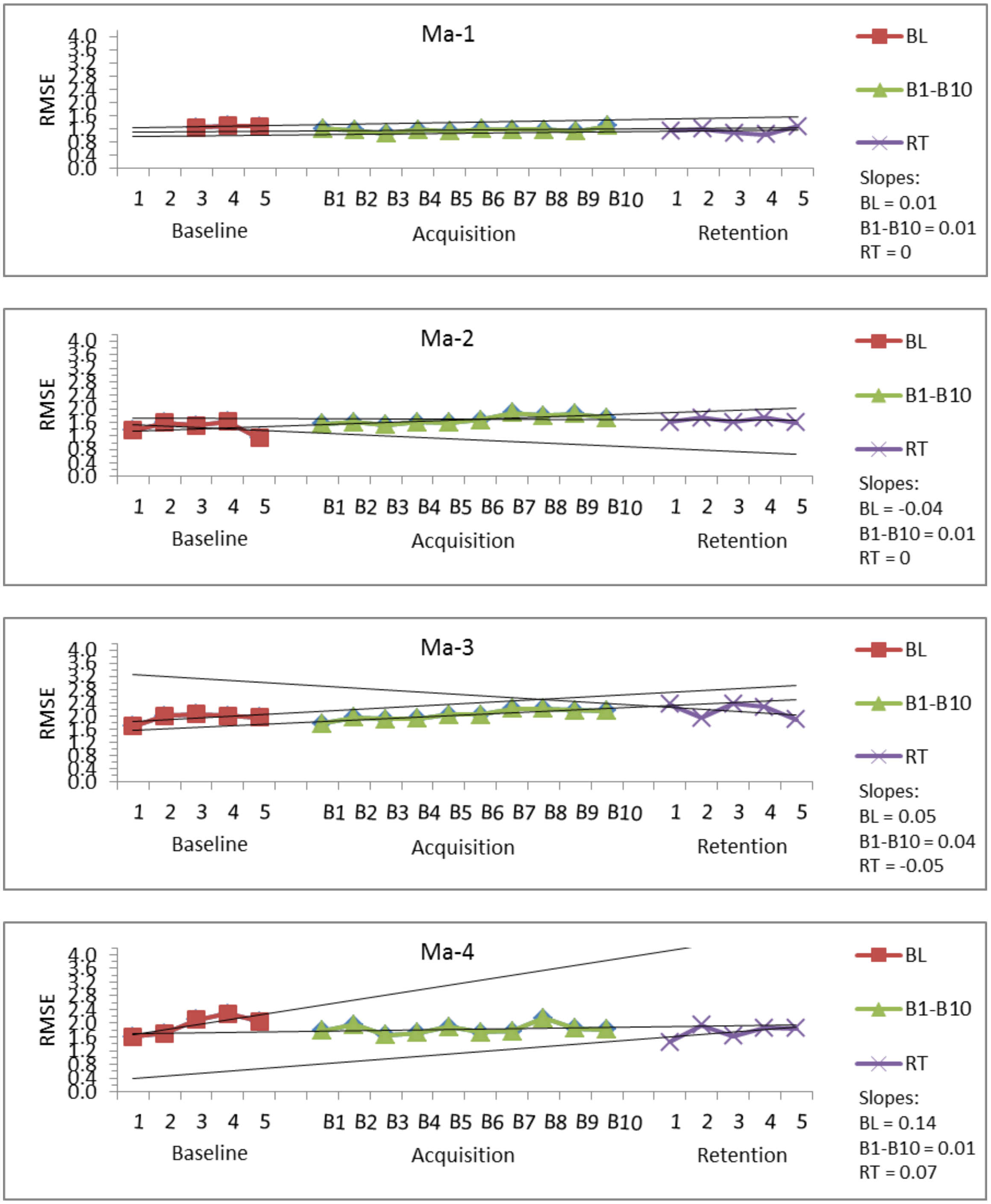


Figure 21 EFOA Participant-32 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

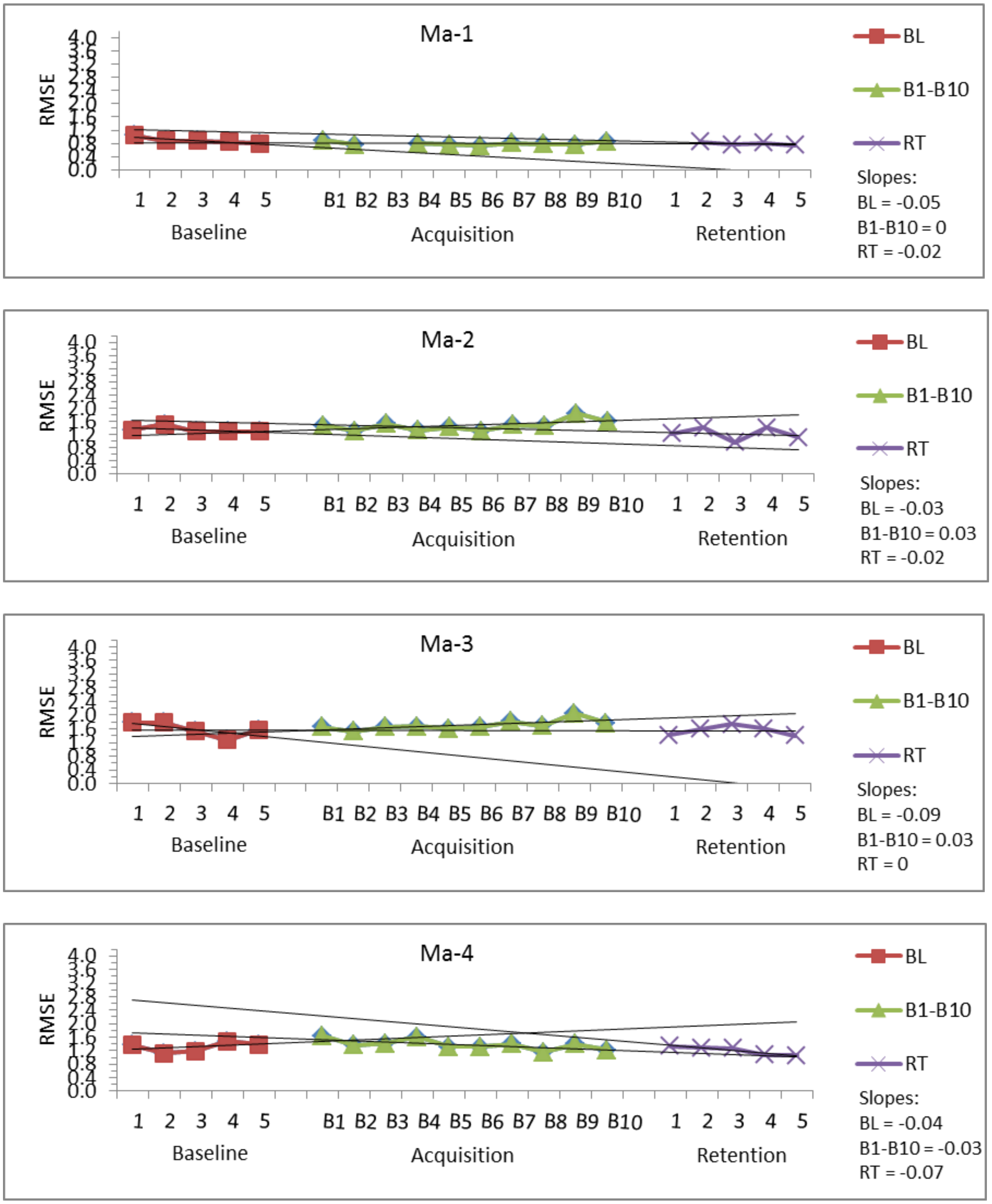


Figure 22 EFOA Participant-33 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

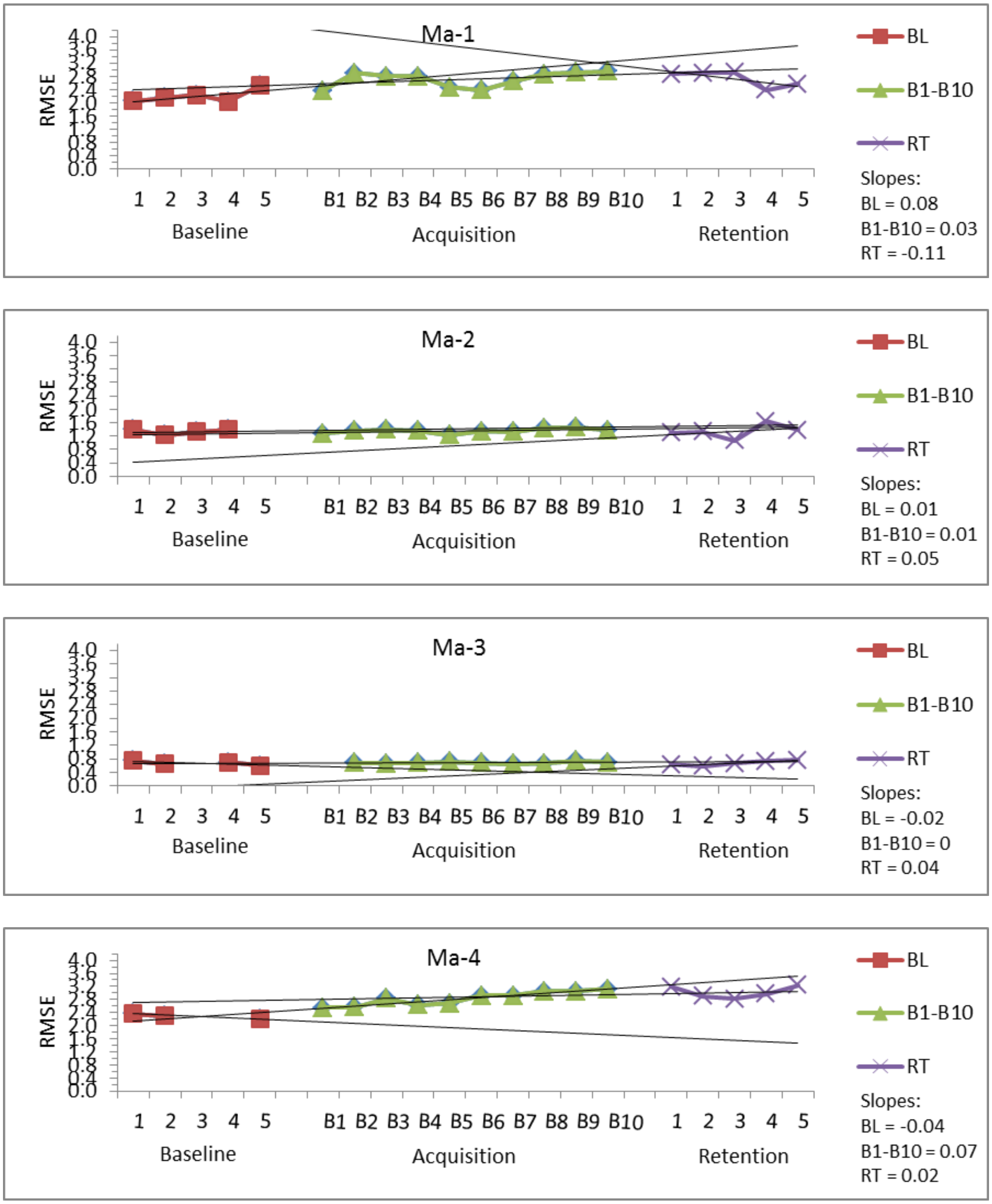


Figure 23 EFOA Participant-38 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

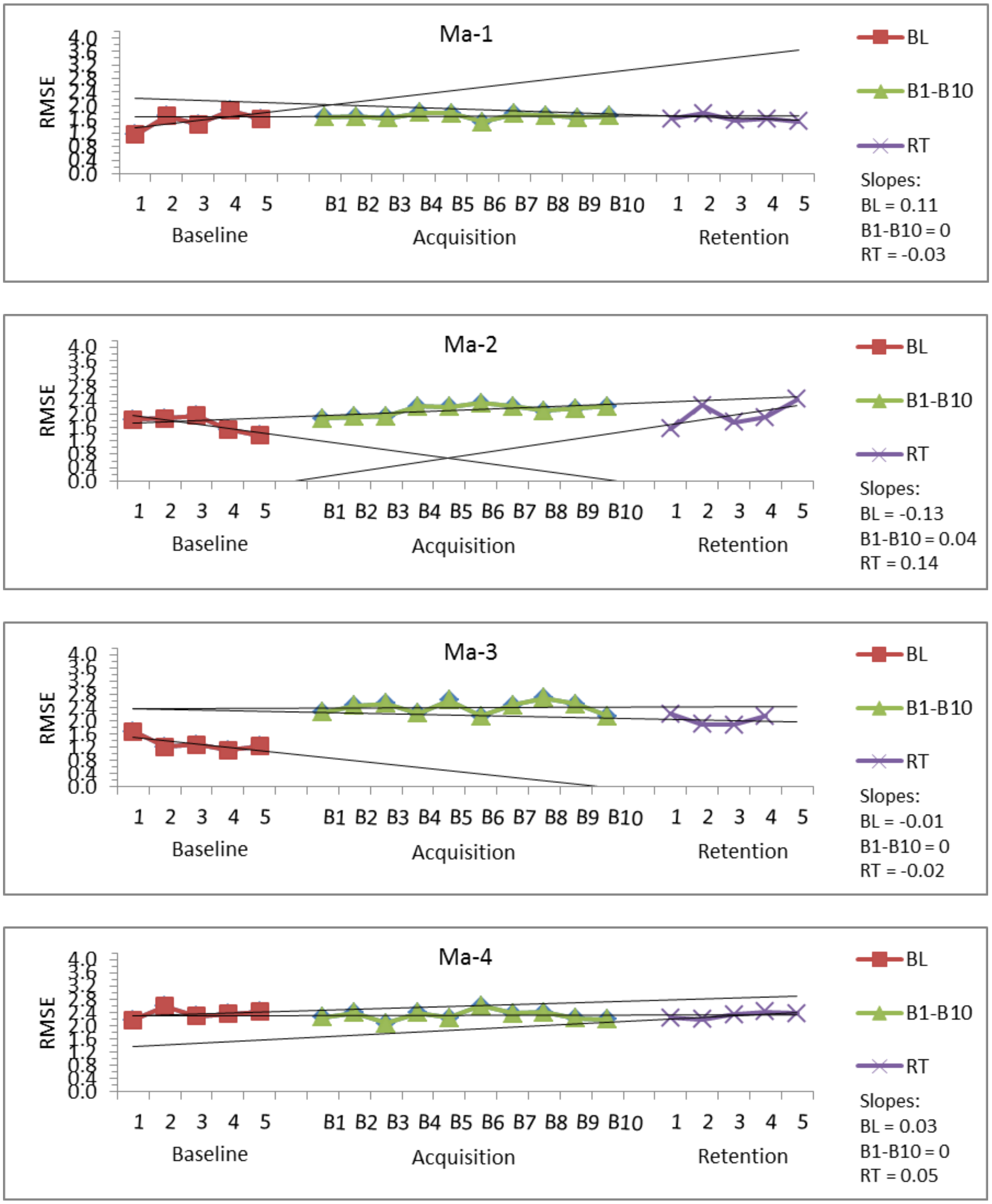


Figure 24 EFOA Participant-42 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

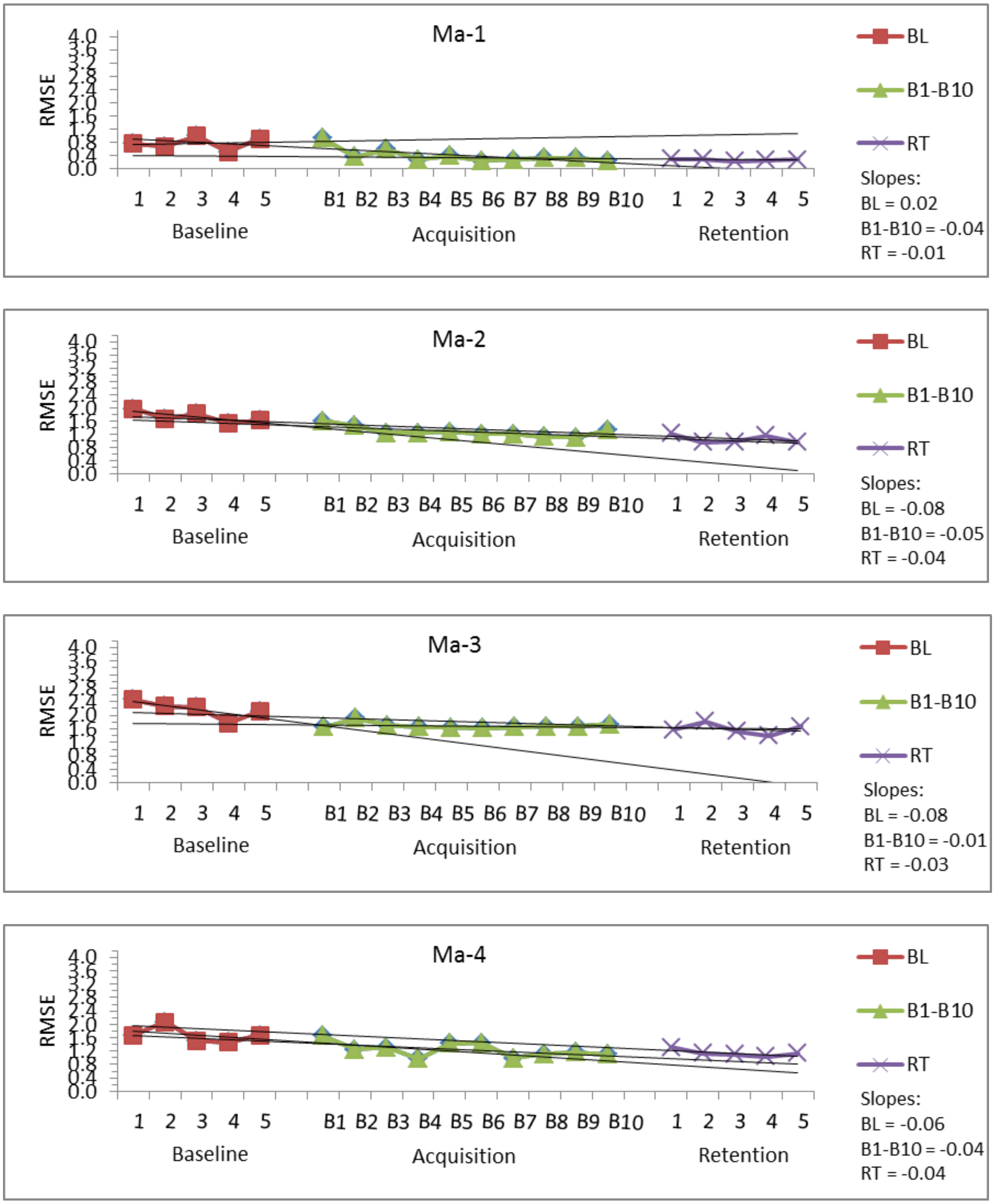


Figure 25 EFOA Participant-43 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

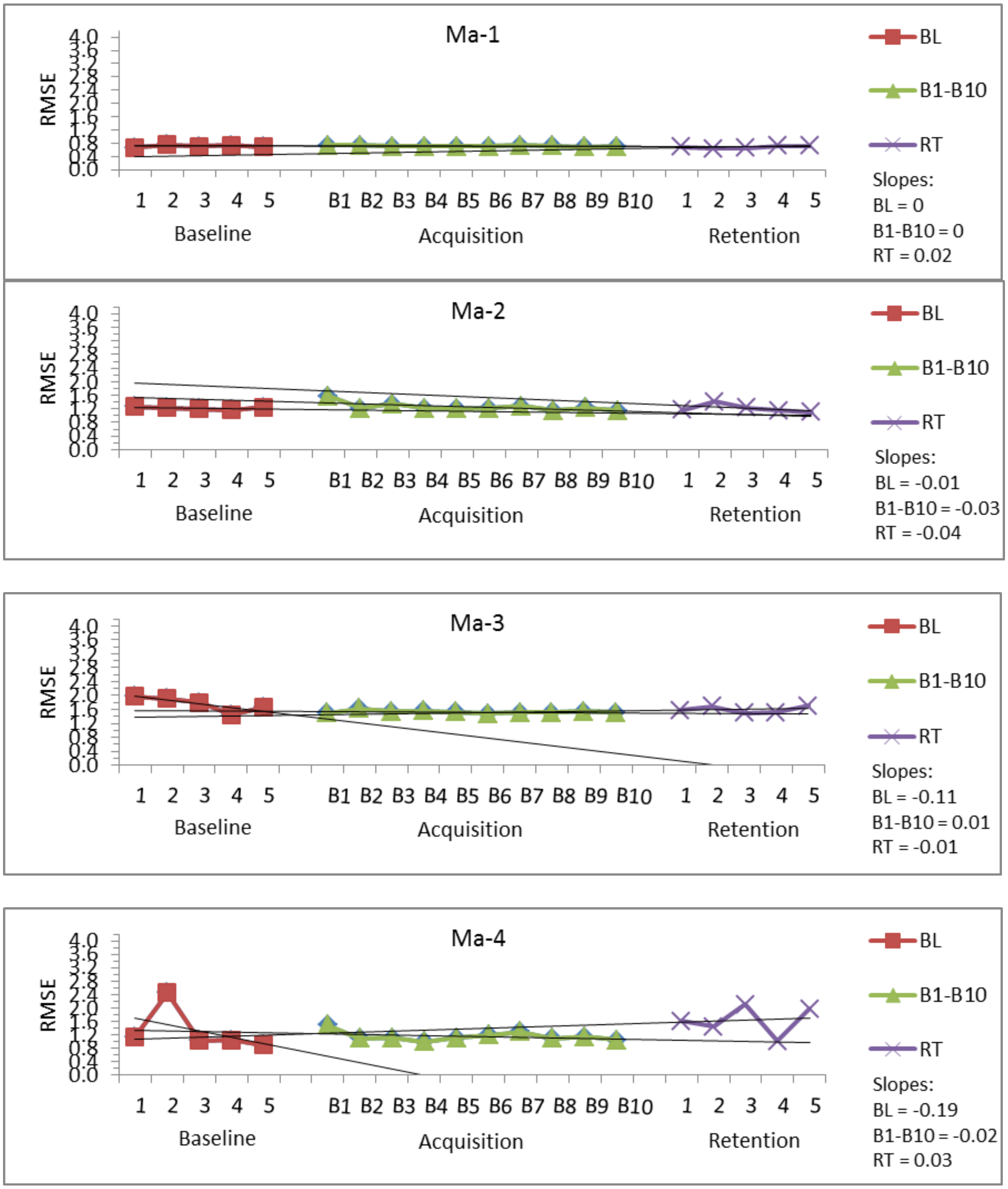


Figure 26 EFOA Participant-46 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

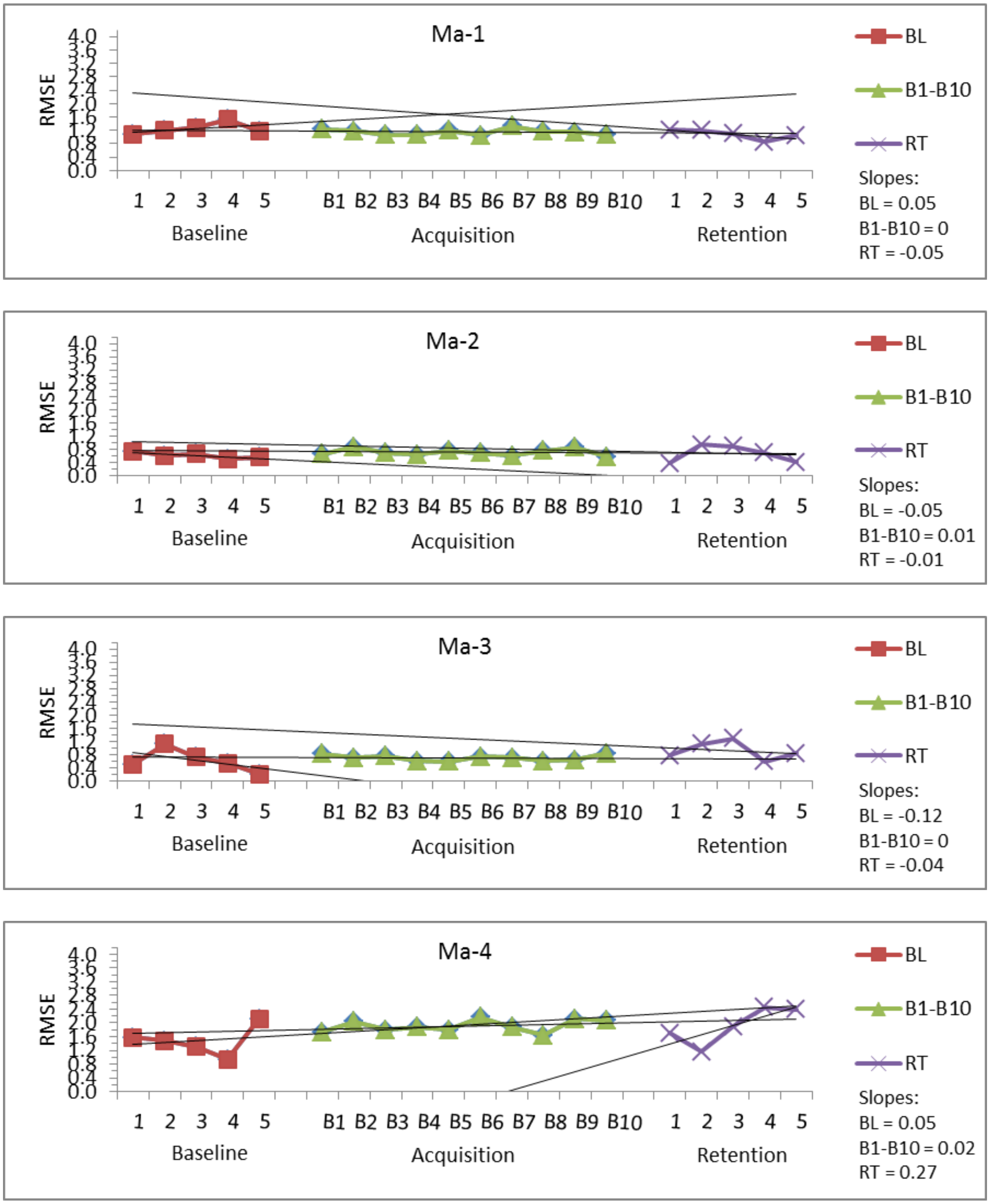


Figure 27 EFOA Participant-47 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

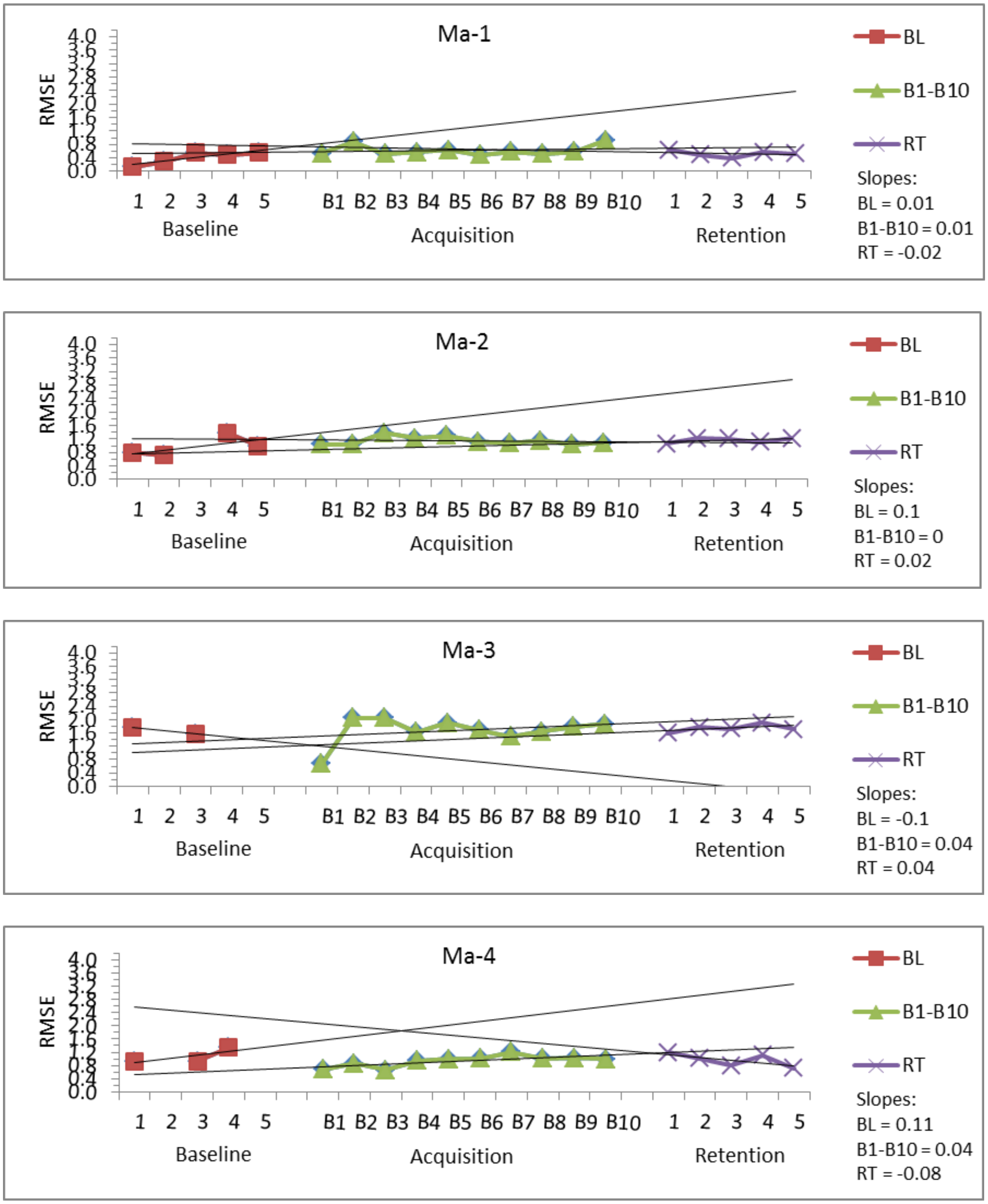


Figure 28 EFOA Participant-55 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. EFOA = external focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

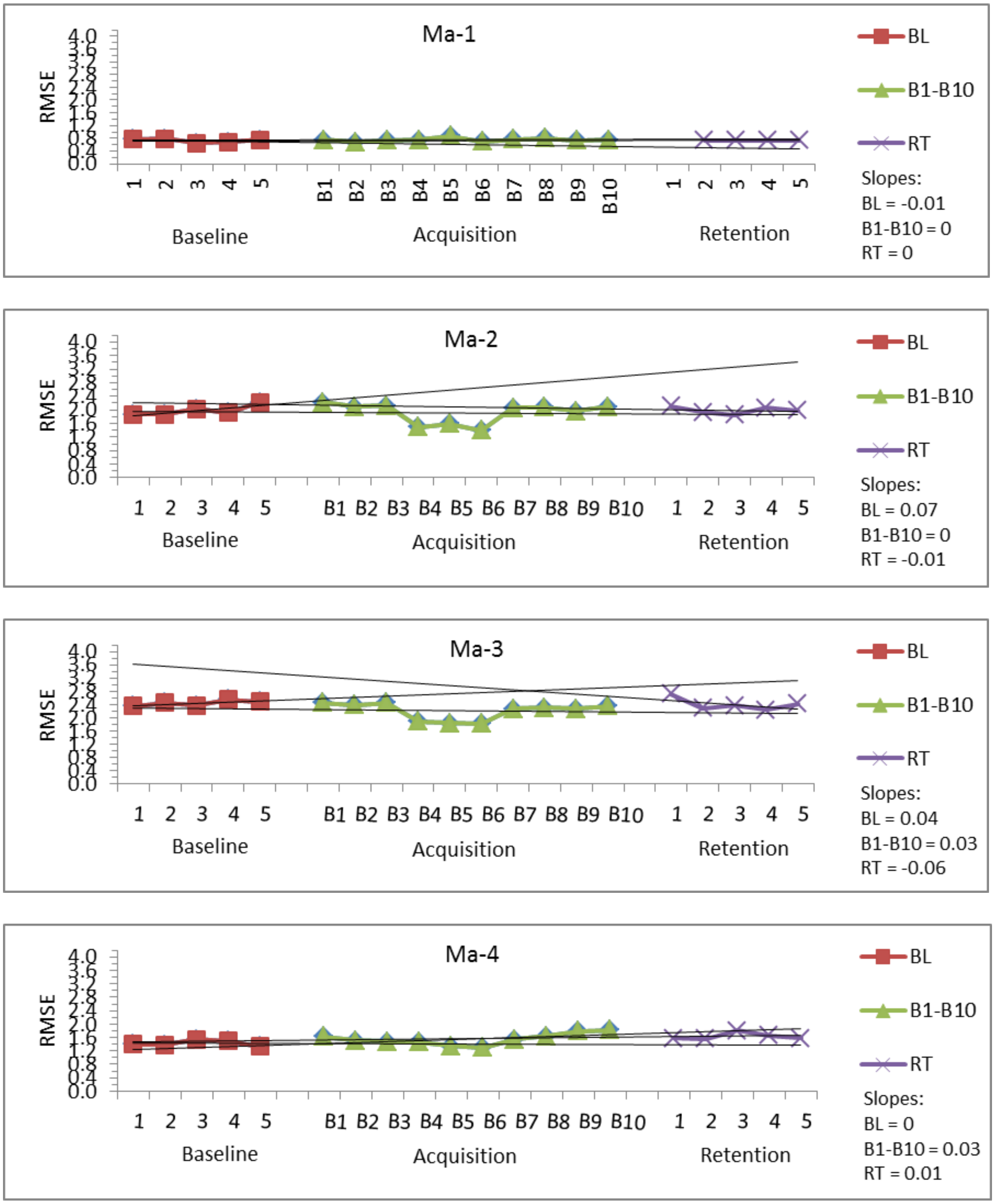


Figure 29 IFOA Participant-2 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

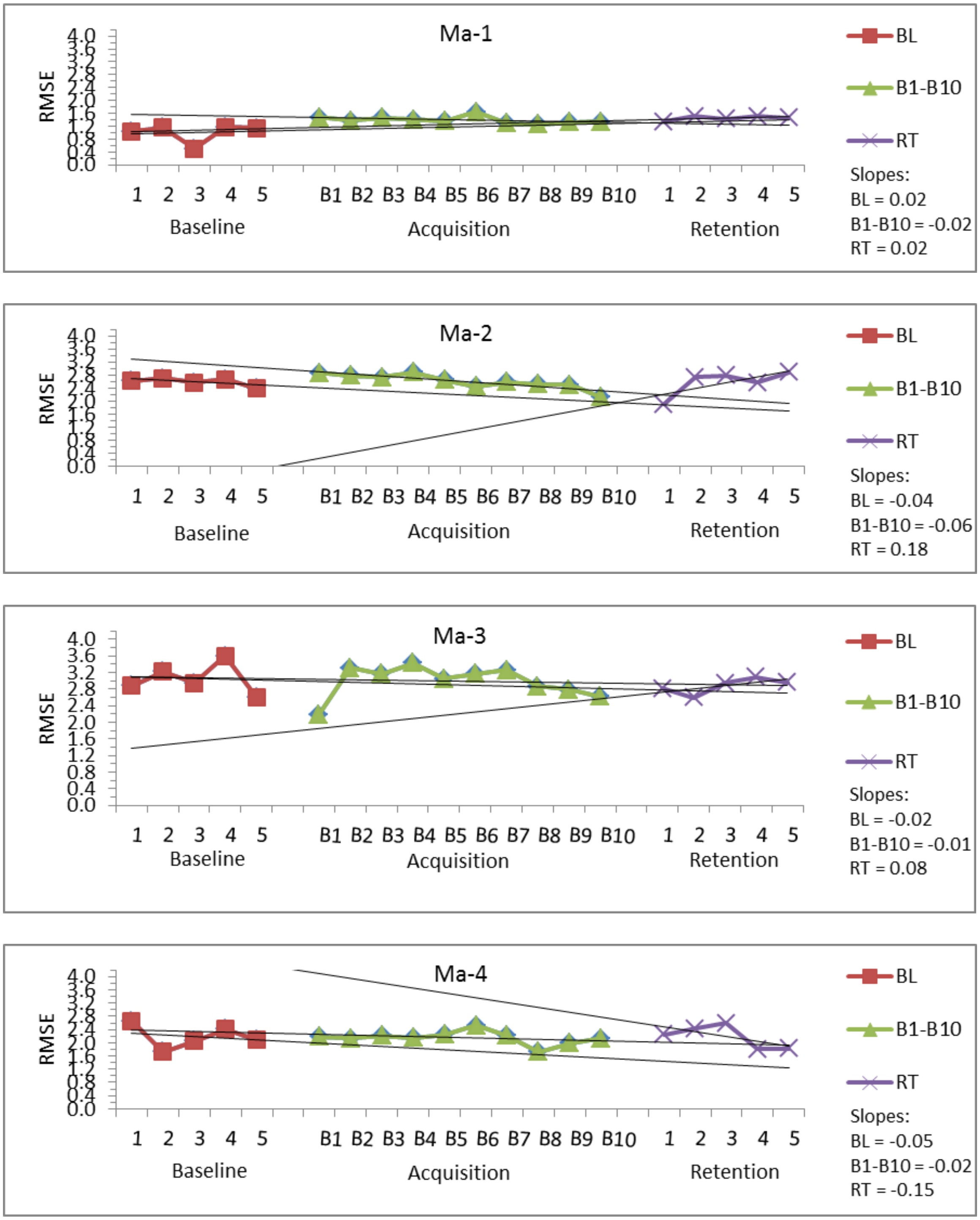


Figure 30 IFOA Participant-5 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

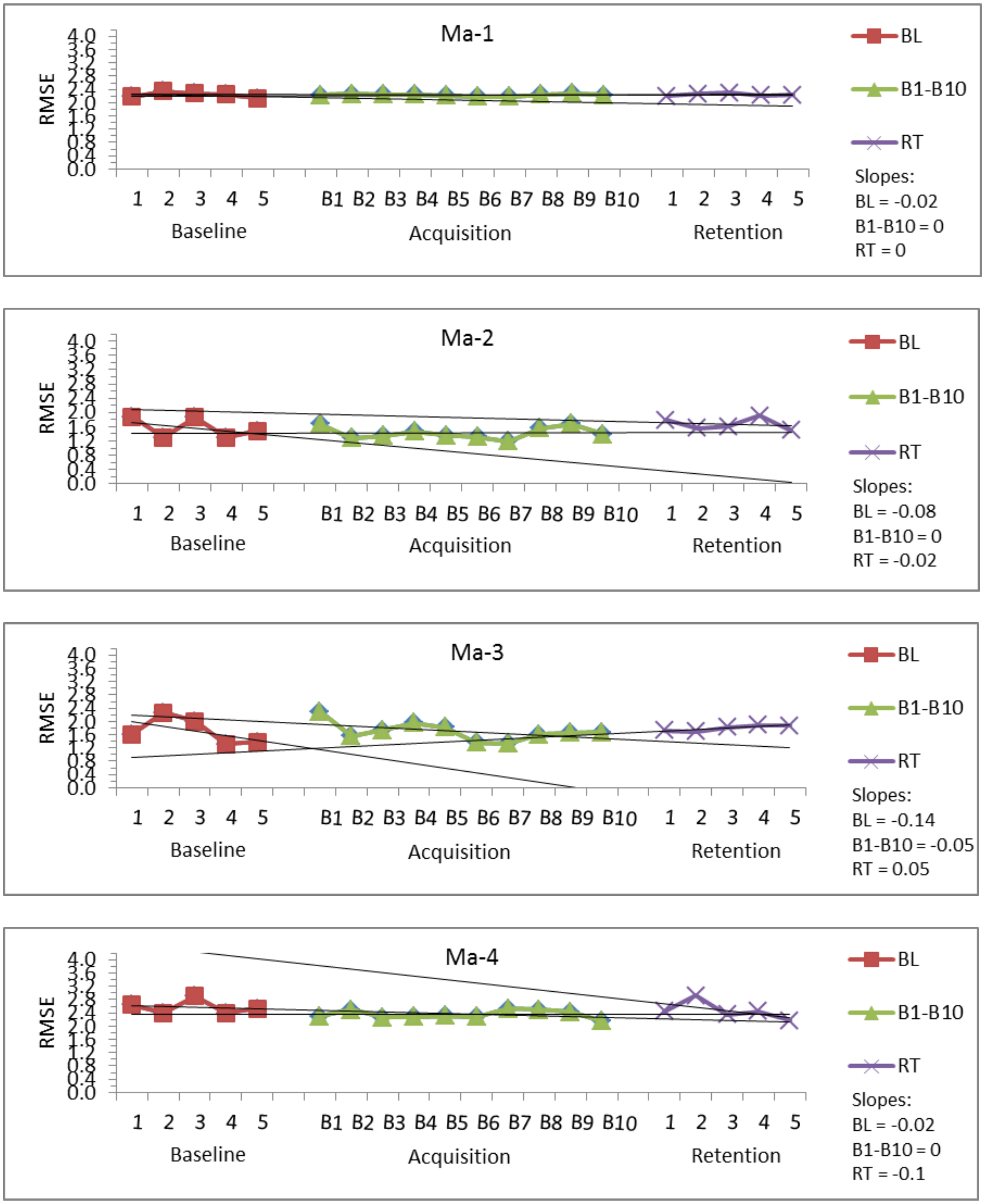


Figure 31 IFOA Participant-7 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

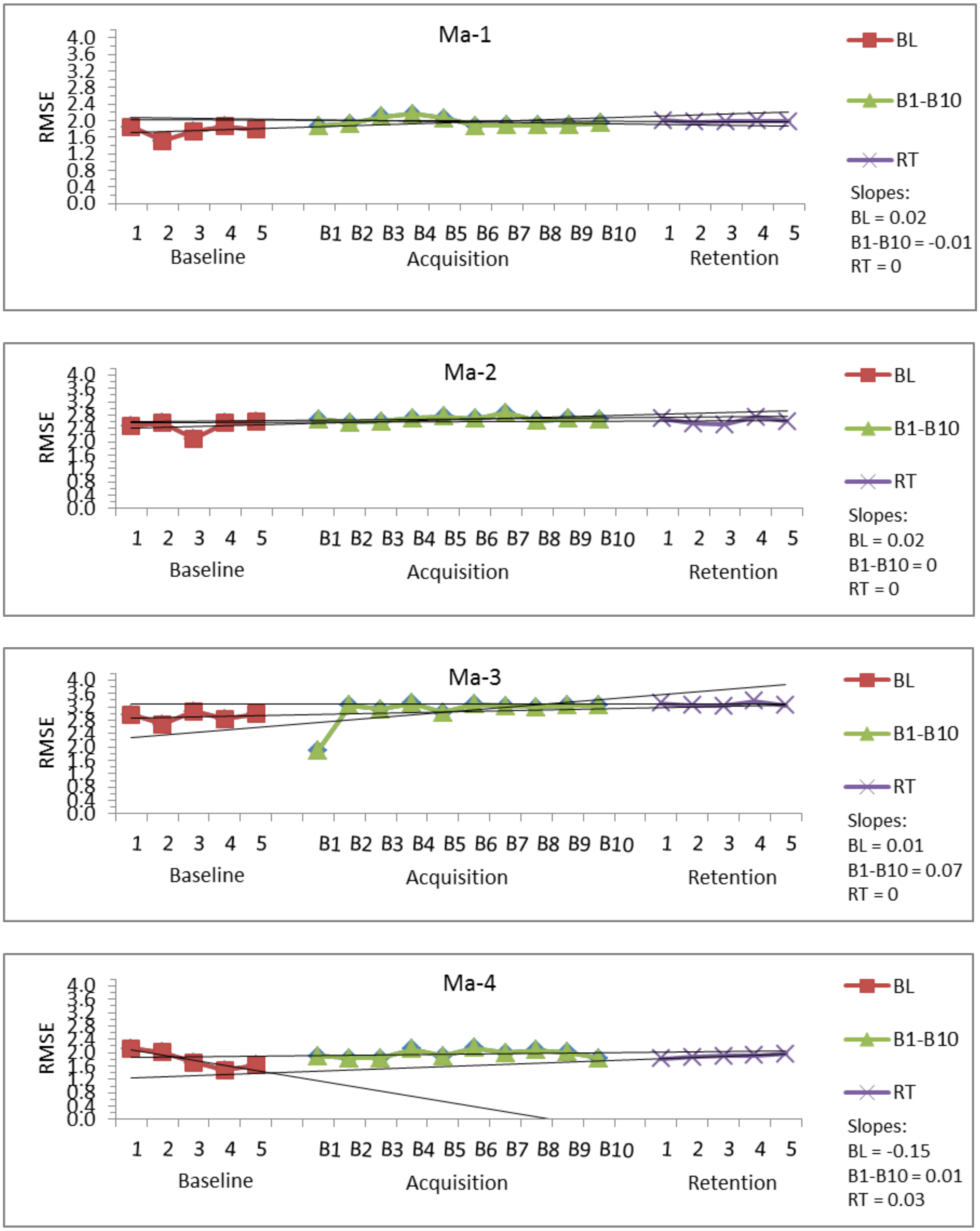
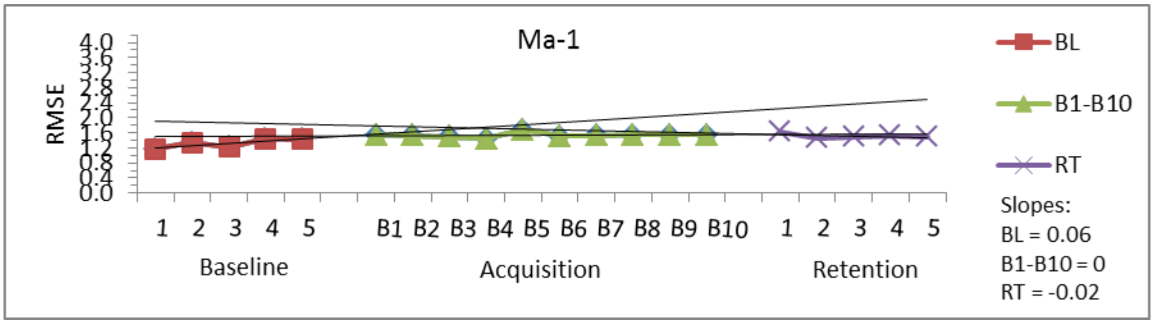
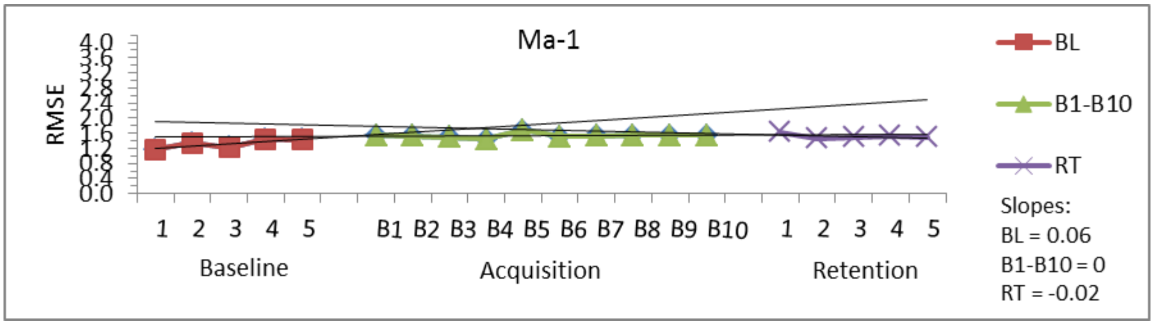
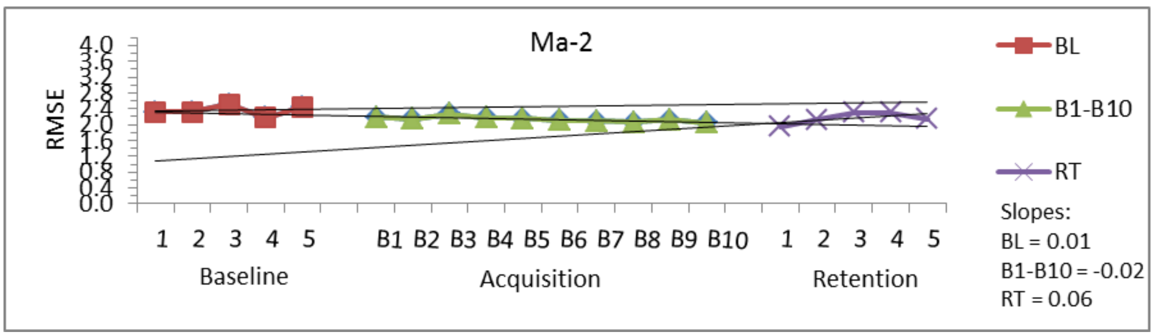


Figure 32 IFOA Participant-11 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.







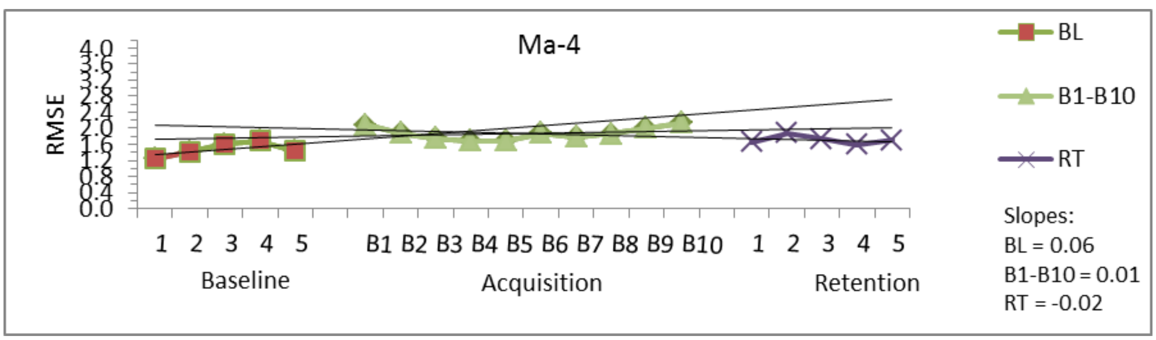


Figure 33 IFOA Participant-14 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

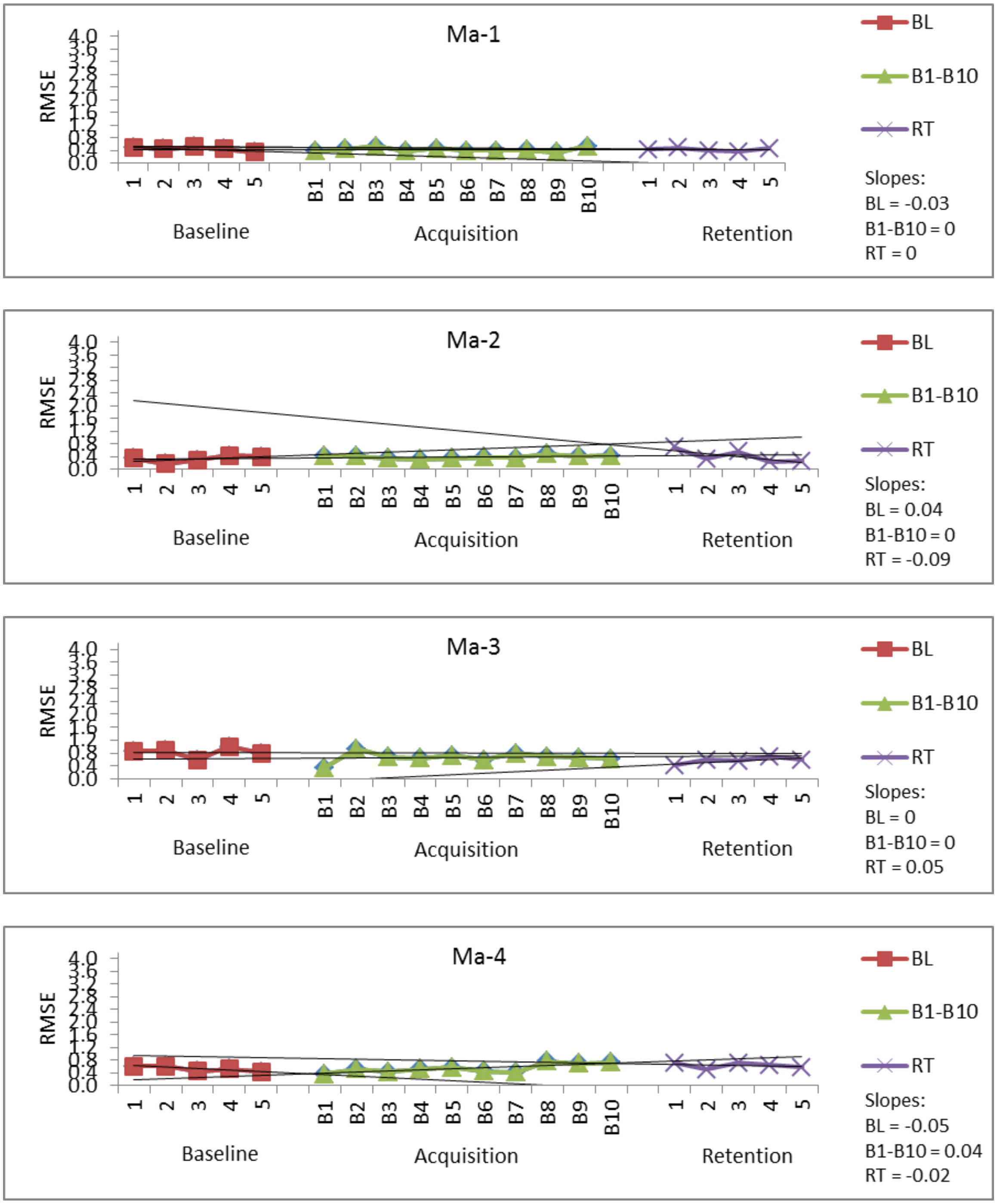


Figure 34 IFOA Participant-24 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

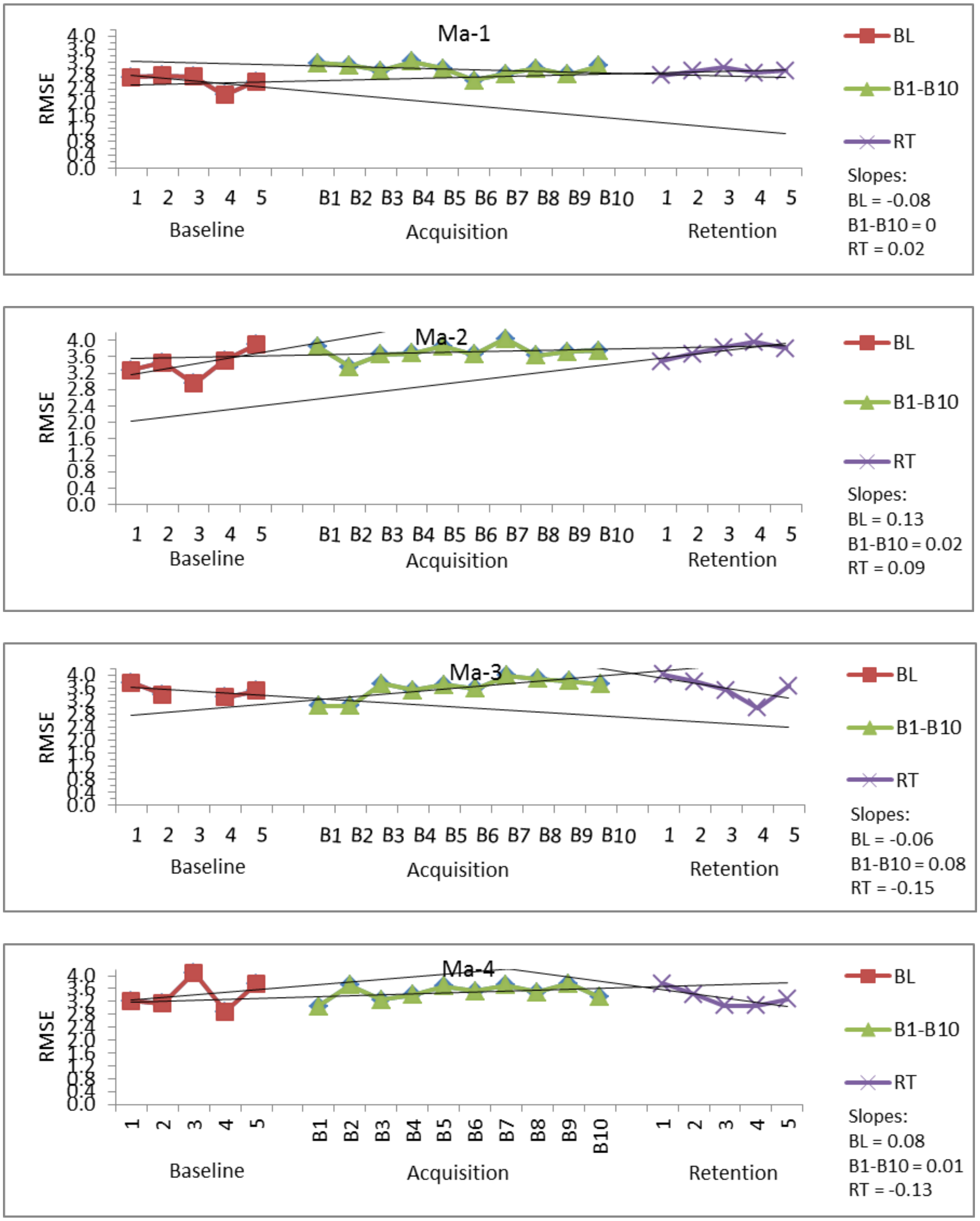


Figure 35 IFOA Participant-26 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

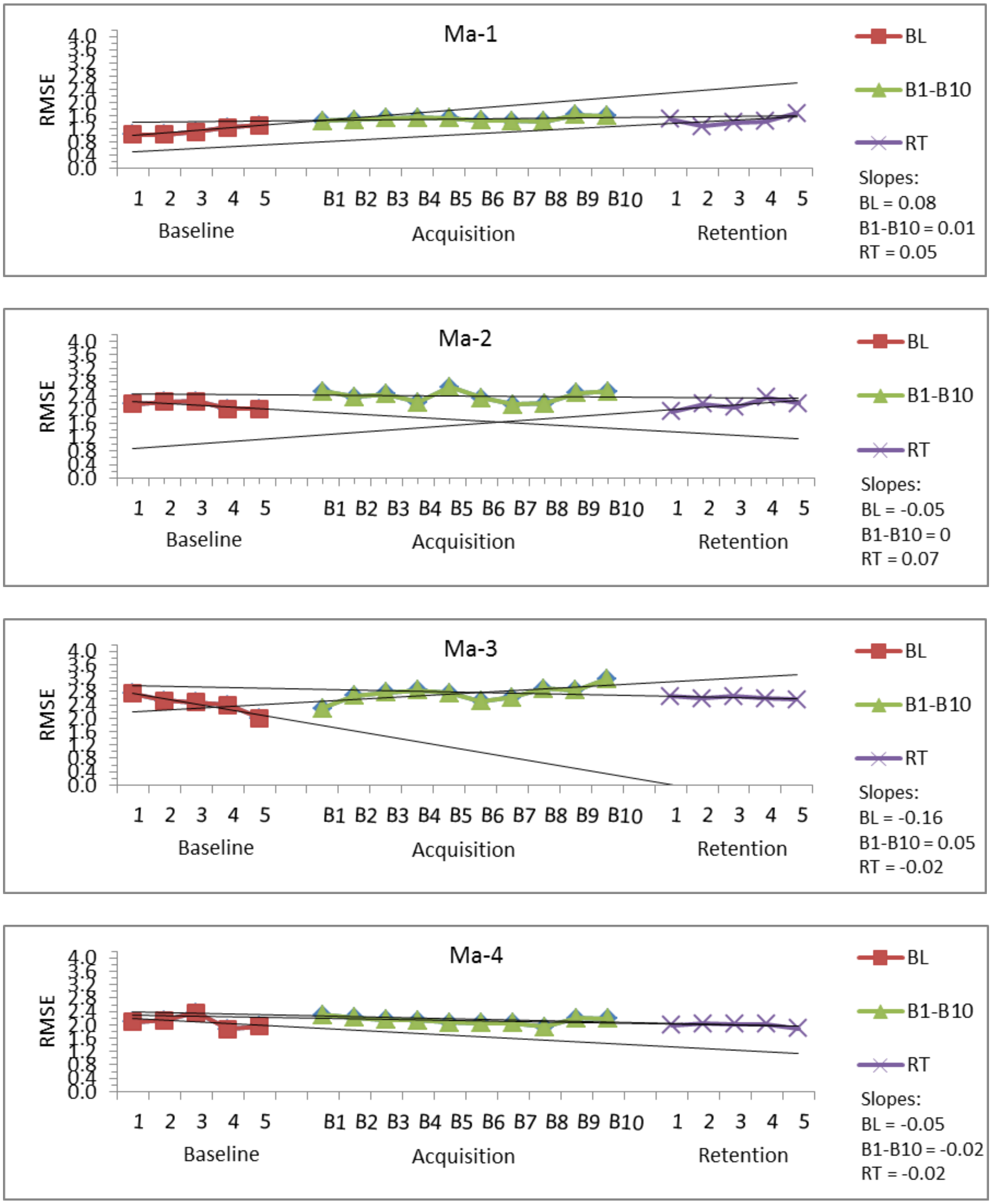


Figure 36 IFOA Participant-27 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

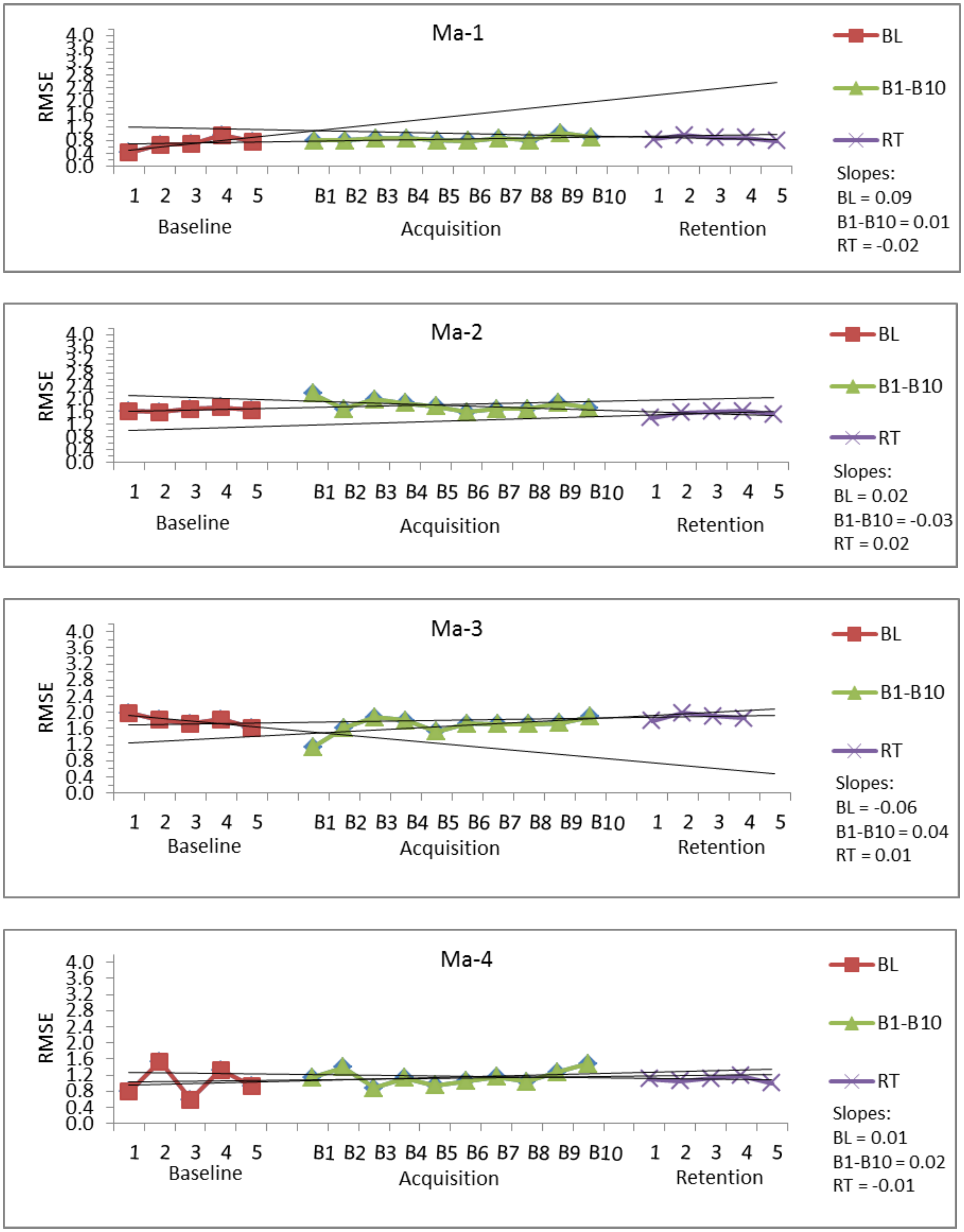


Figure 37 IFOA Participant-34 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

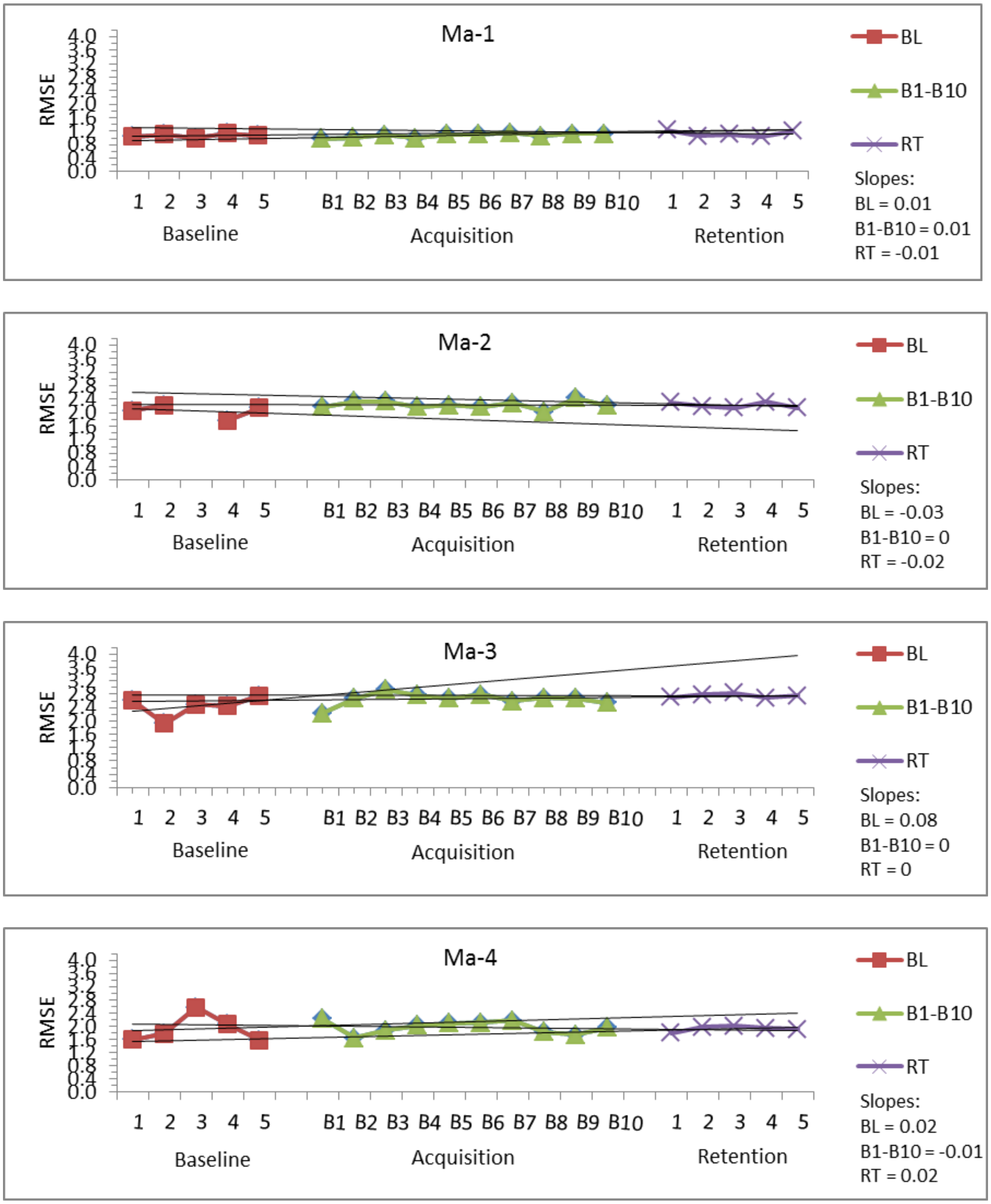


Figure 38 IFOA Participant-39 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

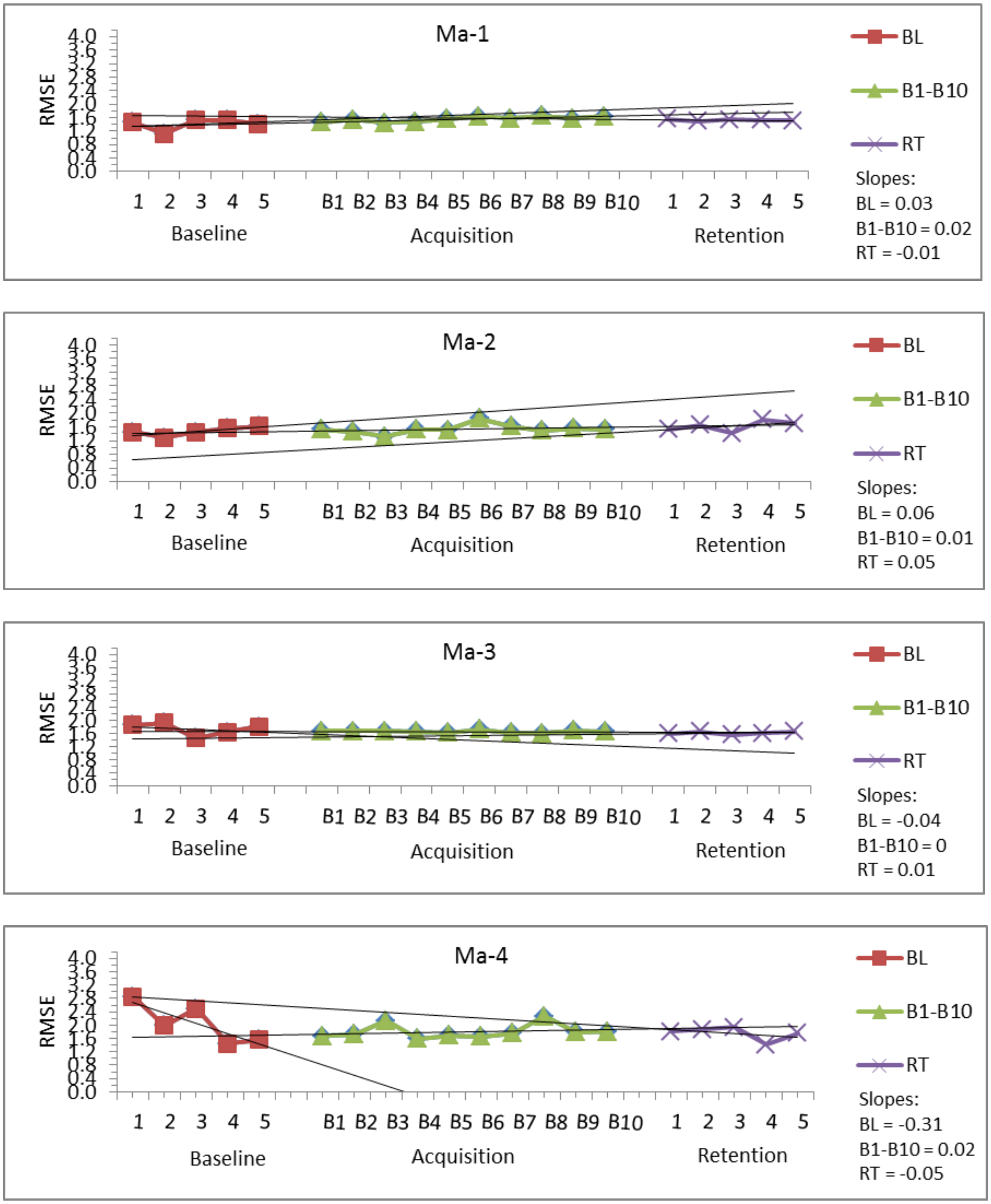


Figure 39 IFOA Participant-40 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

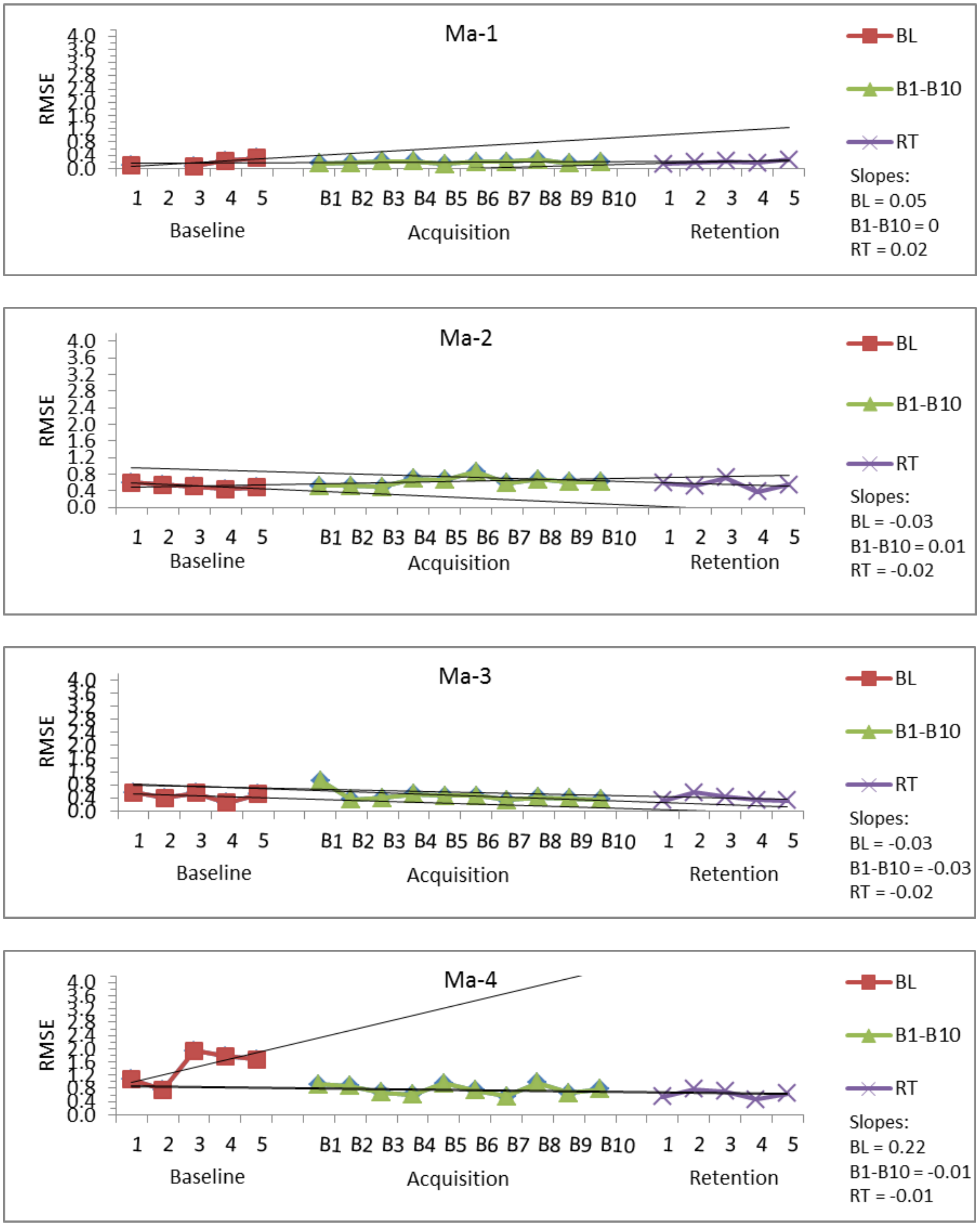


Figure 40 IFOA Participant-41 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

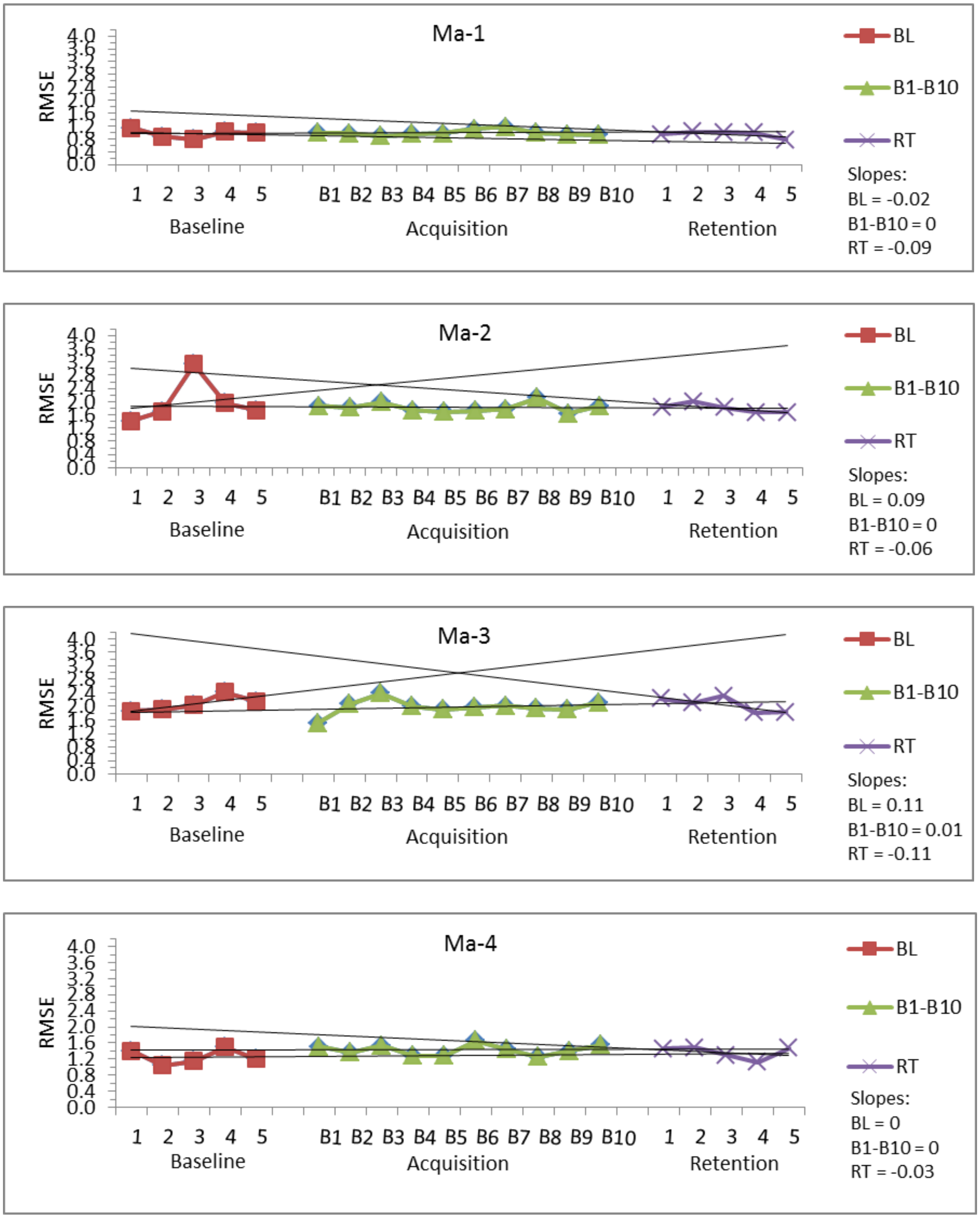


Figure 41 IFOA Participant-44 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

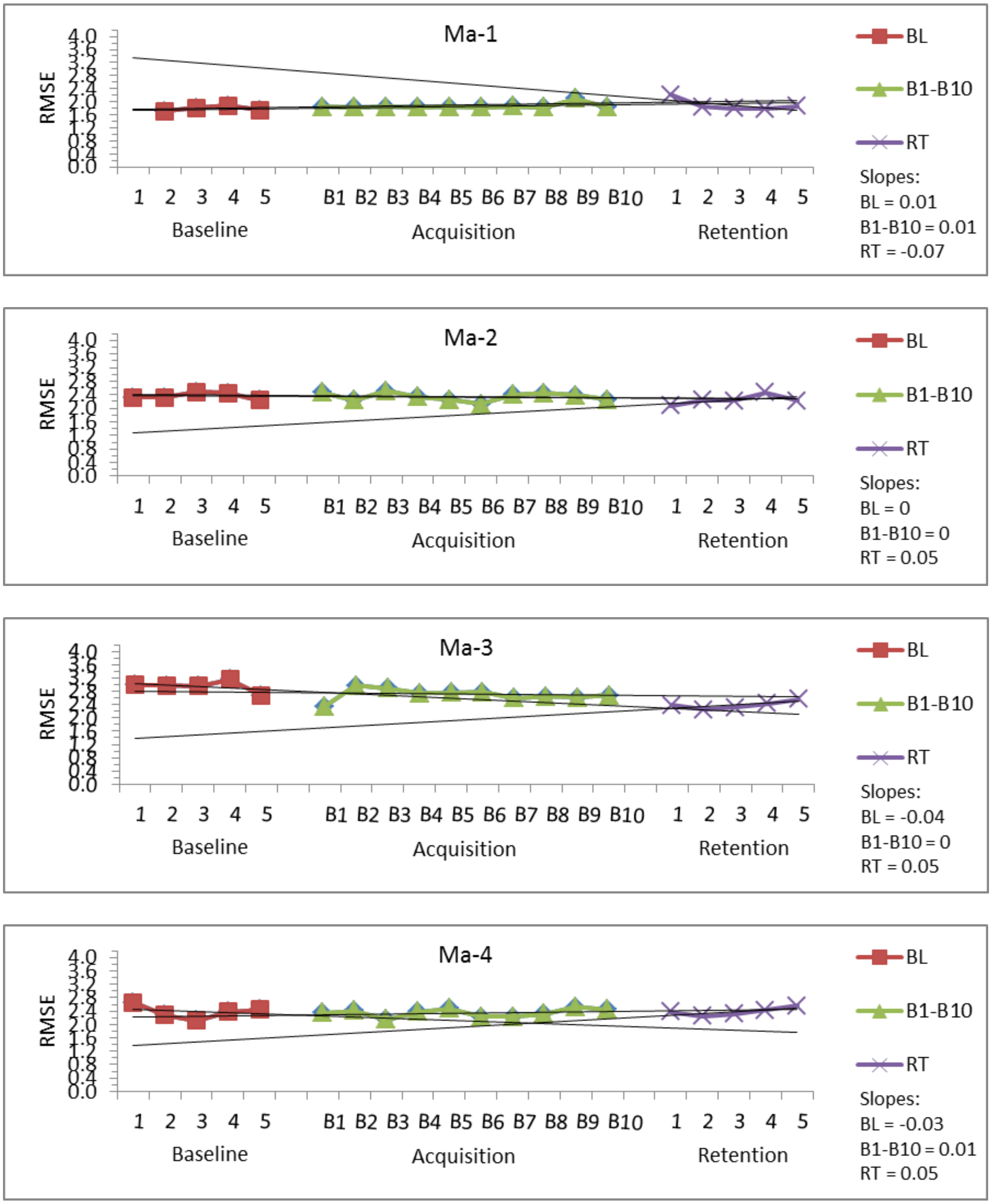


Figure 42 IFOA Participant-50 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

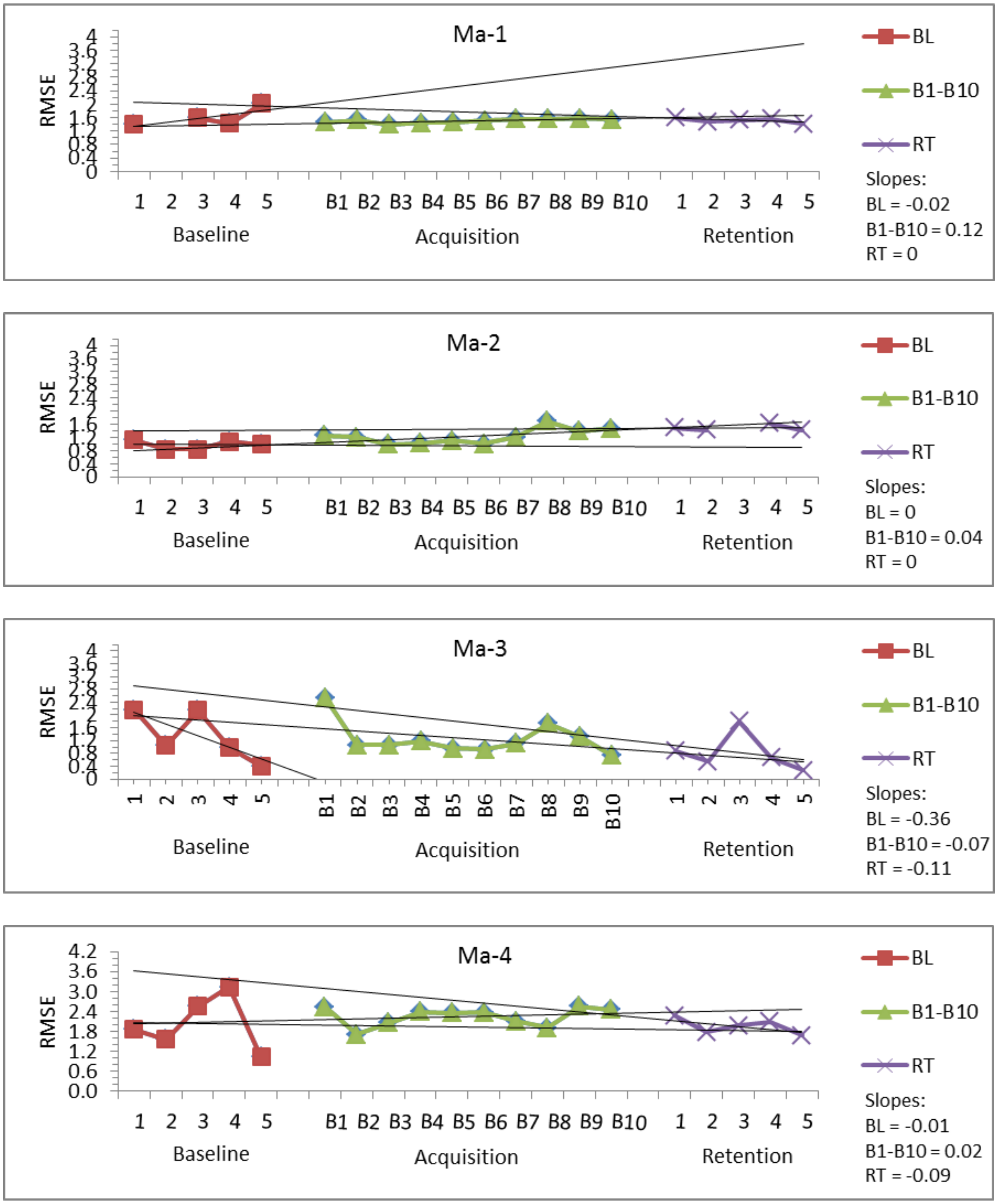


Figure 43 Control group Participant-3 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

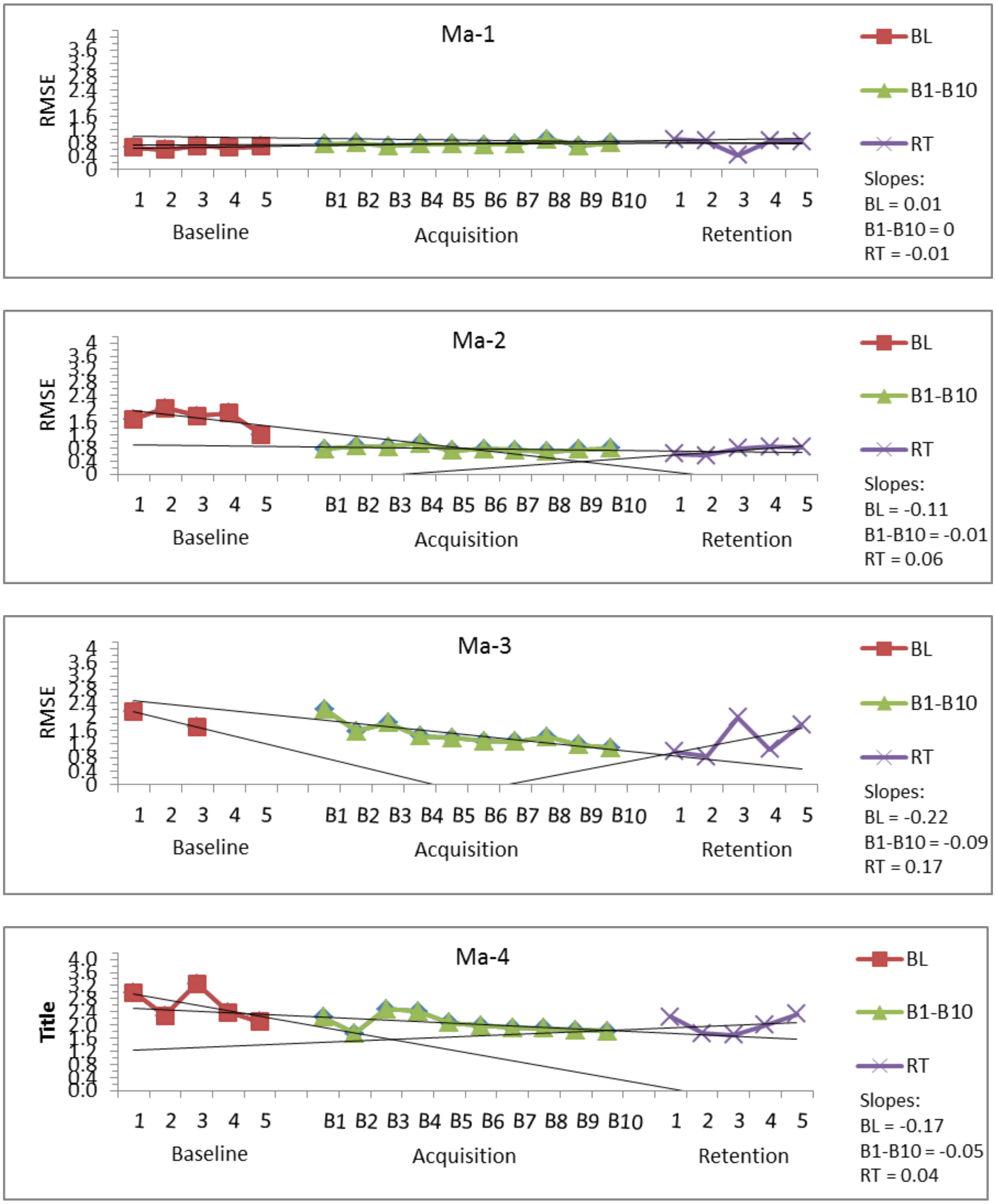


Figure 44 Control group Participant-4 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

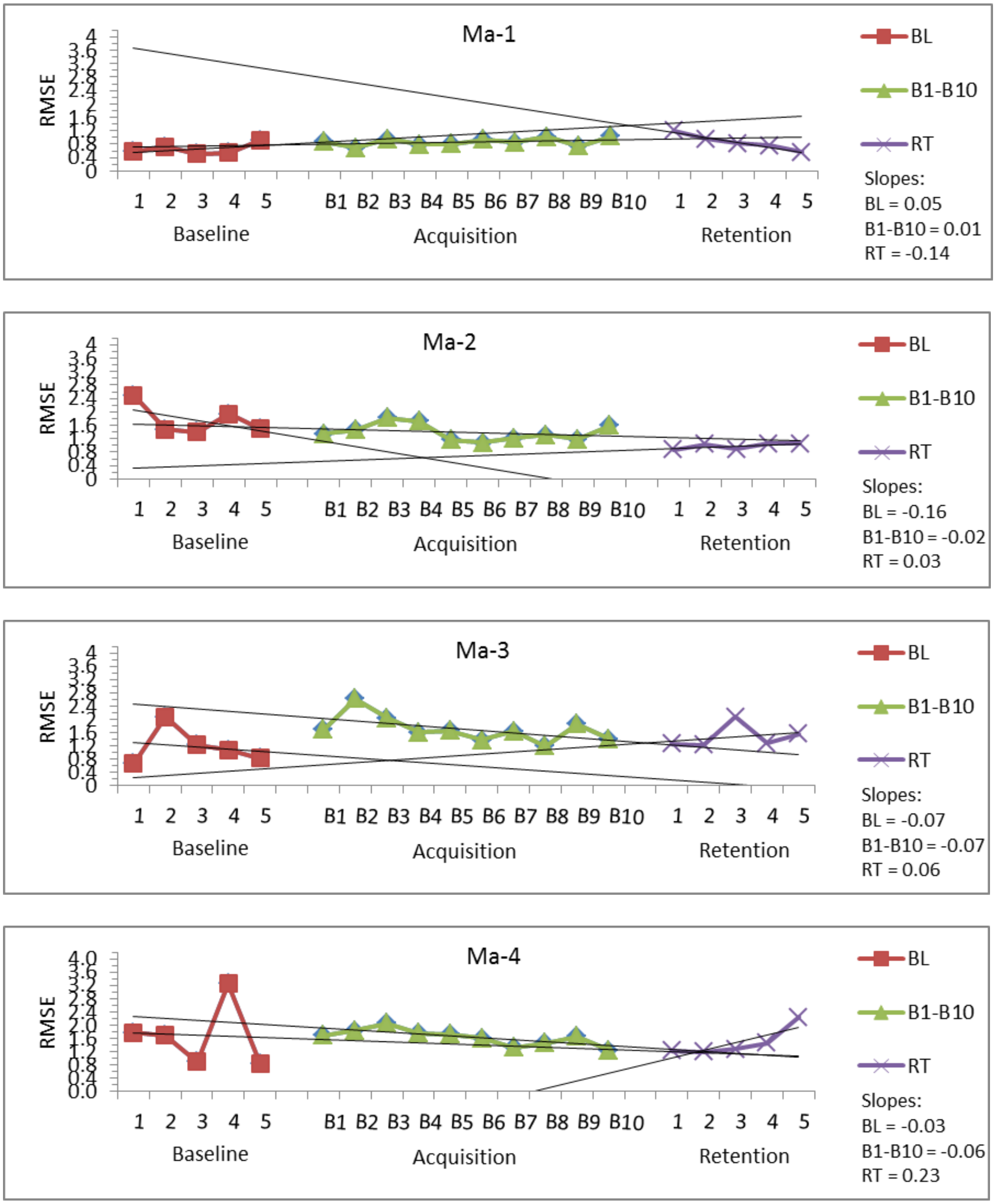


Figure 45 Control group Participant-12 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

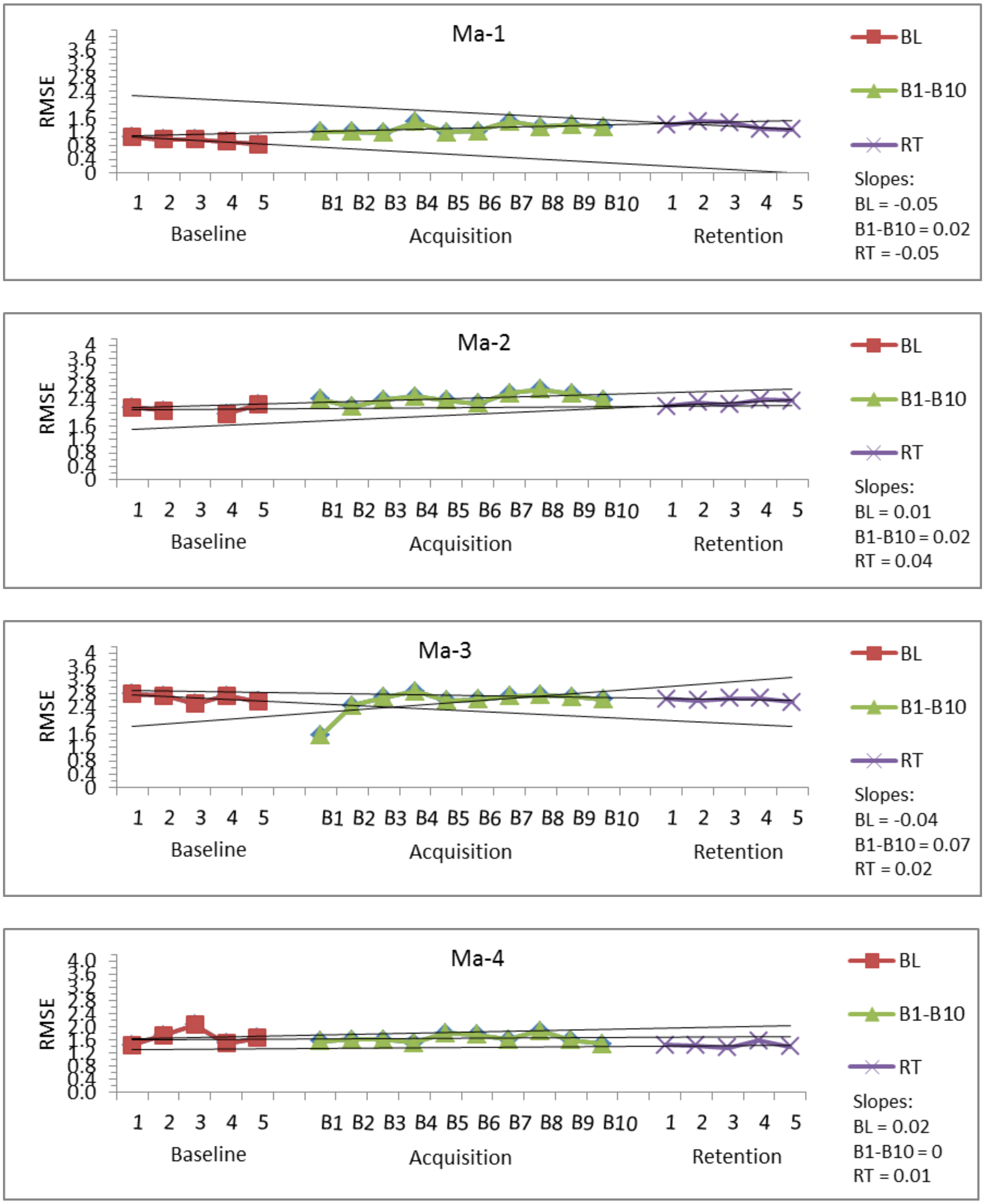


Figure 46 Control group Participant-16 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

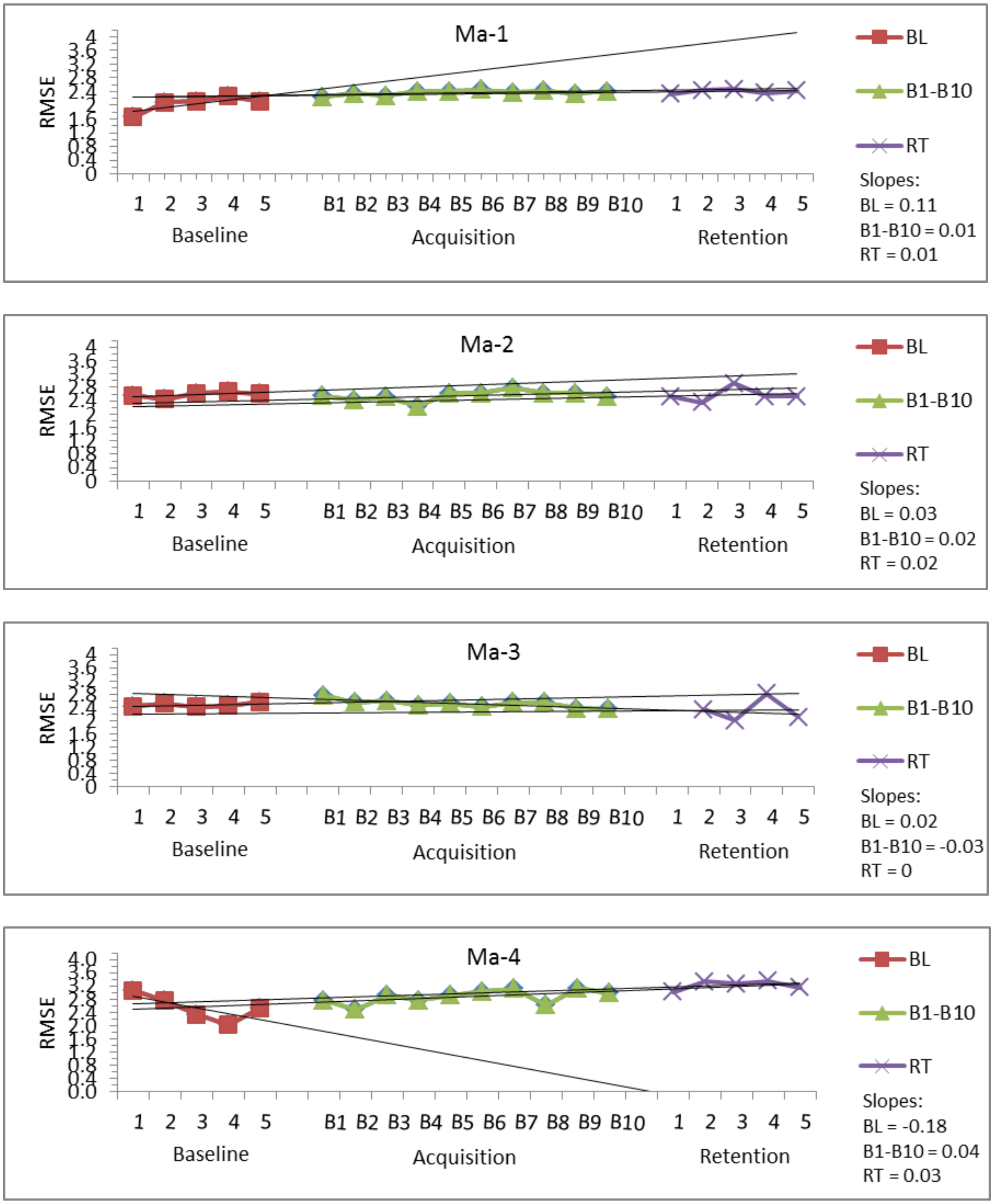


Figure 47 Control group Participant-18 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

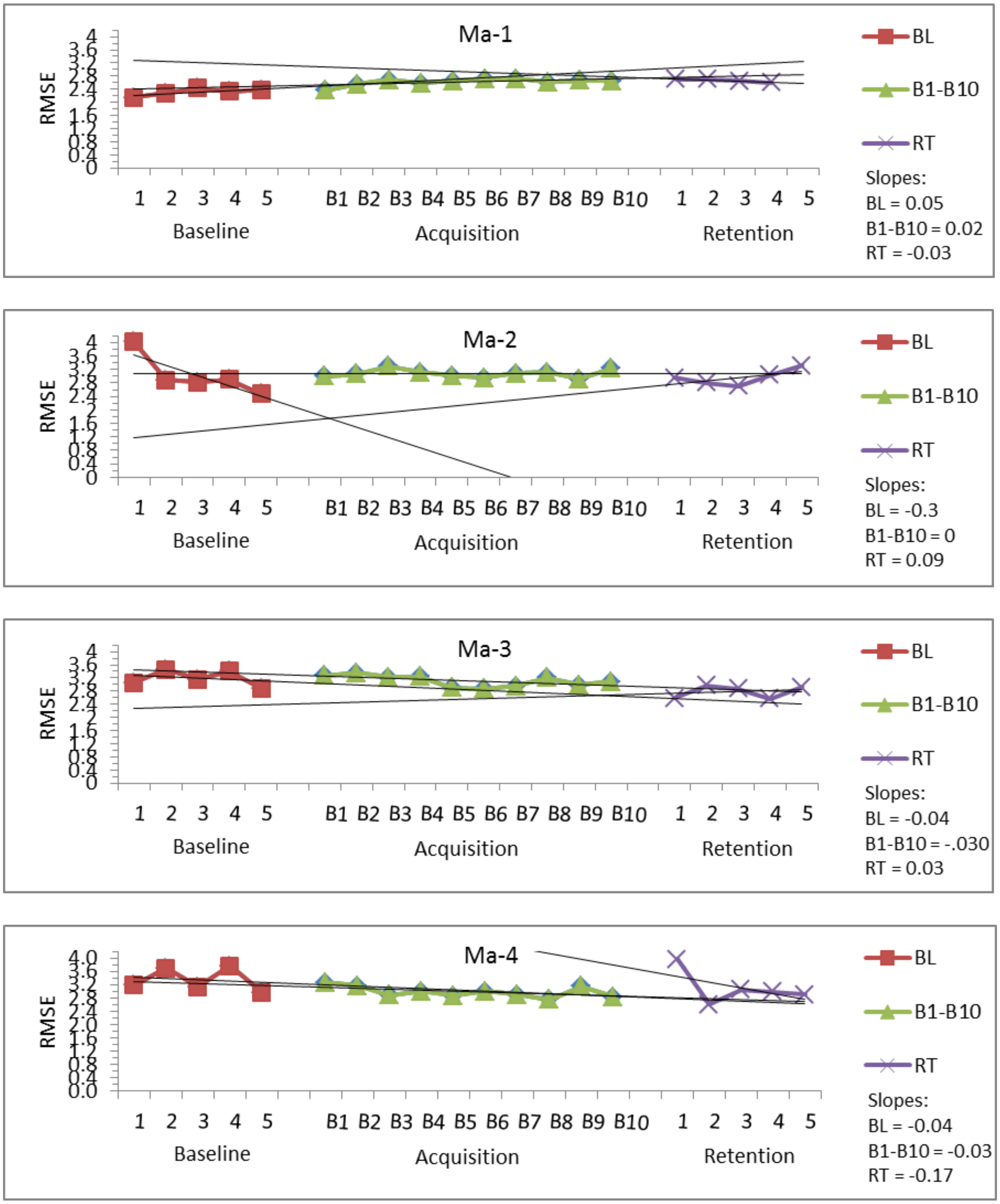


Figure 48 Control group Participant-19 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

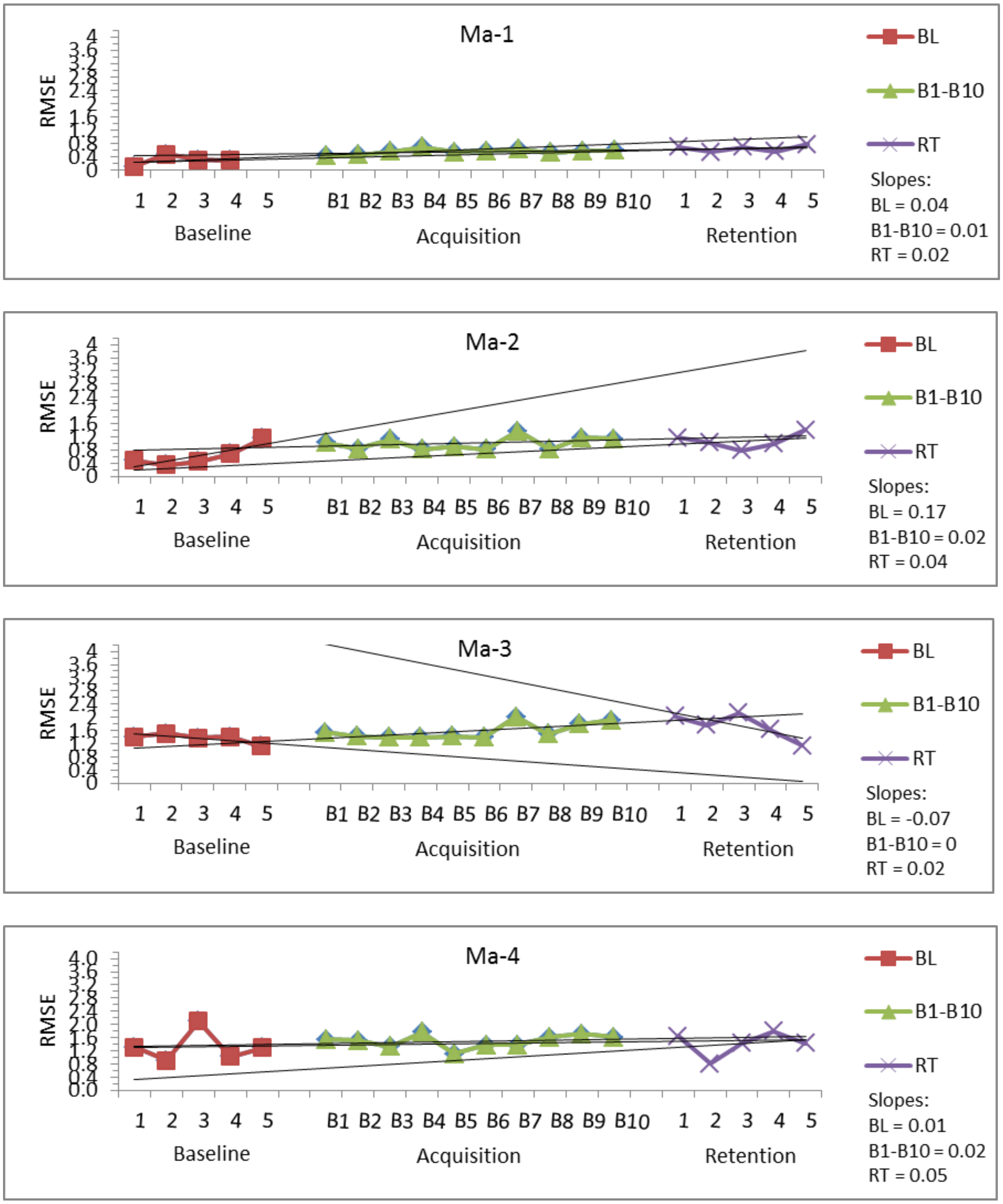


Figure 49 Control group Participant-21 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

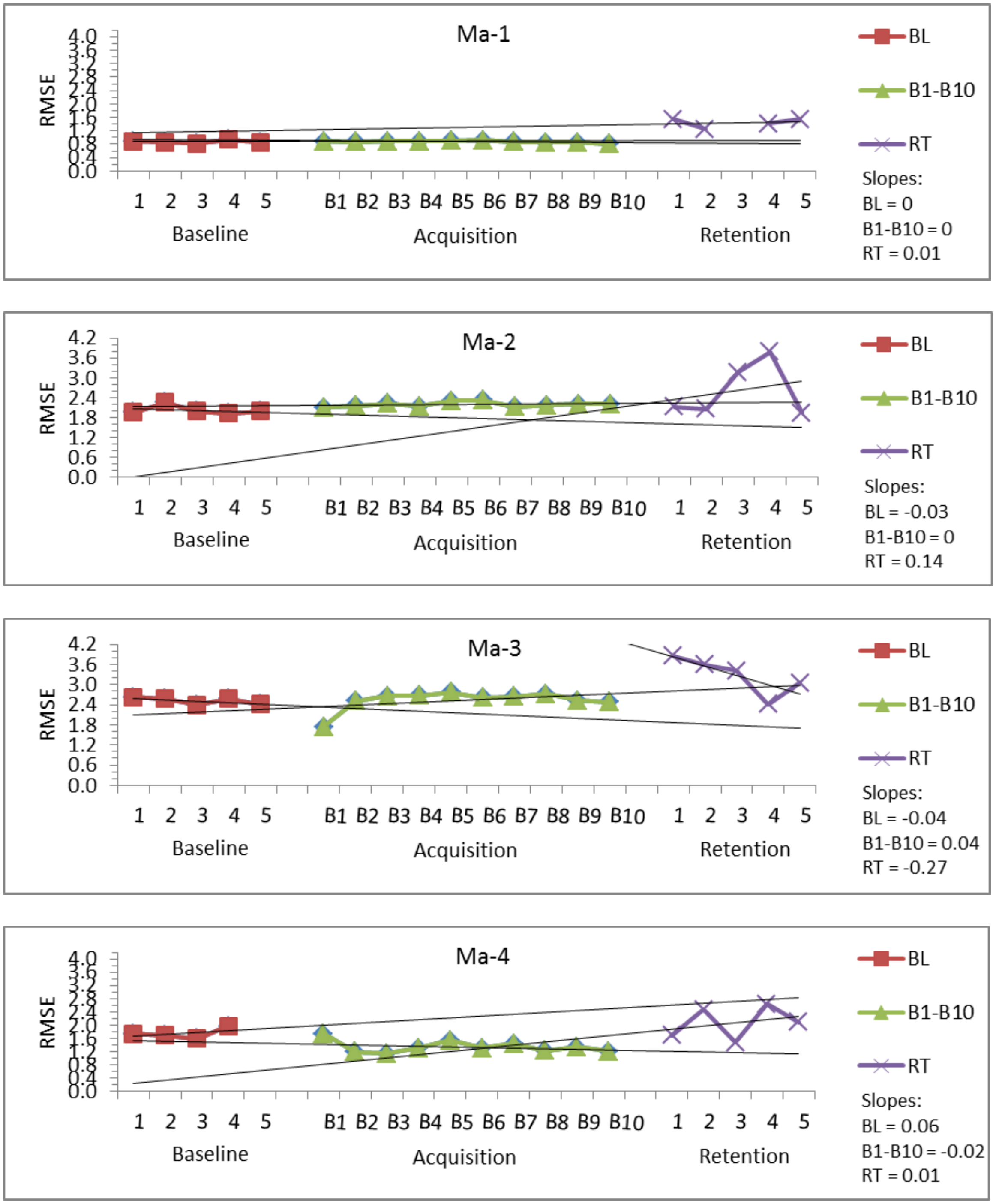


Figure 50 Control group Participant-23 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

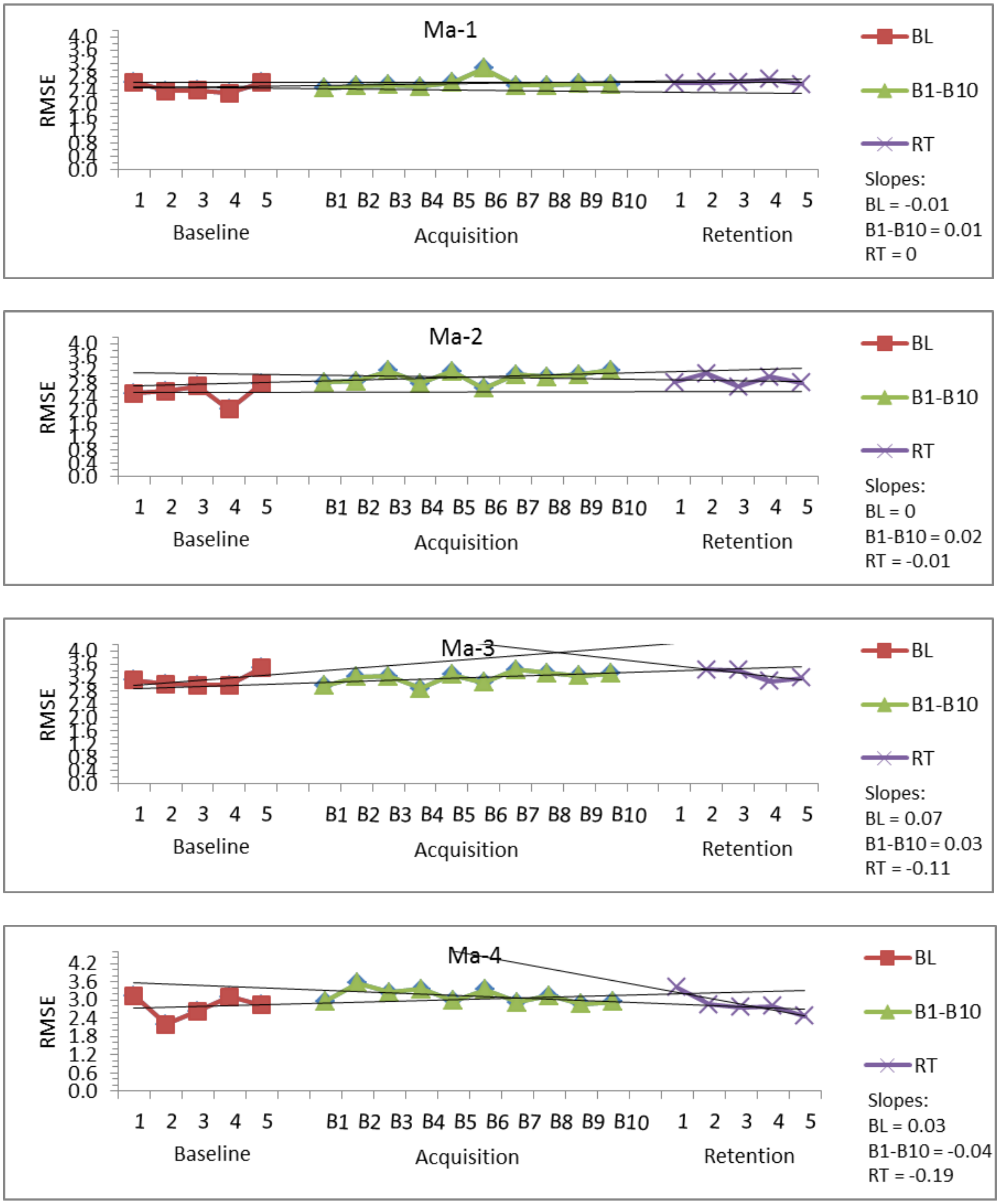


Figure 51 Control group Participant-28 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

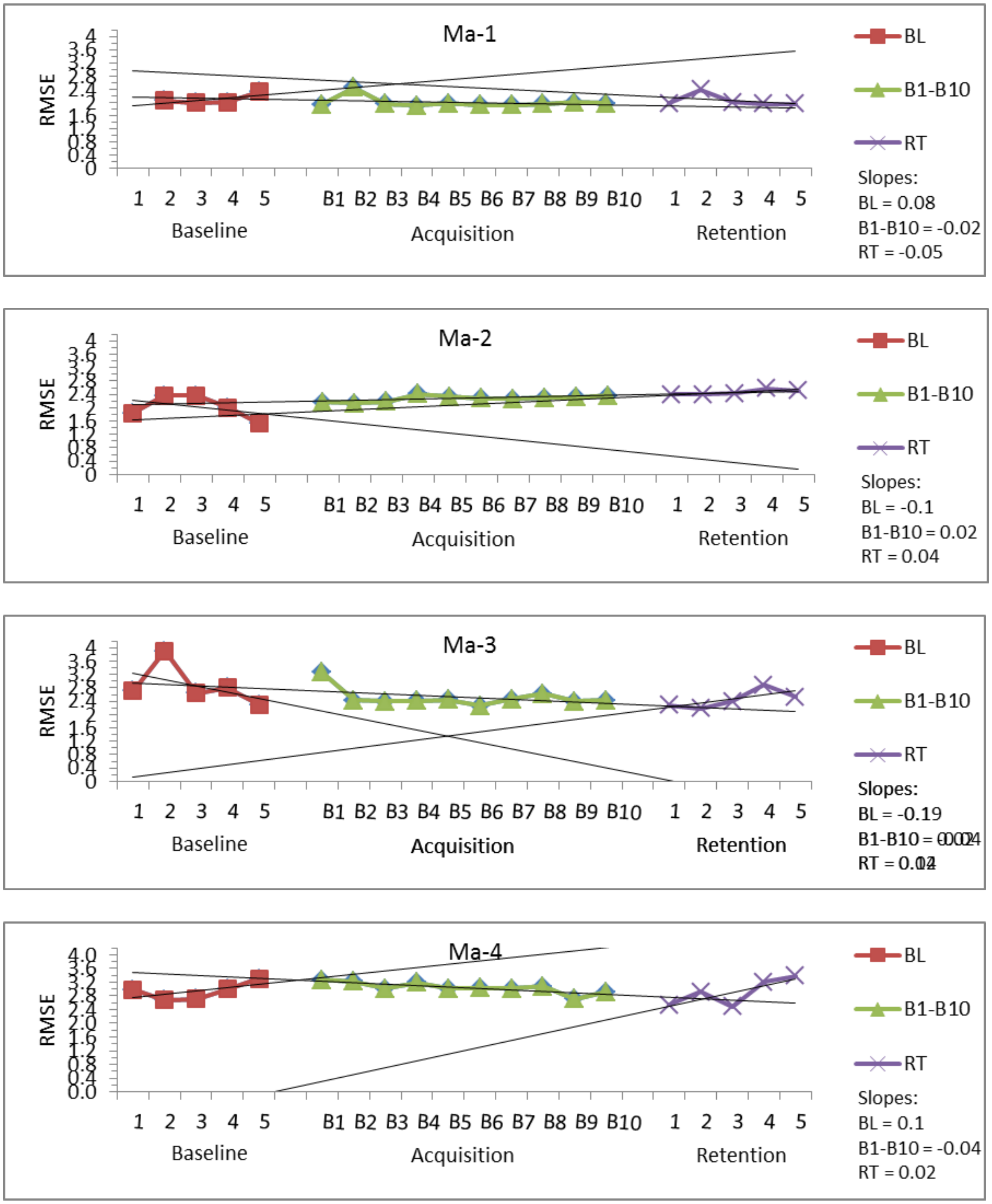


Figure 52 Control group Participant-30 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

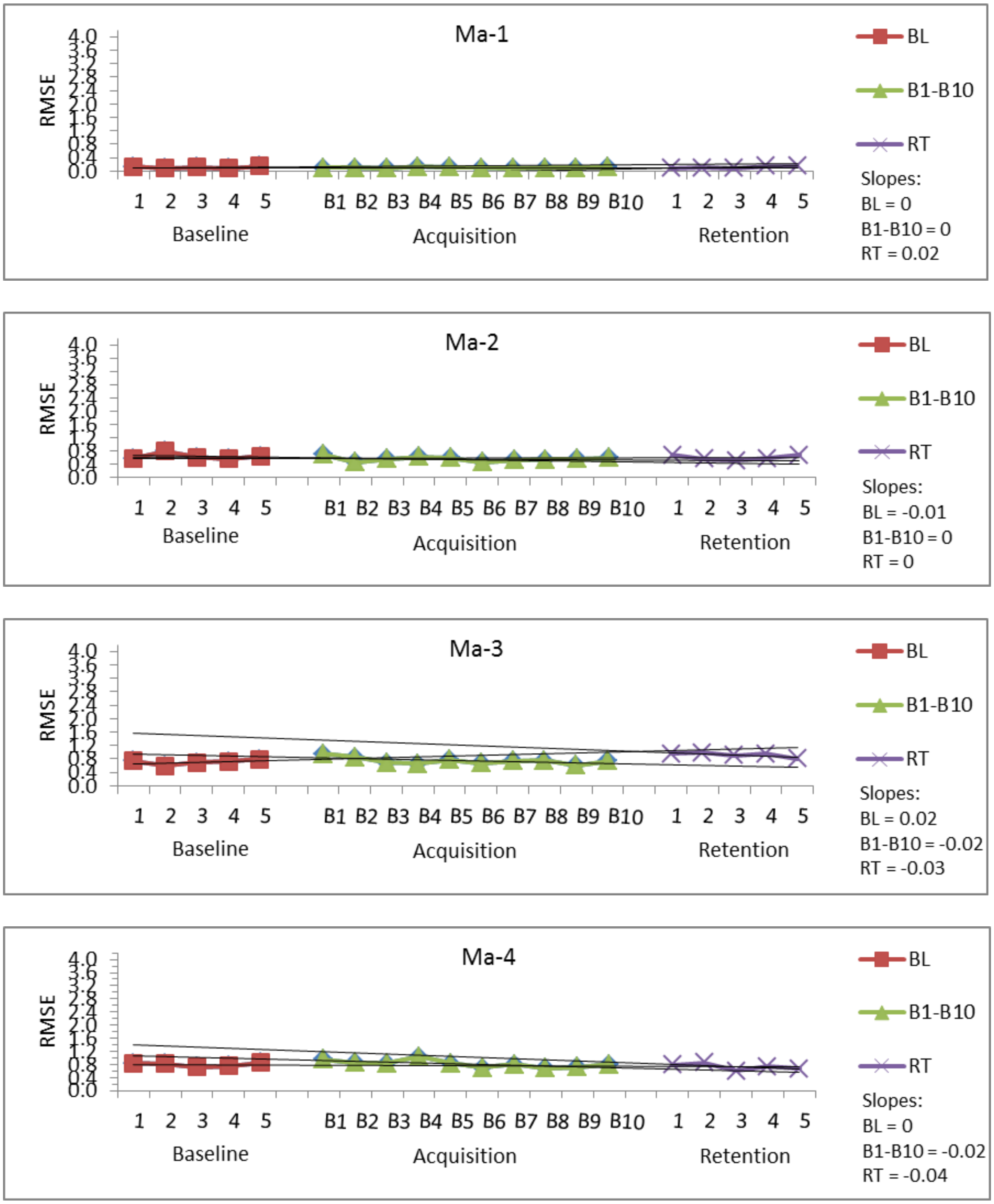


Figure 53 Control group Participant-31 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

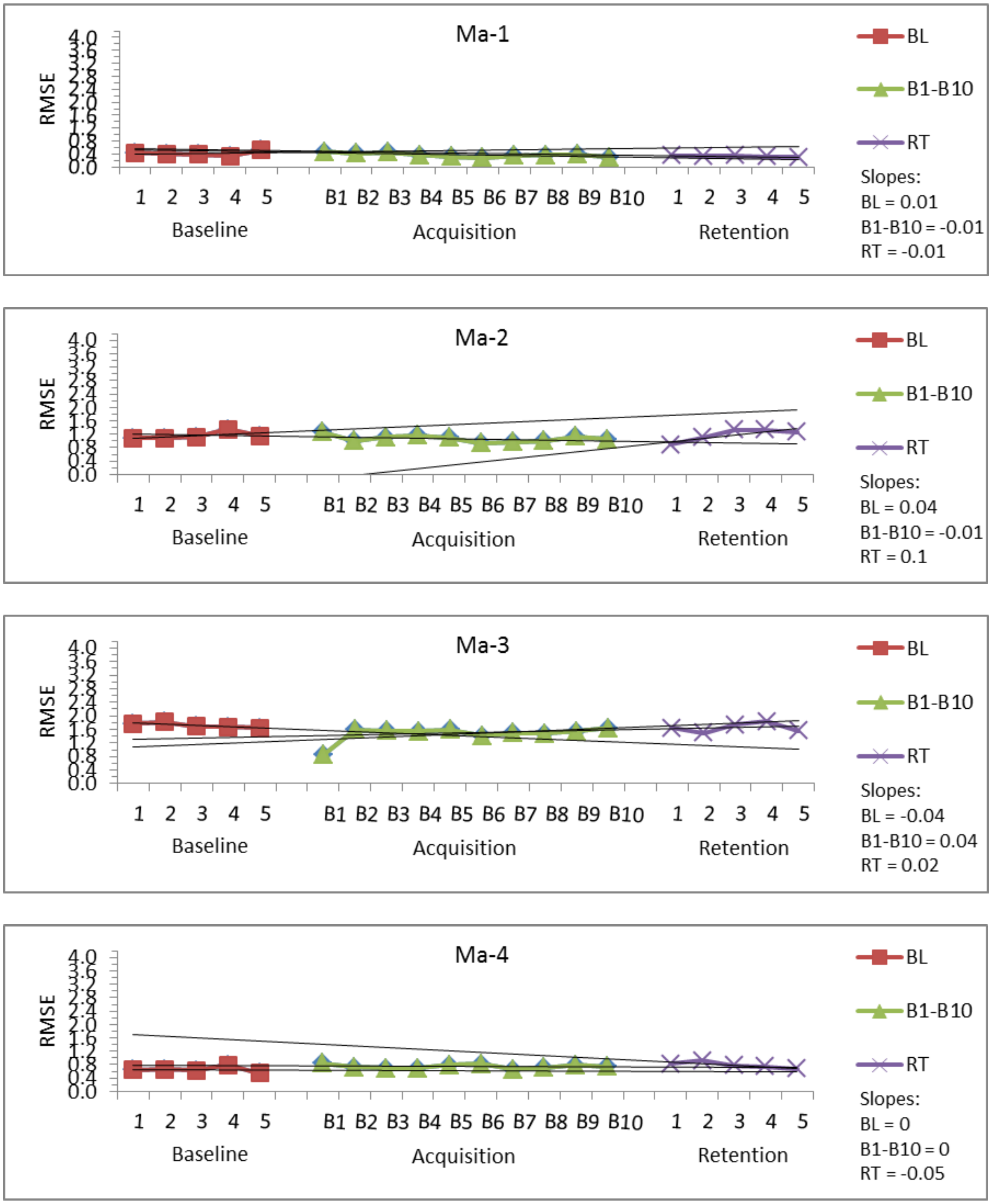


Figure 54 Control group Participant-35 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

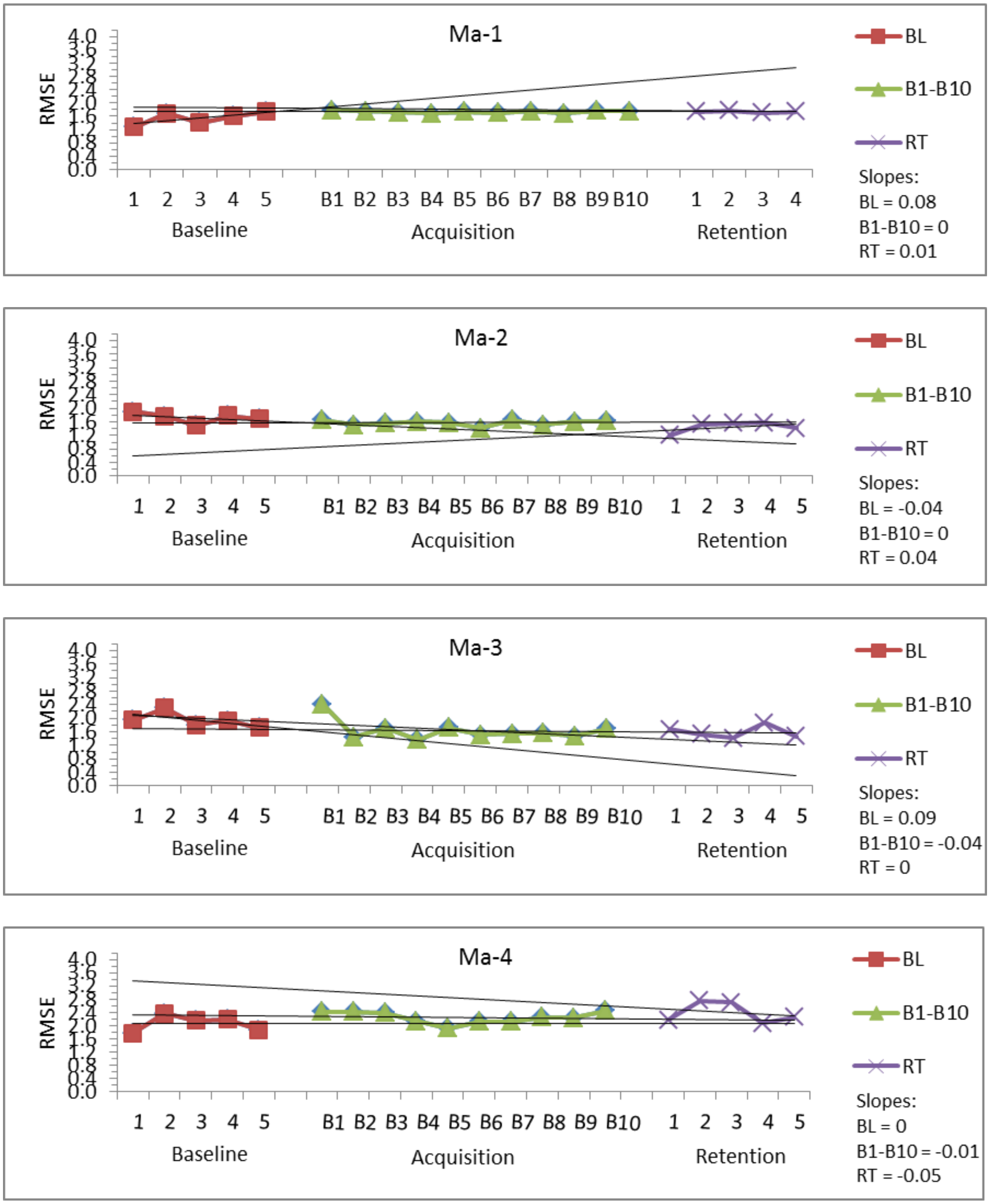


Figure 55 Control group Participant-37 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

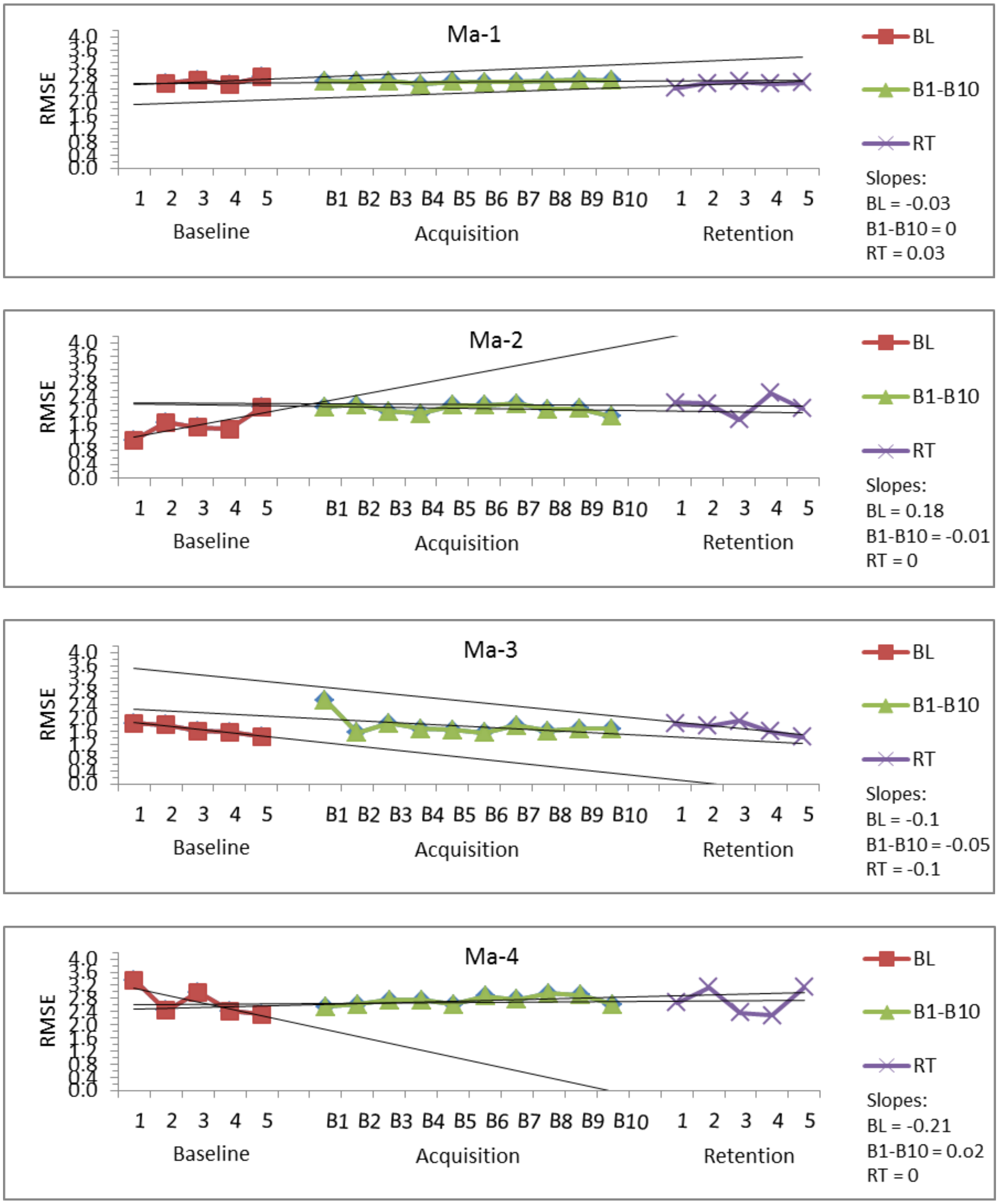


Figure 56 Control group Participant-48 RMSE for the practice words (Ma-1, Ma-2, Ma-3, Ma-4) during baseline, acquisition phase, and retention test. IFOA = internal focus of attention, BL = baseline, B1-B10 = blocks of practice, RT = retention test.

APPENDIX F

SINGLE SUBJECT DATA: RMSE FOR THE PROBED WORDS DURING BASELINE, ACQUISITION PHASE, AND IN TRANSFER TEST

Probes T1 and T2

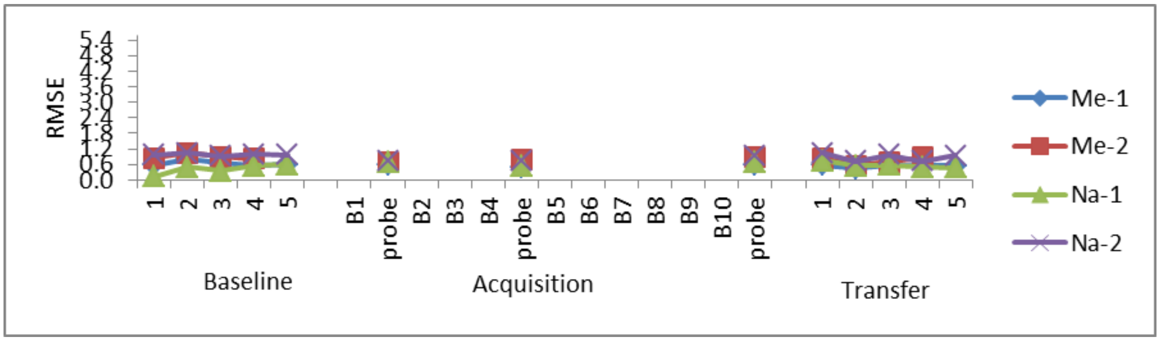


Figure 57 EFOA-Participant-10 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

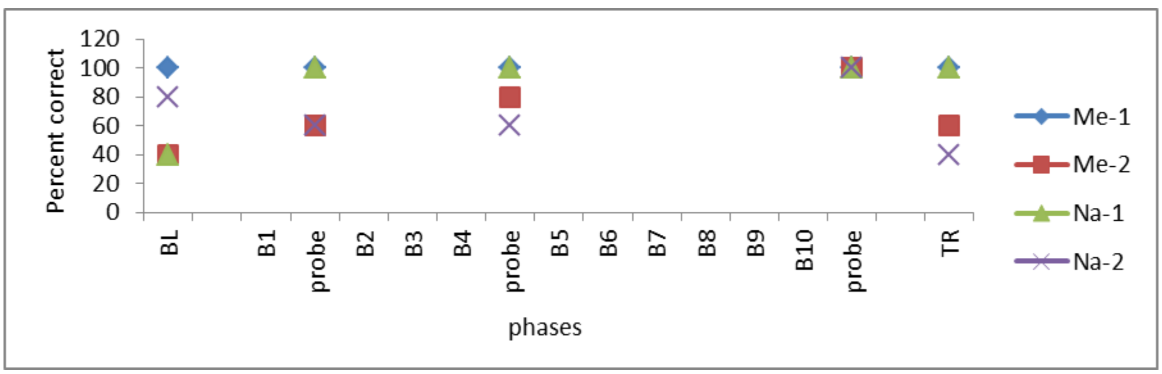


Figure 58 EFOA-Participant-10 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

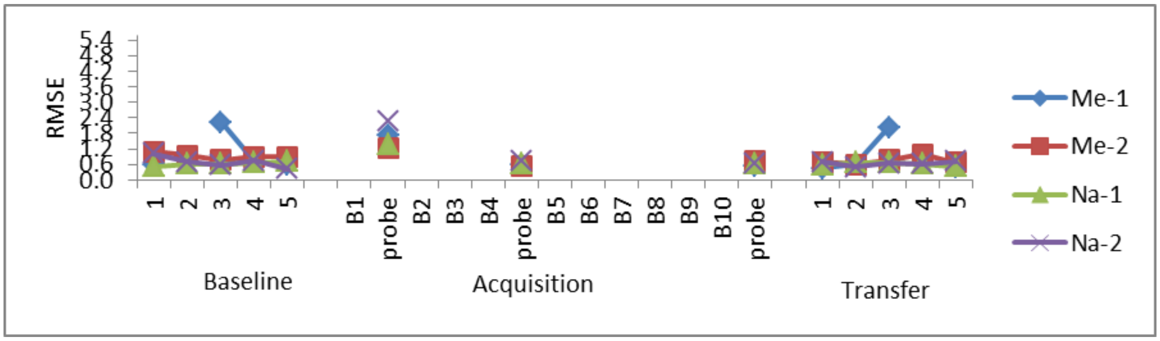


Figure 59 EFOA-Participant-20 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

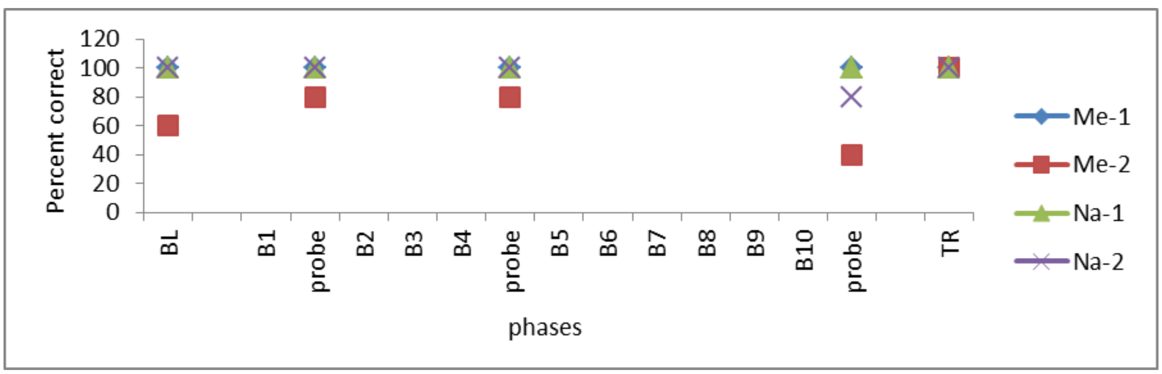


Figure 60 EFOA-Participant-20 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

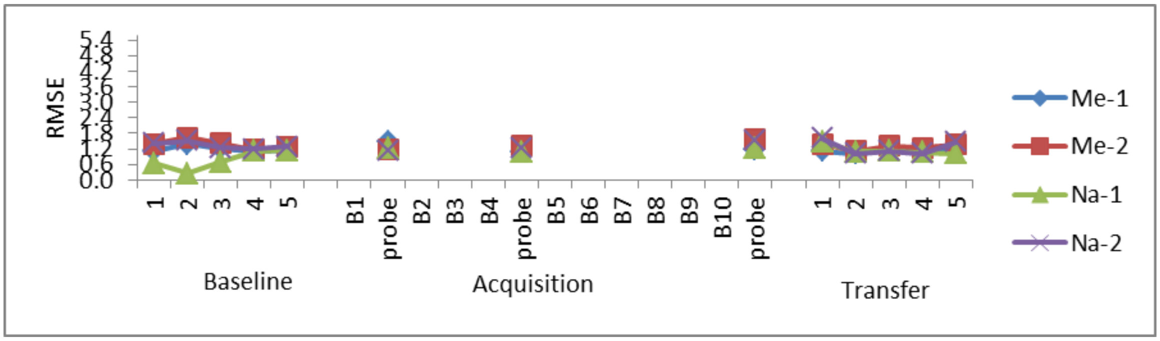


Figure 61 EFOA-Participant-32 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

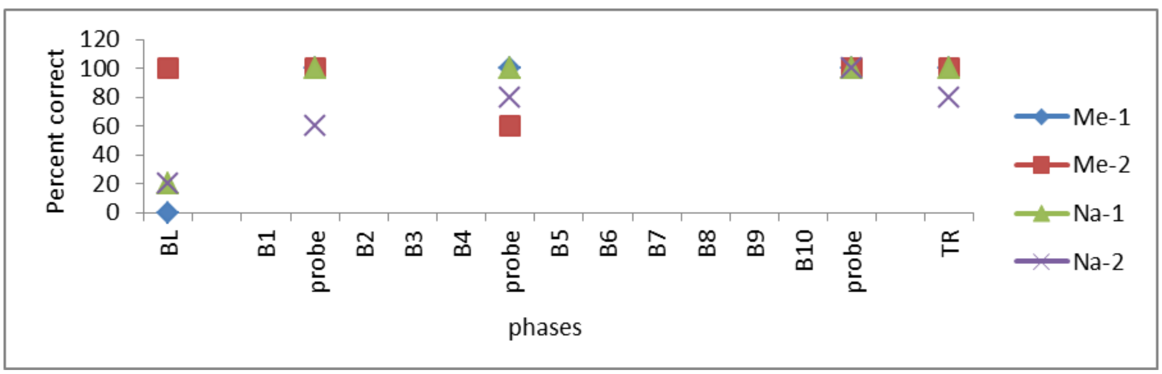


Figure 62 EFOA-Participant-32 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

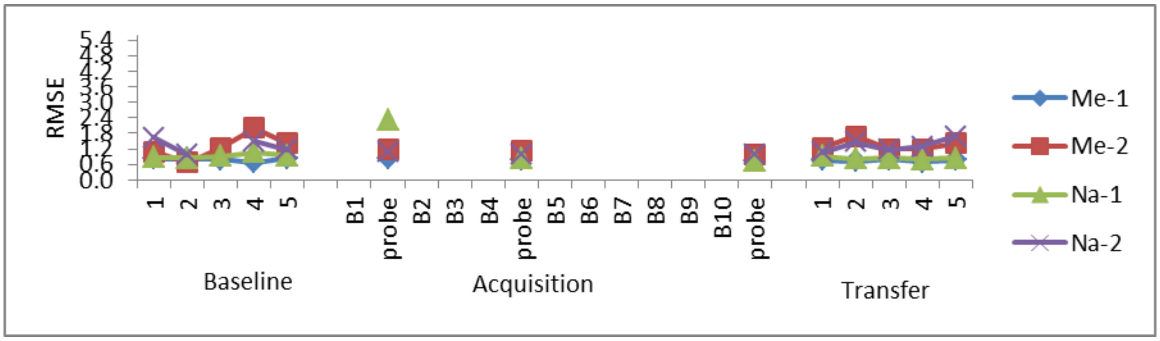


Figure 63 EFOA-Participant-33 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

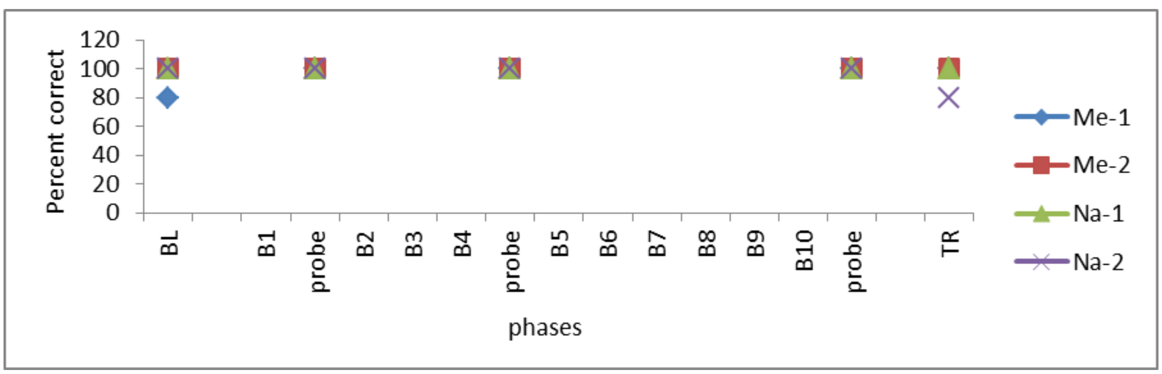


Figure 64 EFOA-Participant-33 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

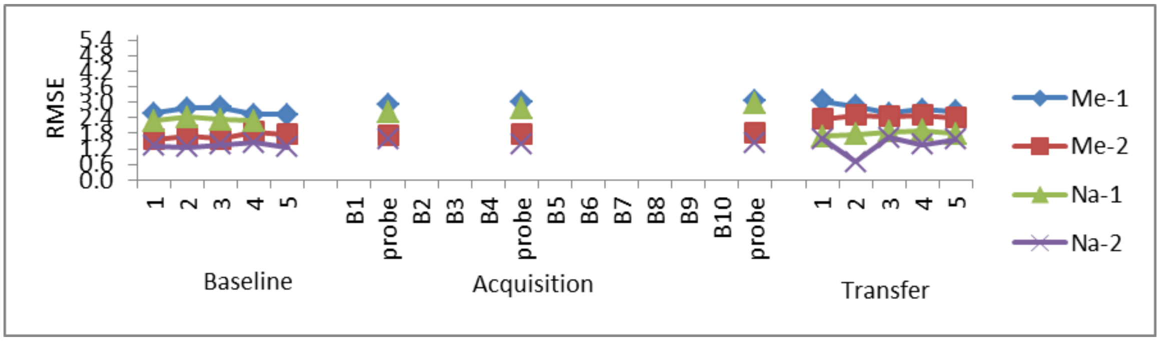


Figure 65 EFOA-Participant-38 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

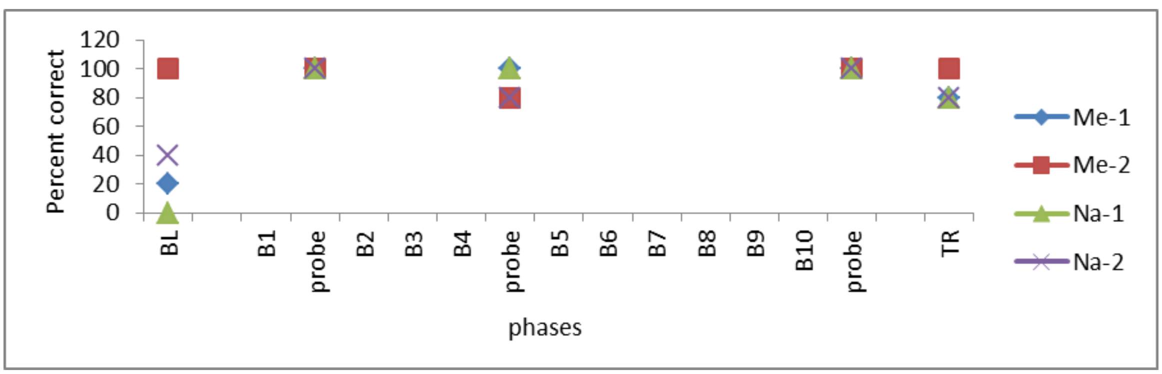


Figure 66 EFOA-Participant-38 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

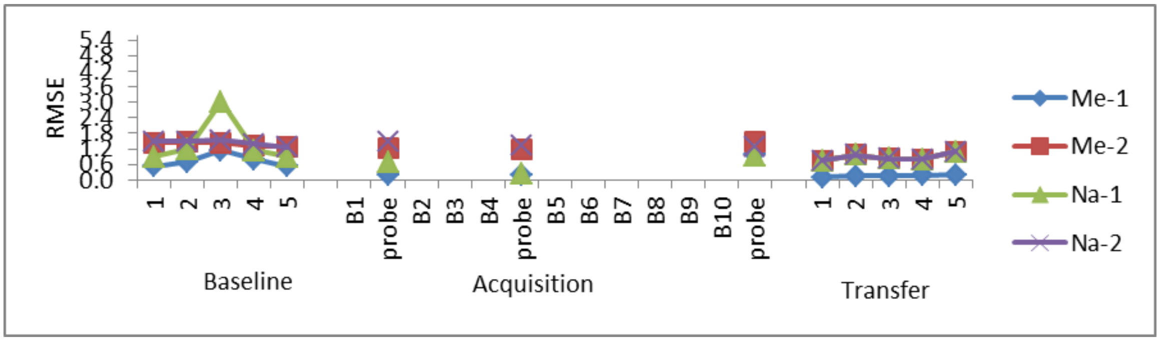


Figure 67   EFOA-Participant-43 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

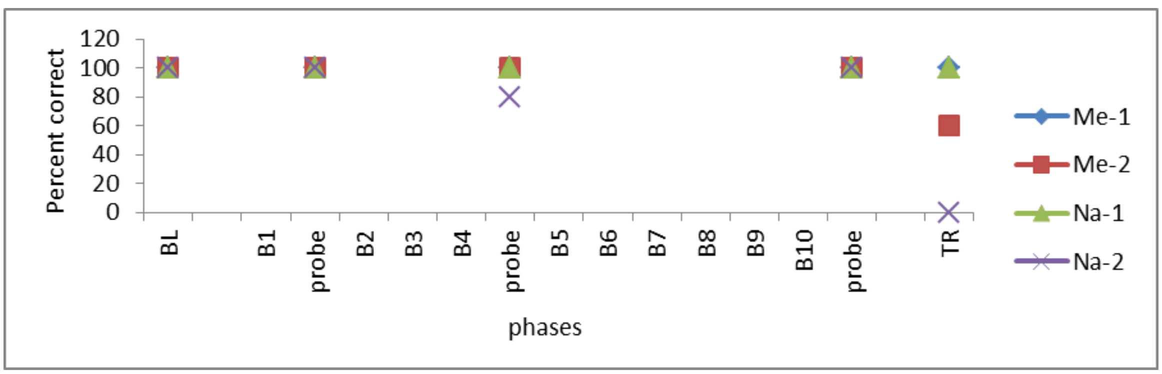


Figure 68   EFOA-Participant-43 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

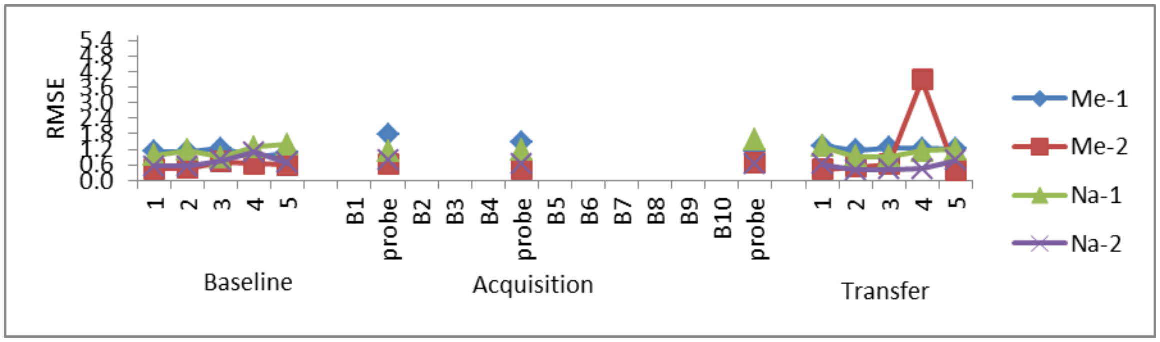


Figure 69 EFOA-Participant-47 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

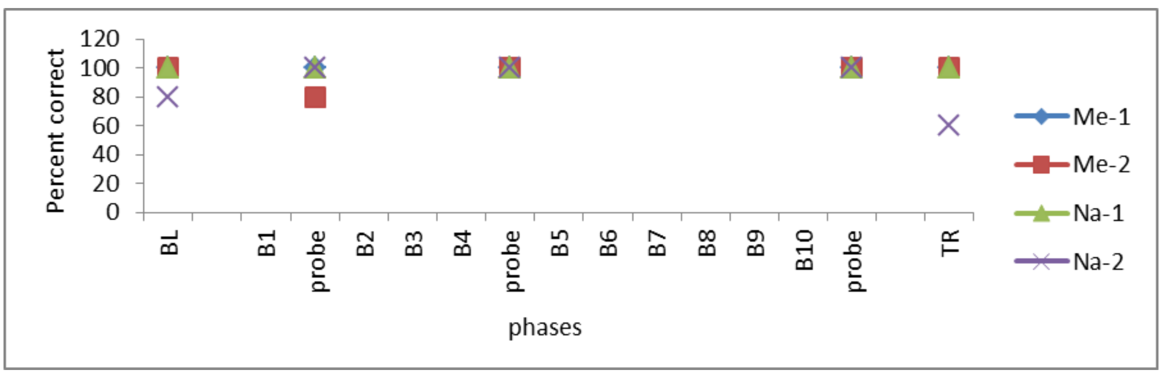


Figure 70 EFOA-Participant-47 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

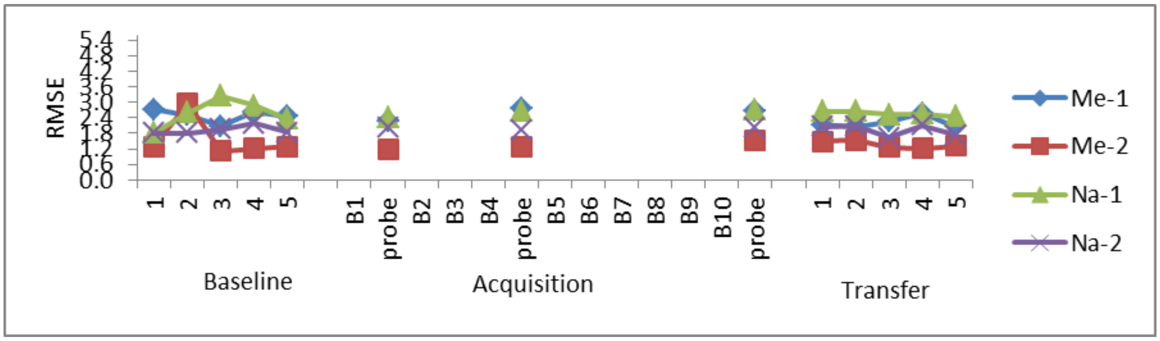


Figure 71 EFOA-Participant-6 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

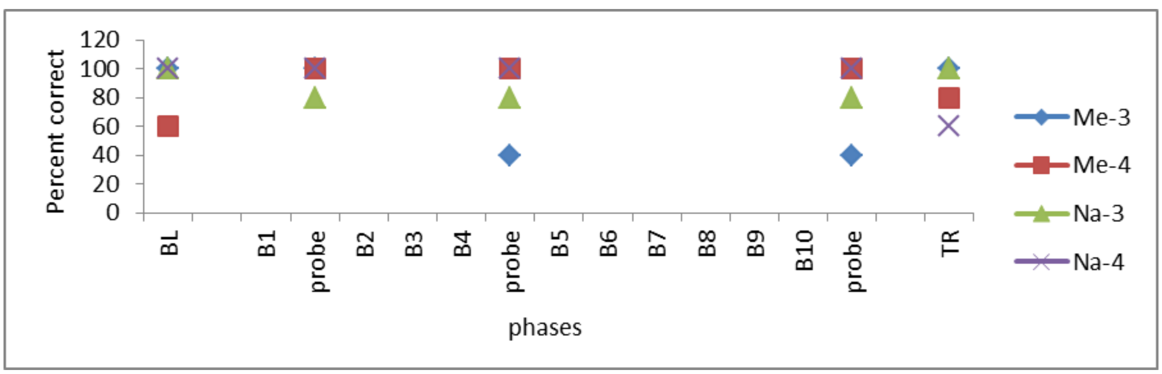


Figure 72 EFOA Participant-6 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

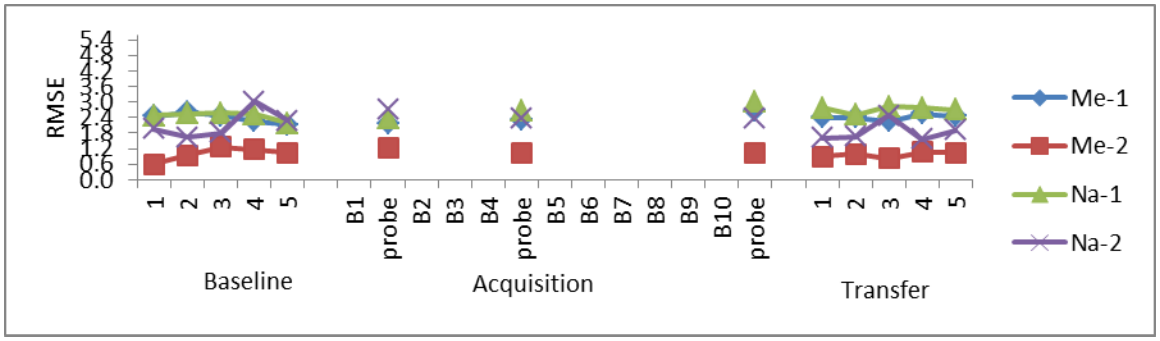


Figure 73 EFOA-Participant-13The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

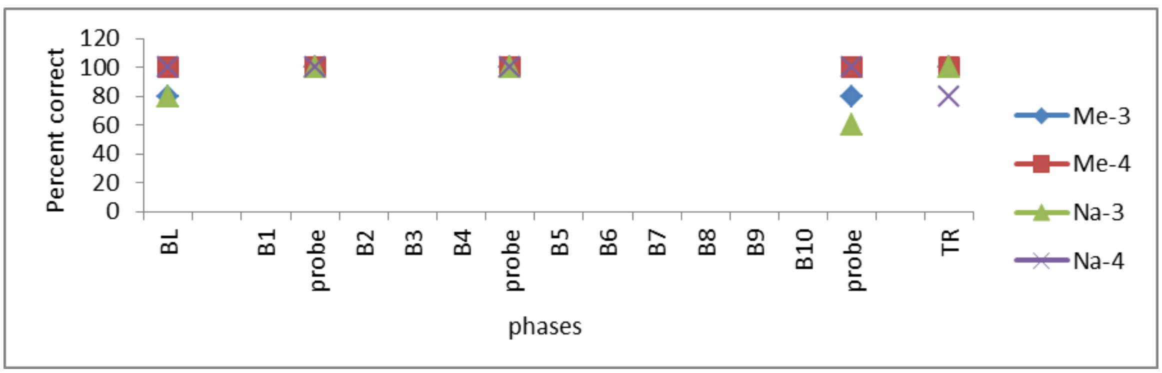


Figure 74 EFOA-Participant-13 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

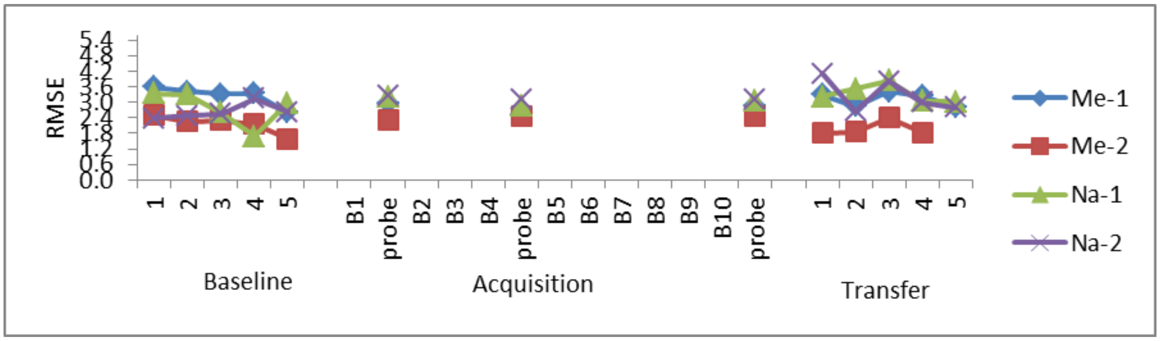


Figure 75 EFOA-Participant-22 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2)during baseline, acquisition phase, and transfer test.

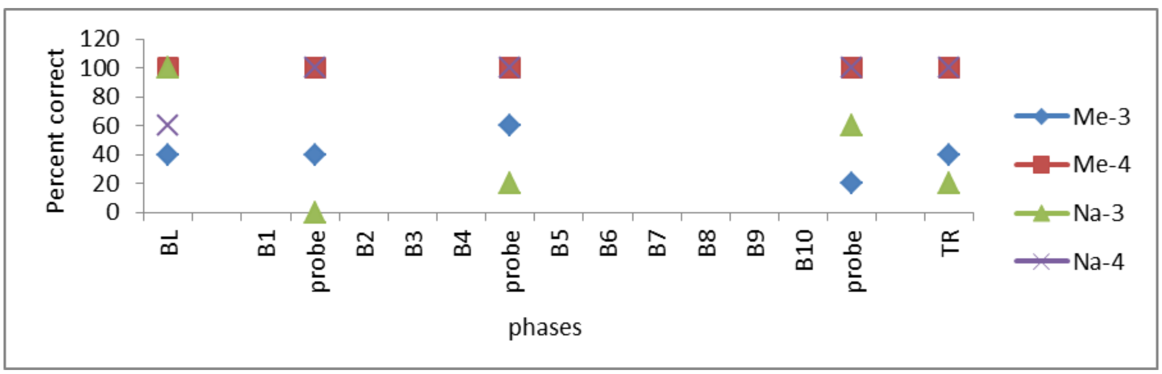


Figure 76 EFOA-Participant-22 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

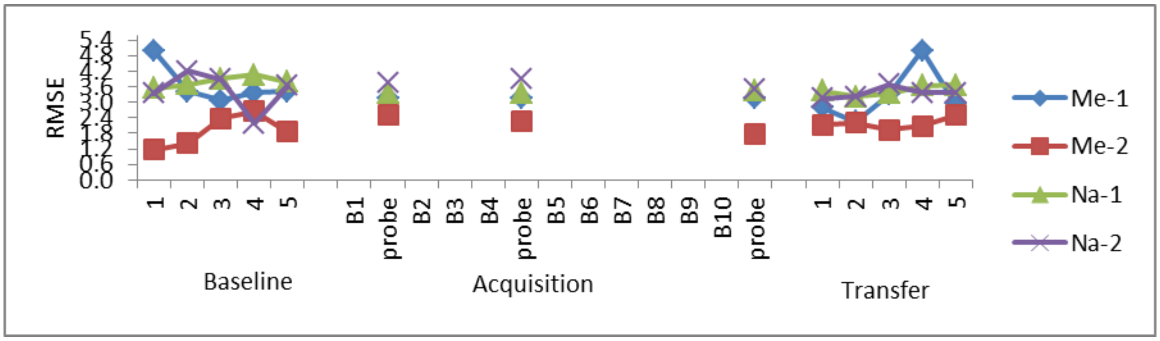


Figure 77 EFOA-Participant-29 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

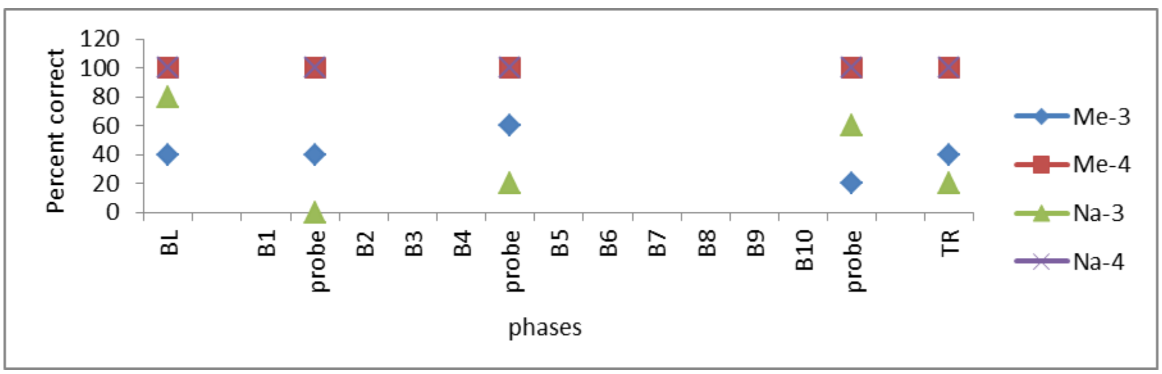


Figure 78 EFOA-Participant-29 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

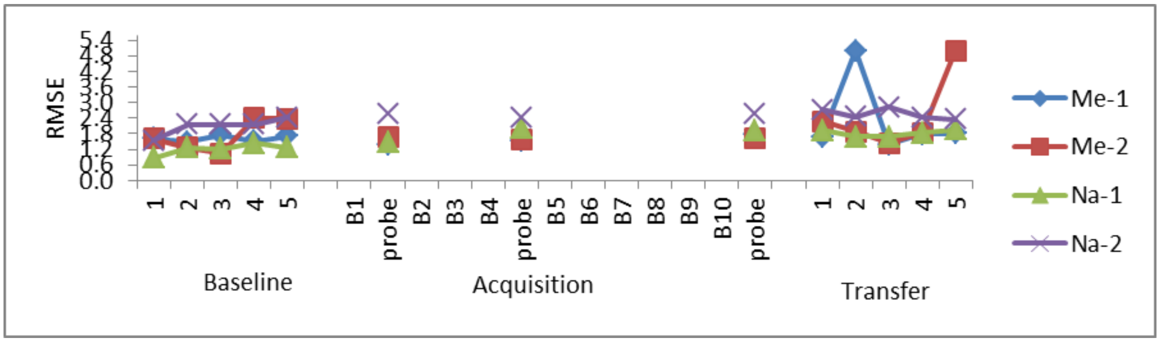


Figure 79 EFOA-Participant-42 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

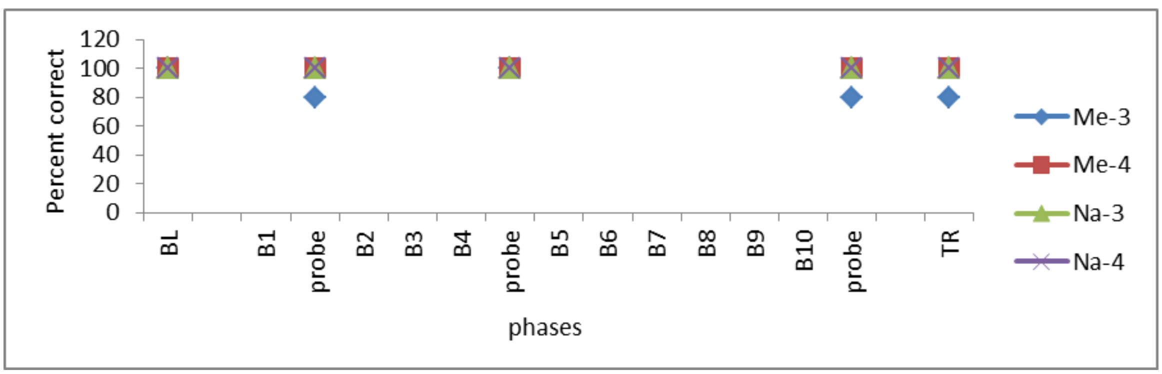


Figure 80 EFOA-Participant-42 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

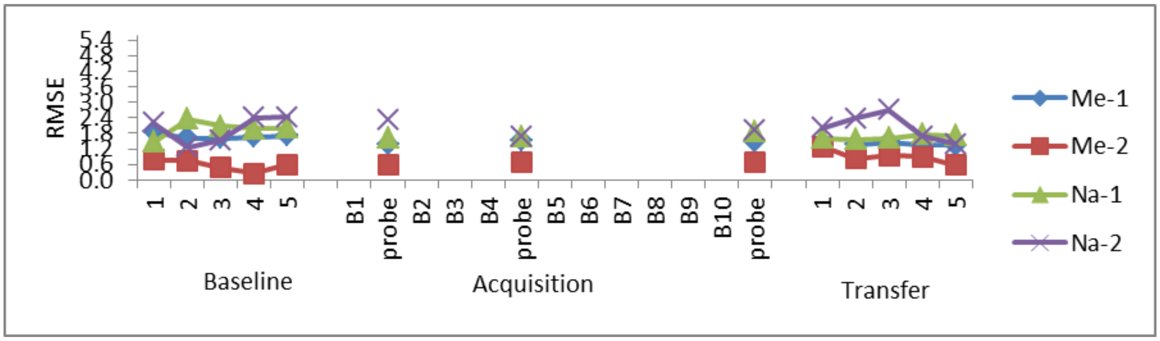


Figure 81 EFOA-Participant-46 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

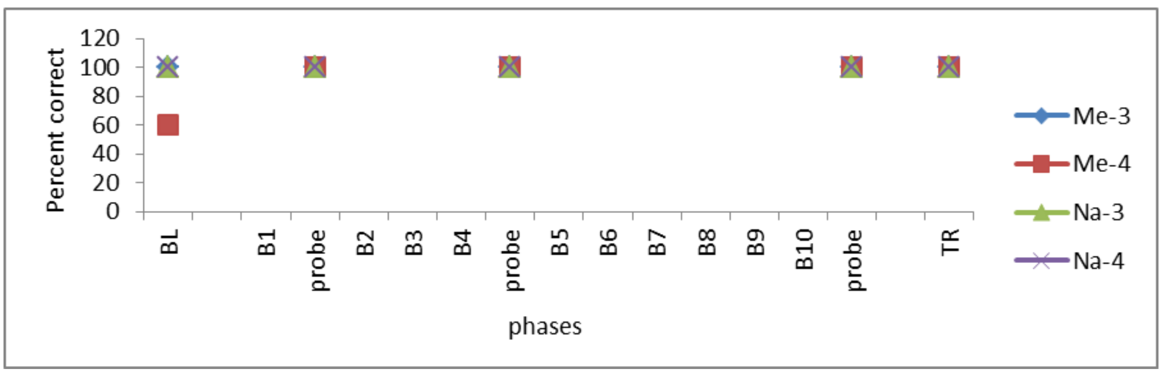


Figure 82 EFOA-Participant-46 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

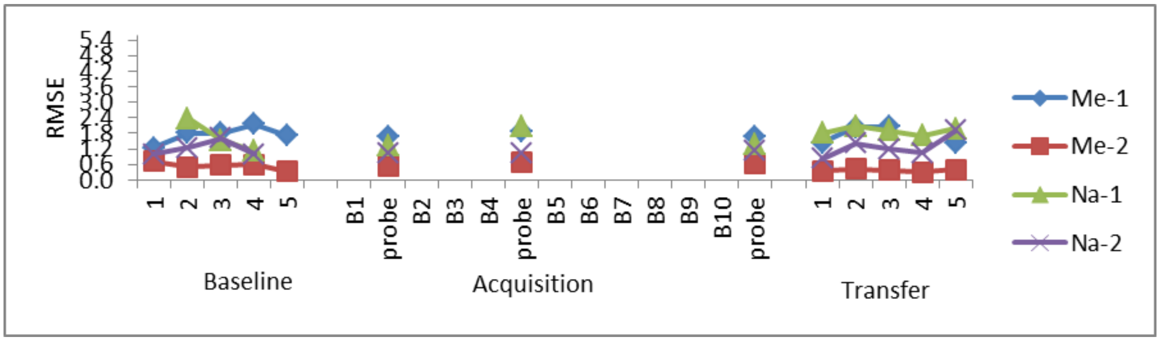


Figure 83 EFOA-Participant-55 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

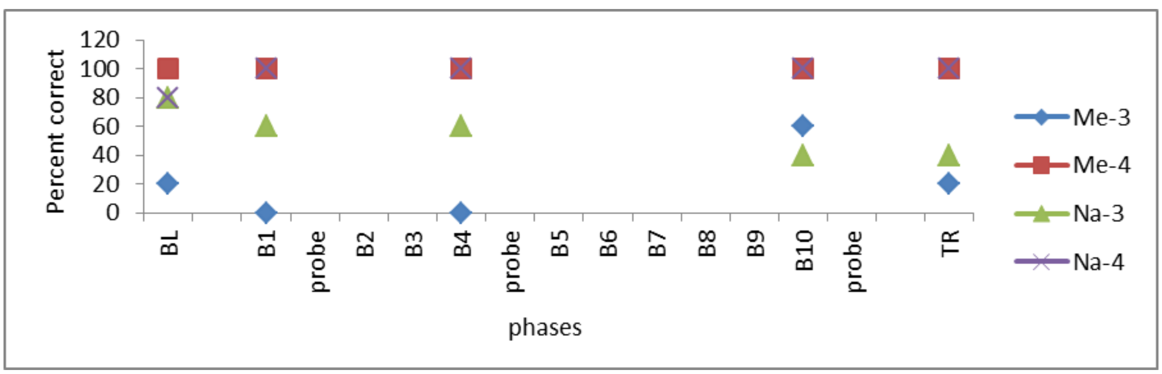


Figure 84 EFOA-Participant-55 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

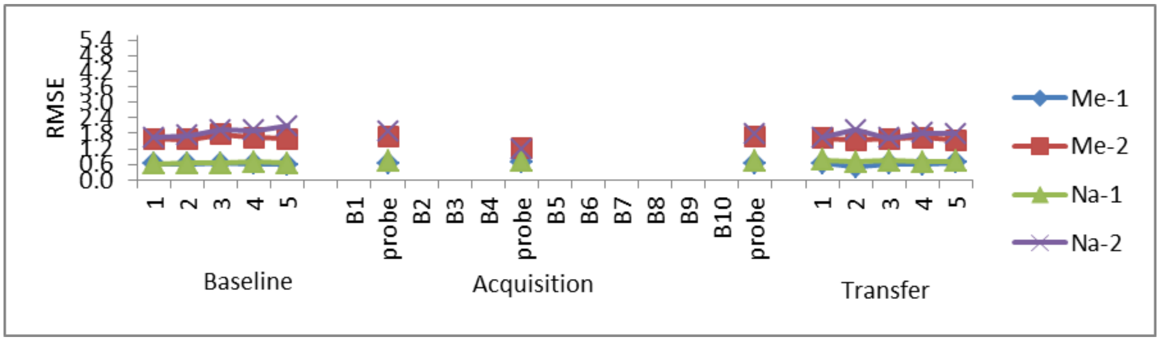


Figure 85 IFOA-Participant-2 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

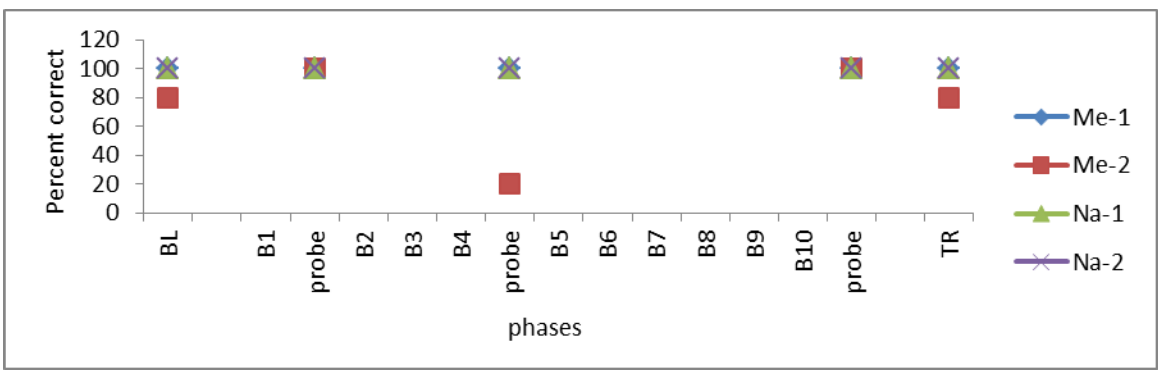


Figure 86 IFOA-Participant-2 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

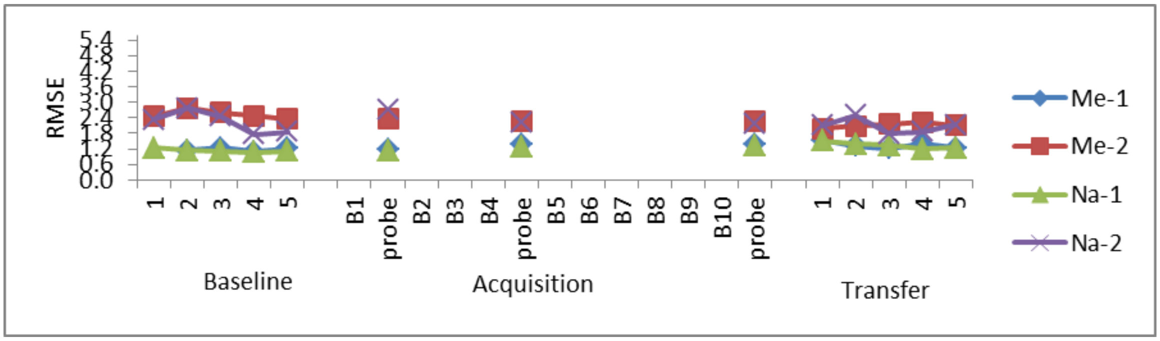


Figure 87 IFOA-Participant-5 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

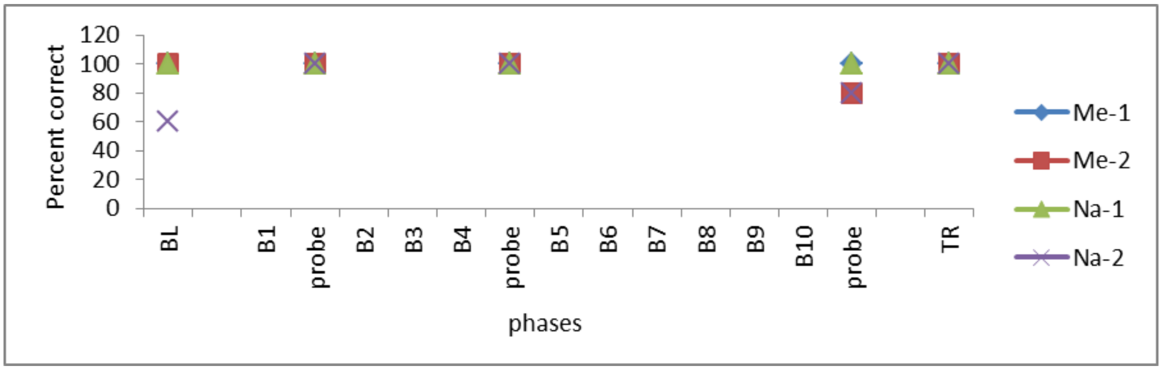


Figure 88 IFOA-Participant-5 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

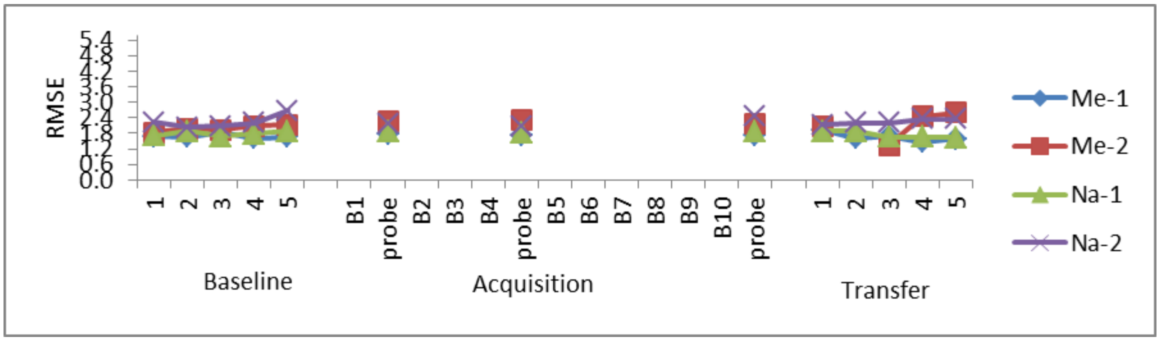


Figure 89 IFOA-Participant-11 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

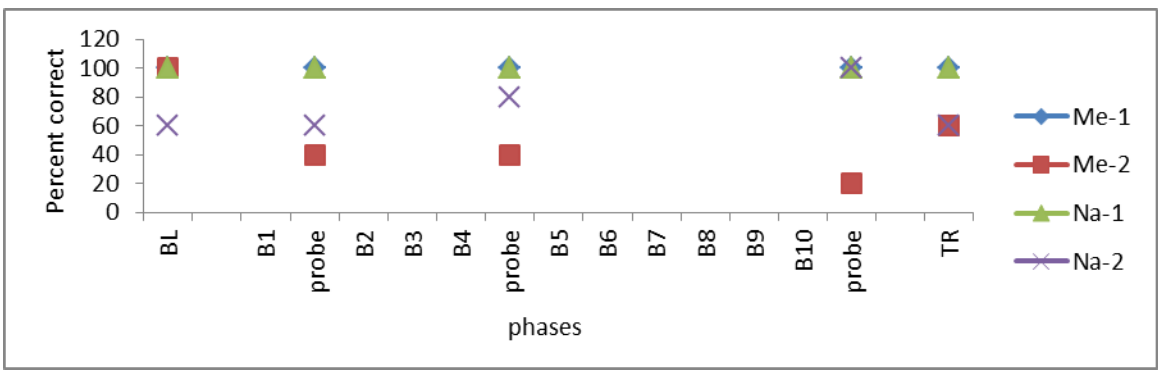


Figure 90 IFOA-Participant-11 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

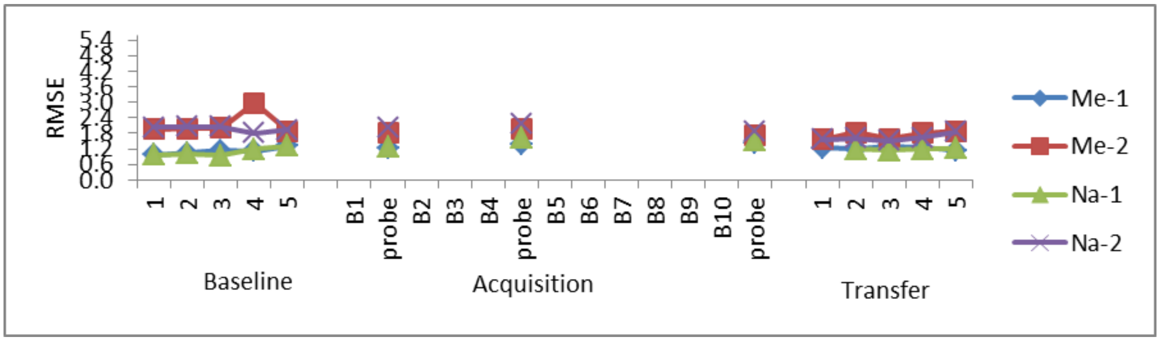


Figure 91 IFOA-Participant-27 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

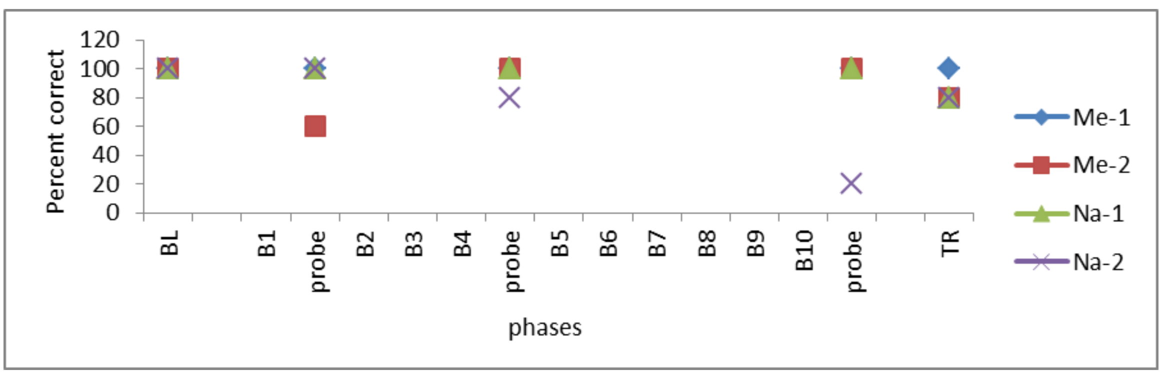


Figure 92 IFOA-Participant-27 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

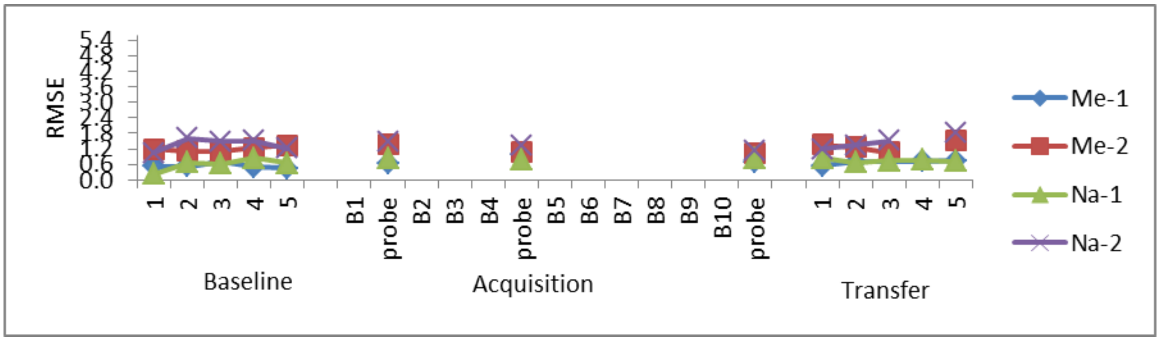


Figure 93 IFOA-Participant-34 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

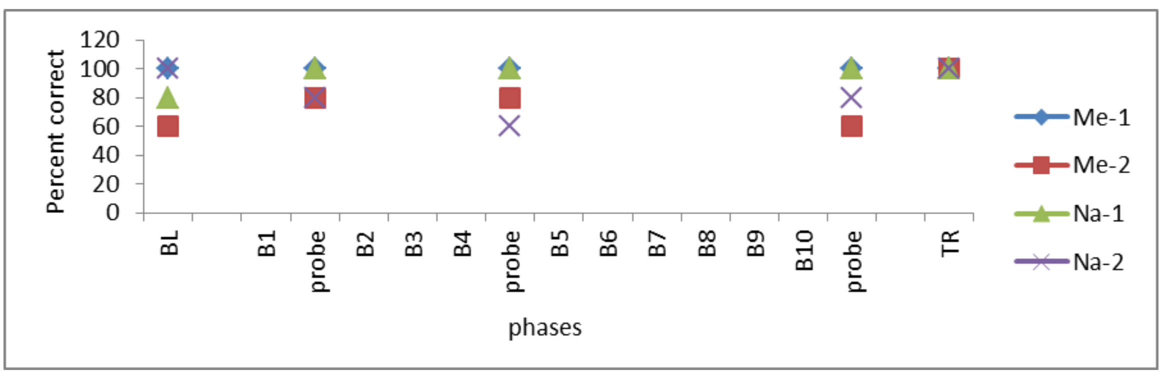


Figure 94 IFOA-Participant-34 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

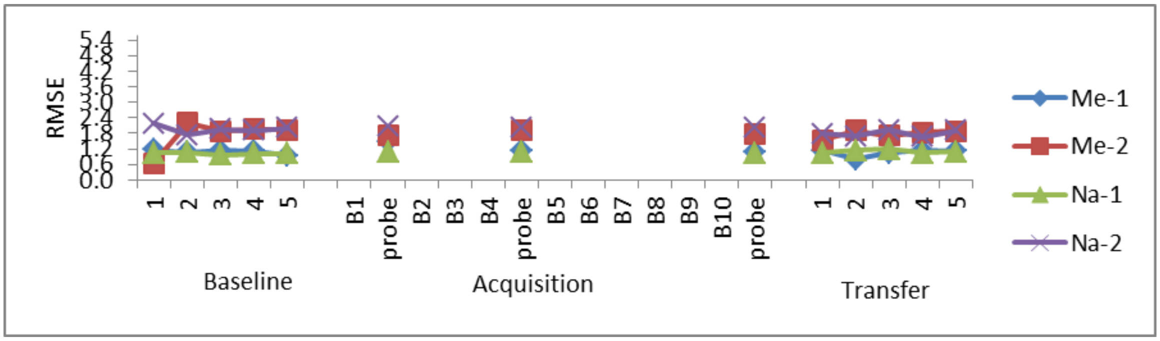


Figure 95 IFOA-Participant-39 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

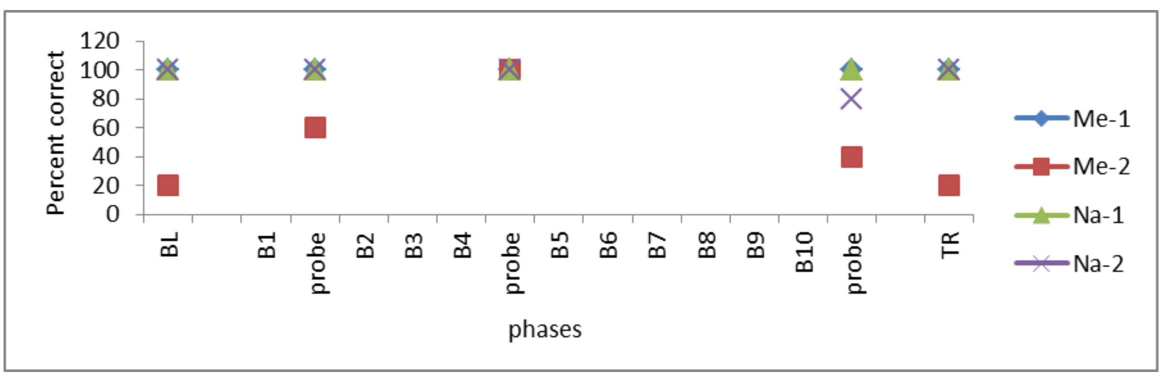


Figure 96 IFOA-Participant-39 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

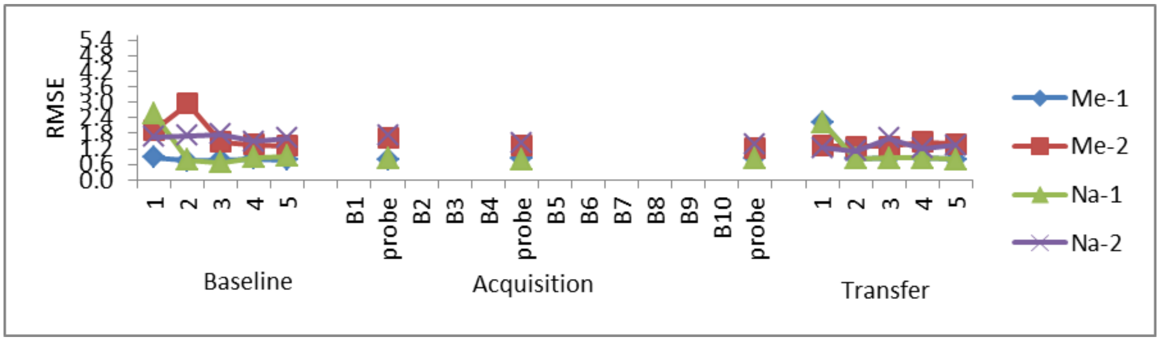


Figure 97 IFOA-Participant-44 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

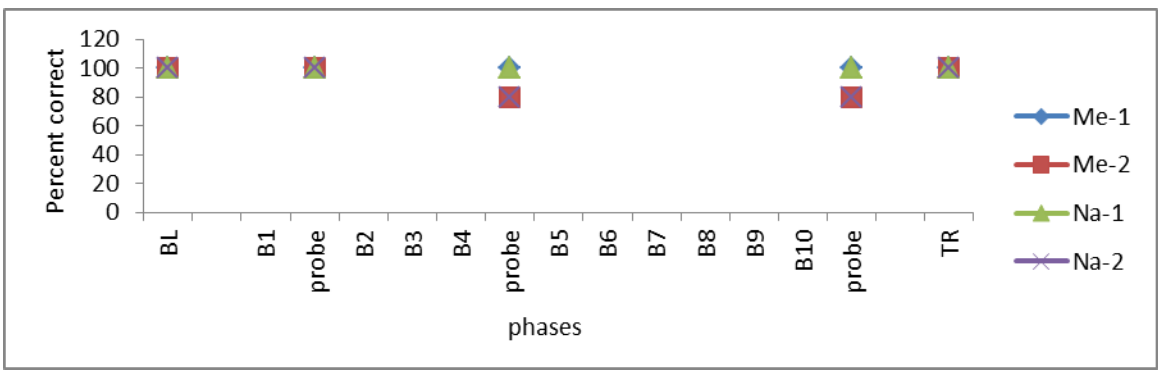


Figure 98 IFOA-Participant-44 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

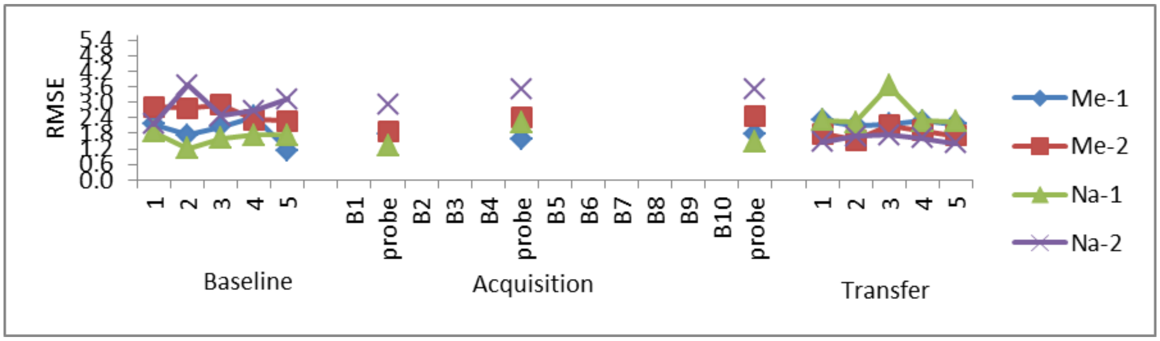


Figure 99 IFOA-Participant-7 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

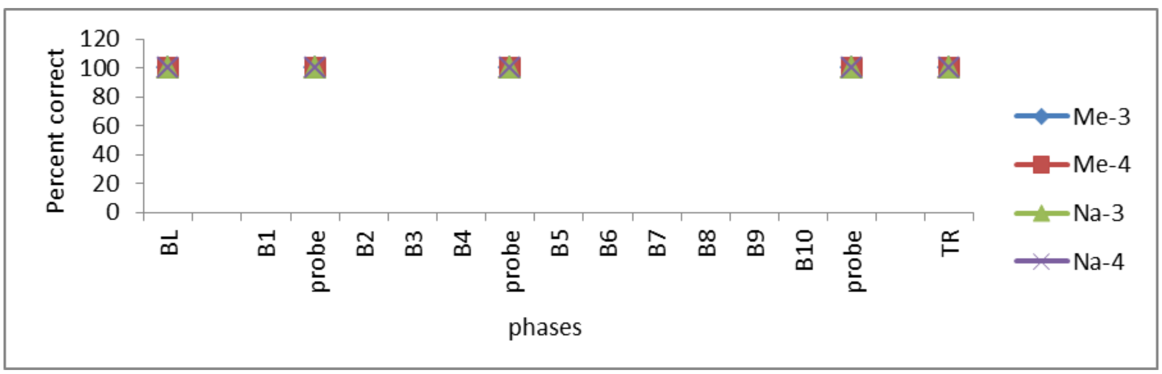


Figure 100 IFOA-Participant-7 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

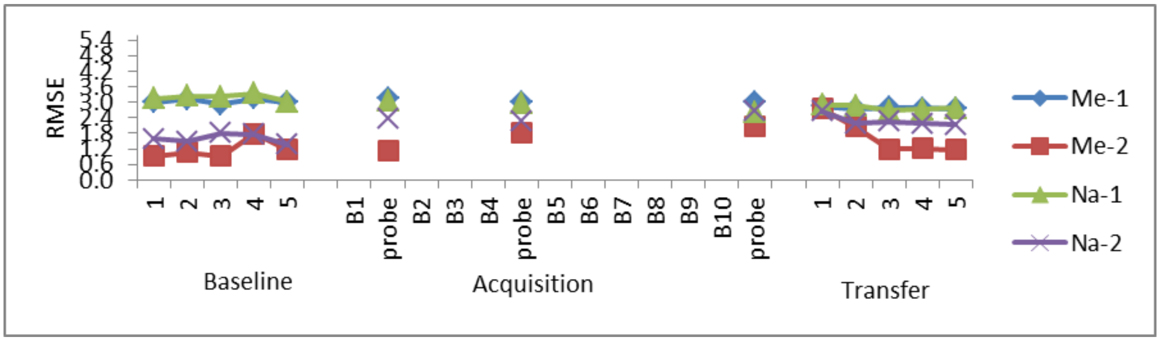


Figure 101 IFOA-Participant-14 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

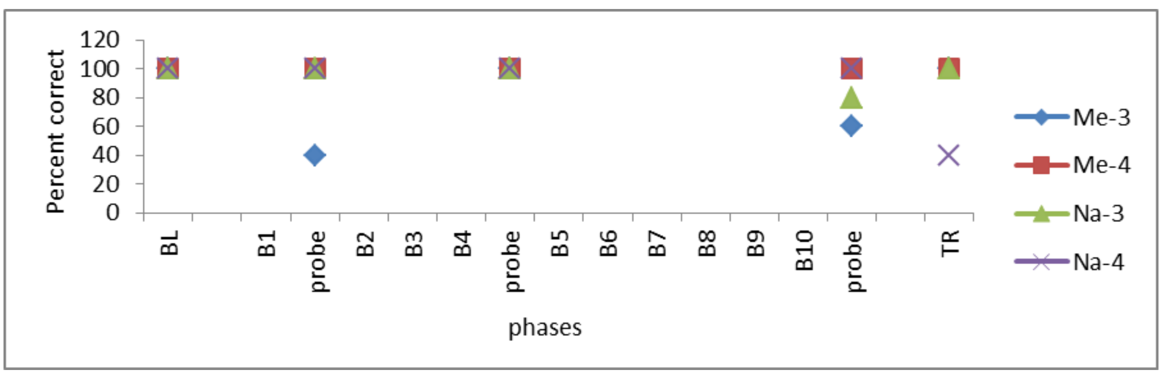


Figure 102 IFOA-Participant-14 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

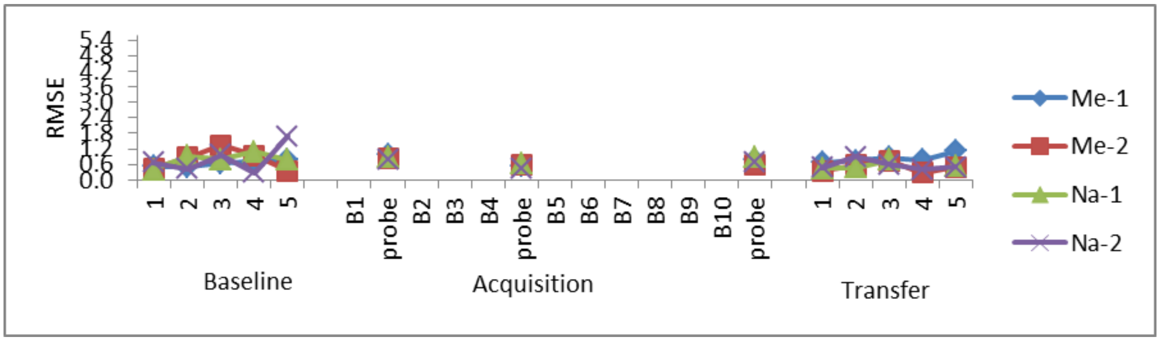


Figure 103 IFOA-Participant-24 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

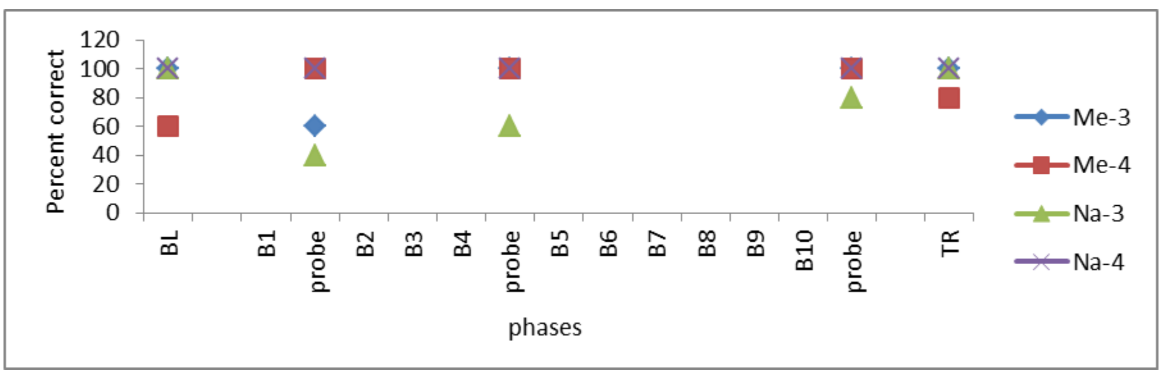


Figure 104 IFOA-Participant-24 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

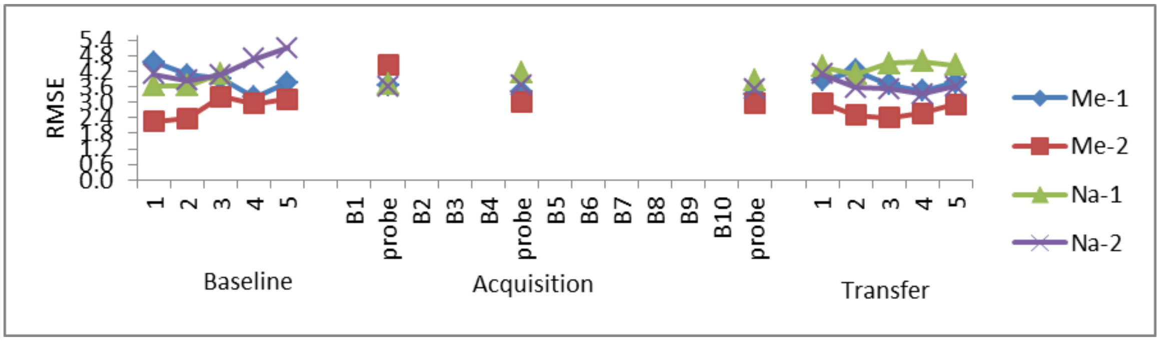


Figure 105 IFOA-Participant-26 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

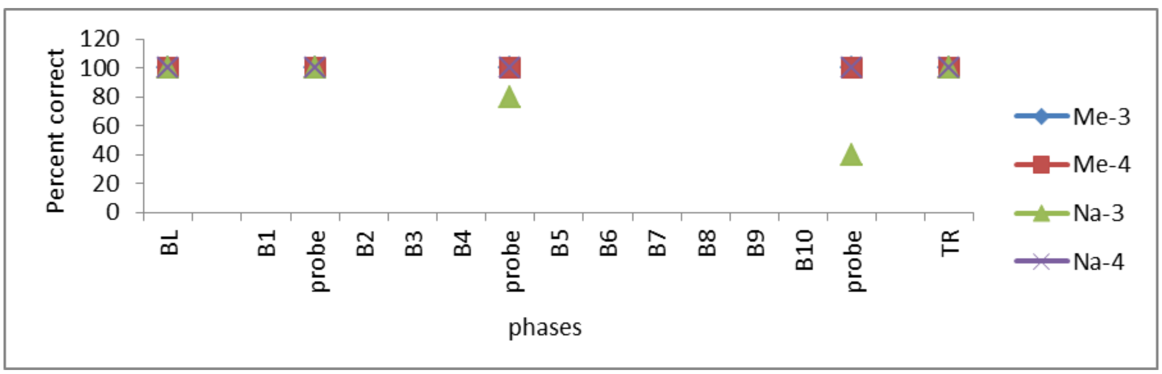


Figure 106 IFOA-Participant-26 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

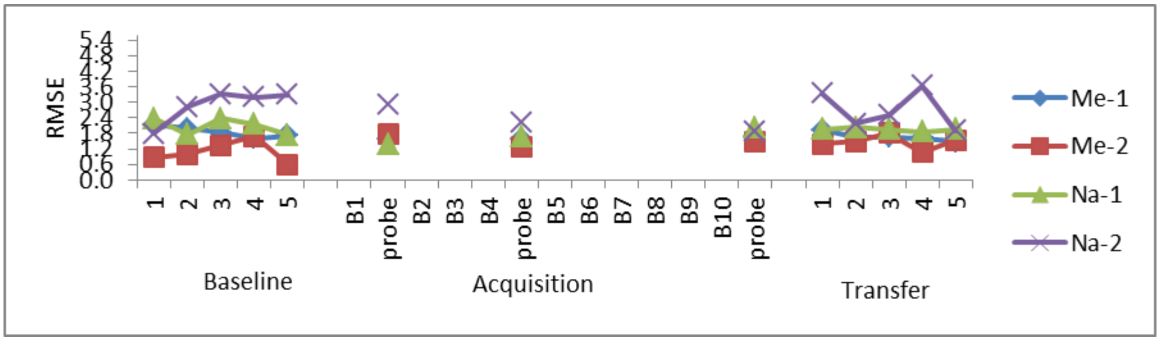


Figure 107 IFOA-Participant-40 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

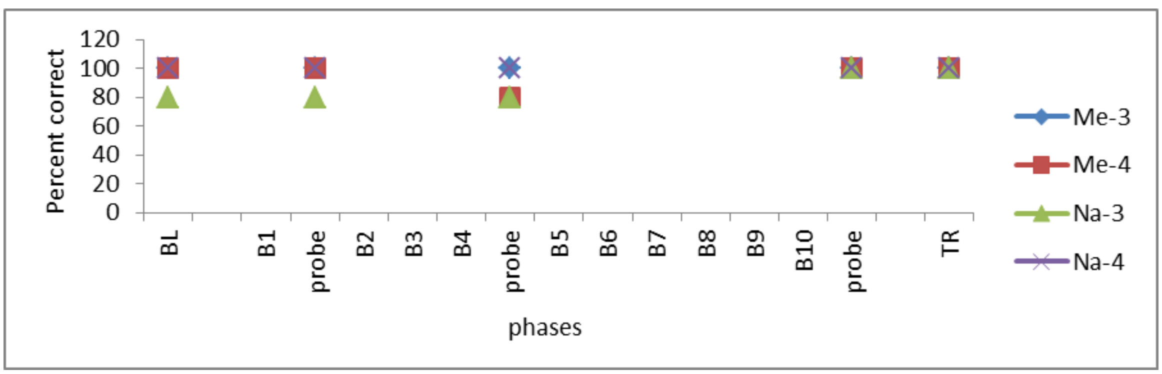


Figure 108 IFOA-Participant-40 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

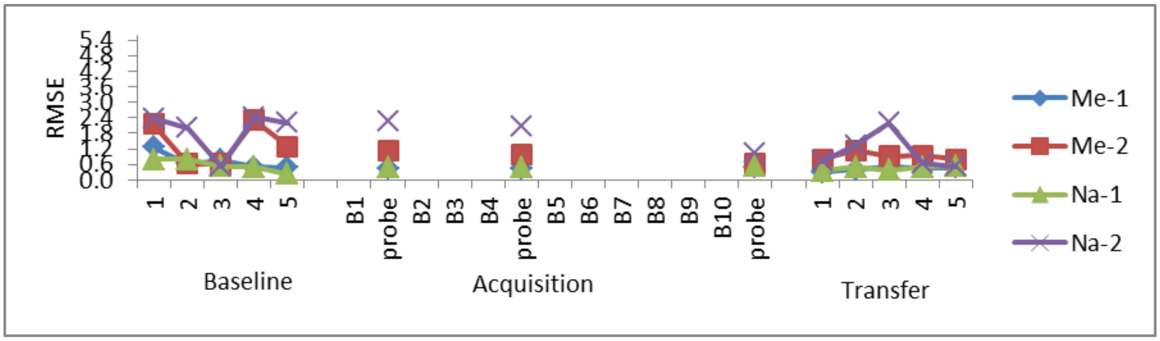


Figure 109 IFOA-Participant-41 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

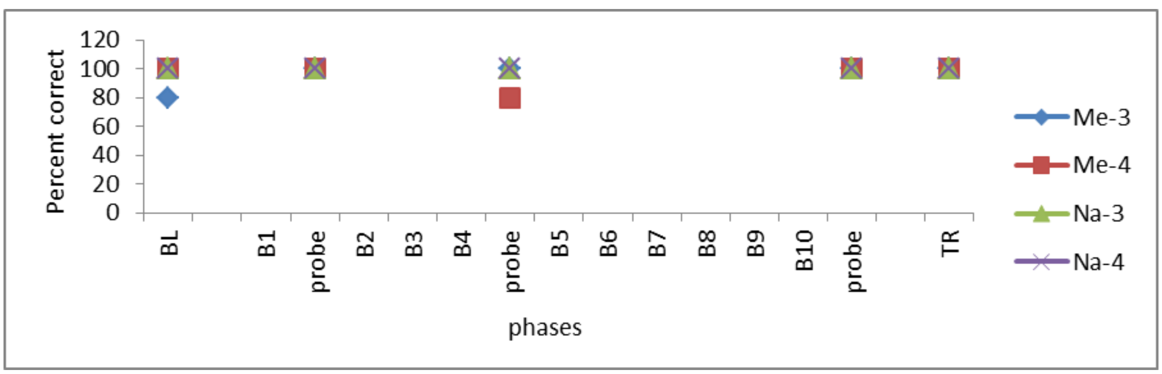


Figure 110 IFOA-Participant-41 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

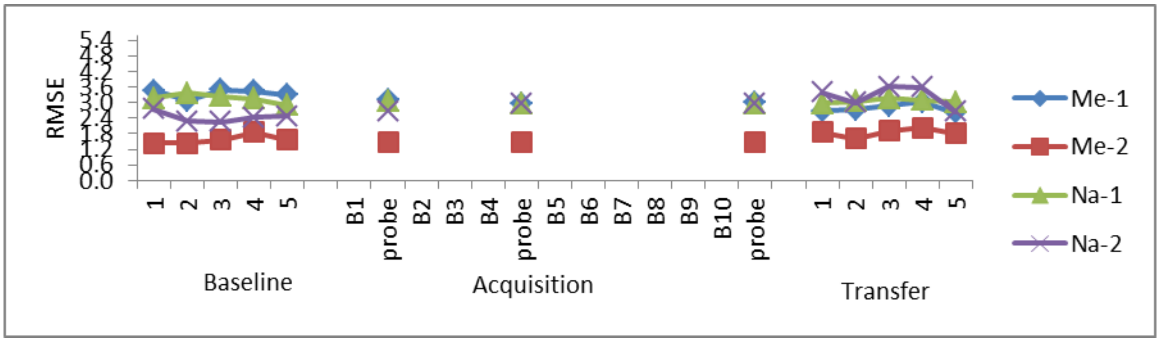


Figure 111 IFOA-Participant-50 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

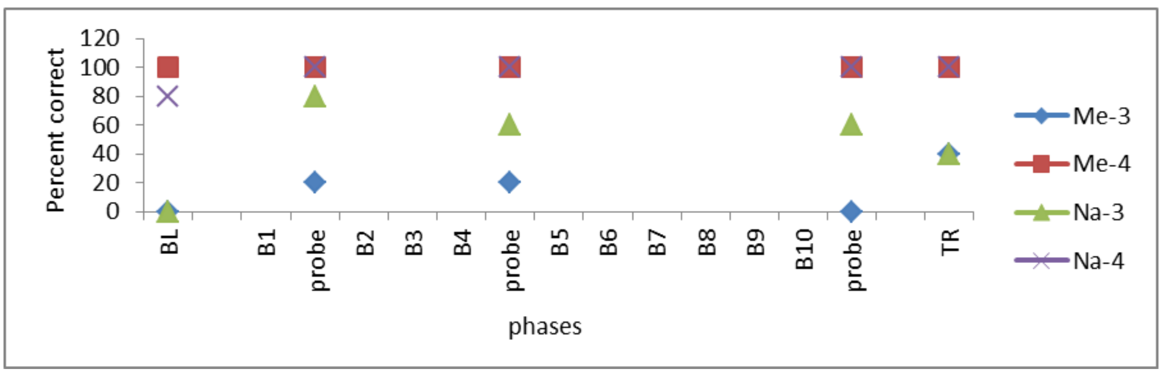


Figure 112 IFOA-Participant-50 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

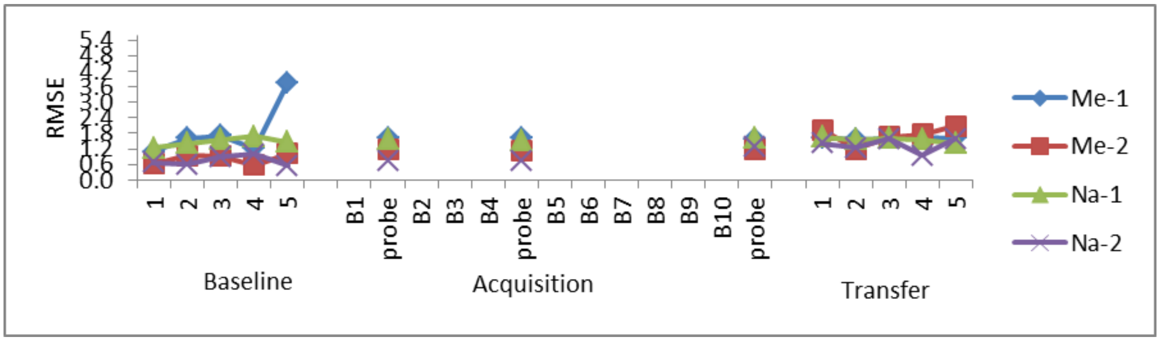


Figure 113   Control group-Participant-3 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

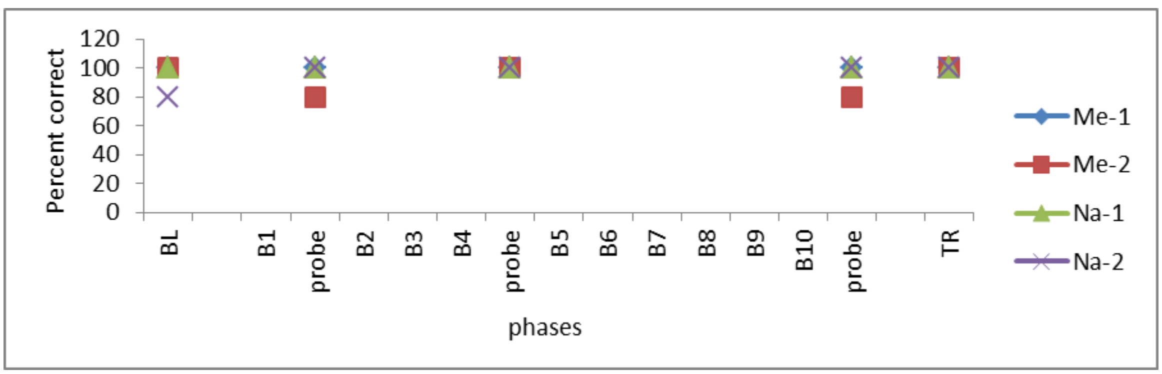


Figure 114   Control group-Participant-3 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

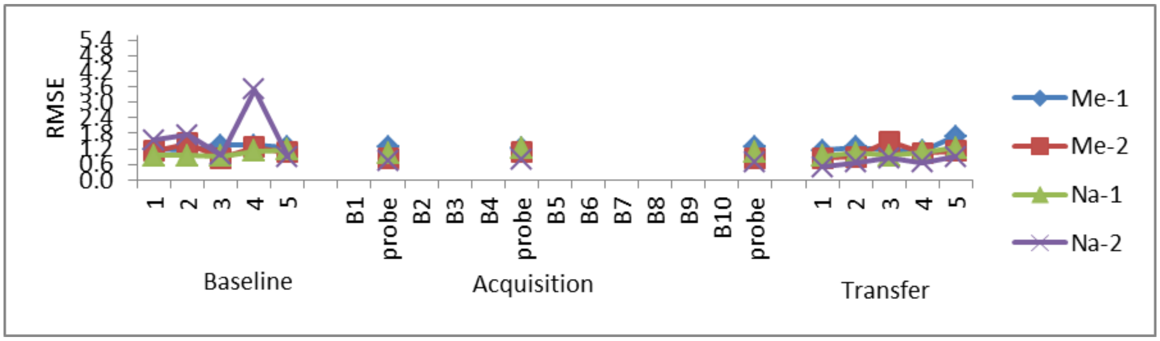


Figure 115 Control-group-Participant-4 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

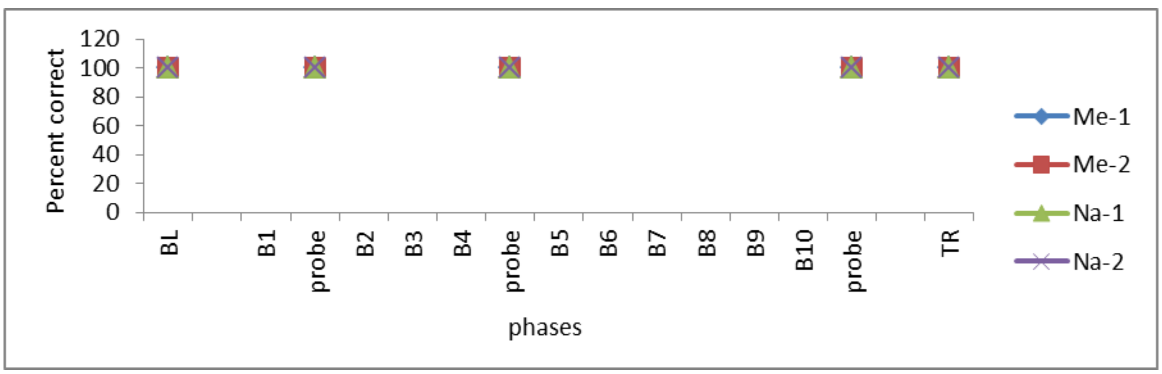


Figure 116 Control-group-Participant-4 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

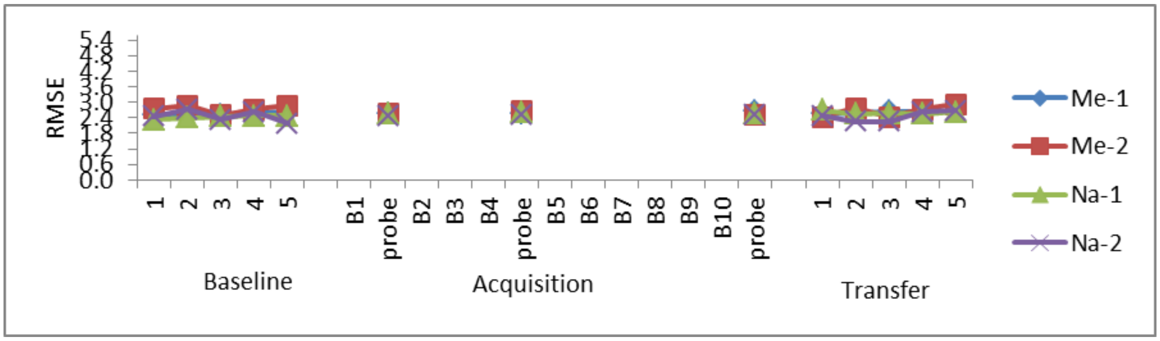


Figure 117 Control-group-Participant-28 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

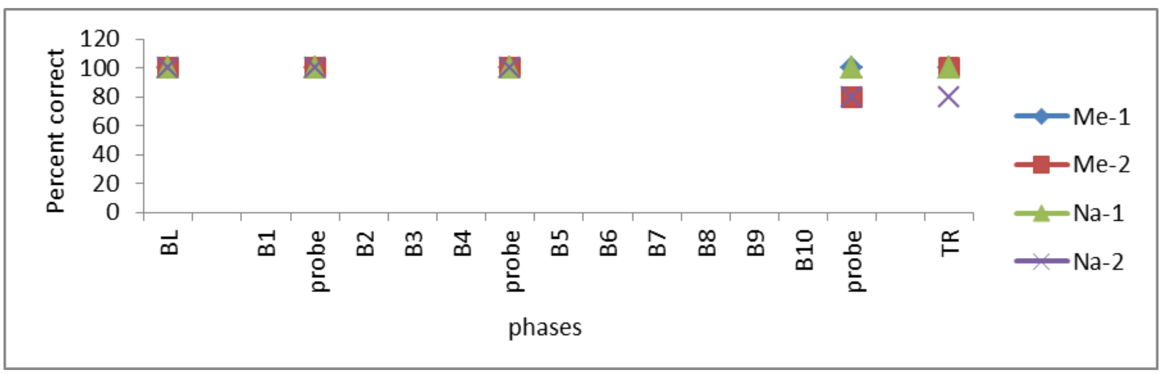


Figure 118 Control-group-Participant-28 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

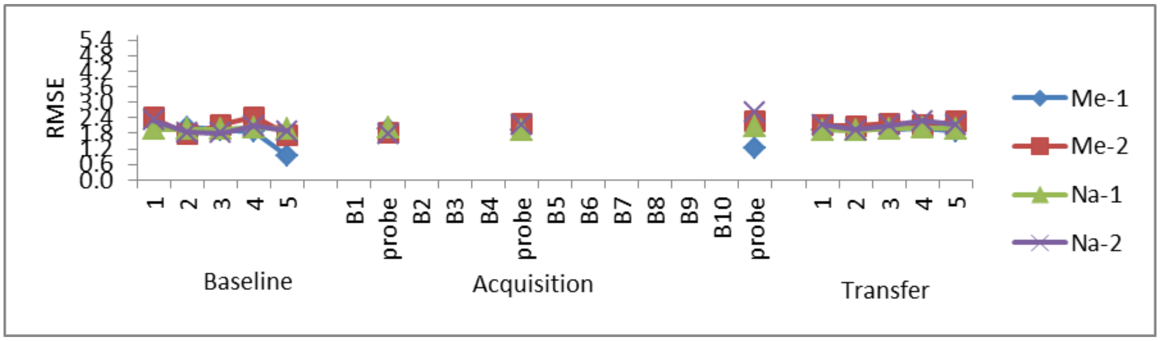


Figure 119 Control-group-Participant-30 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

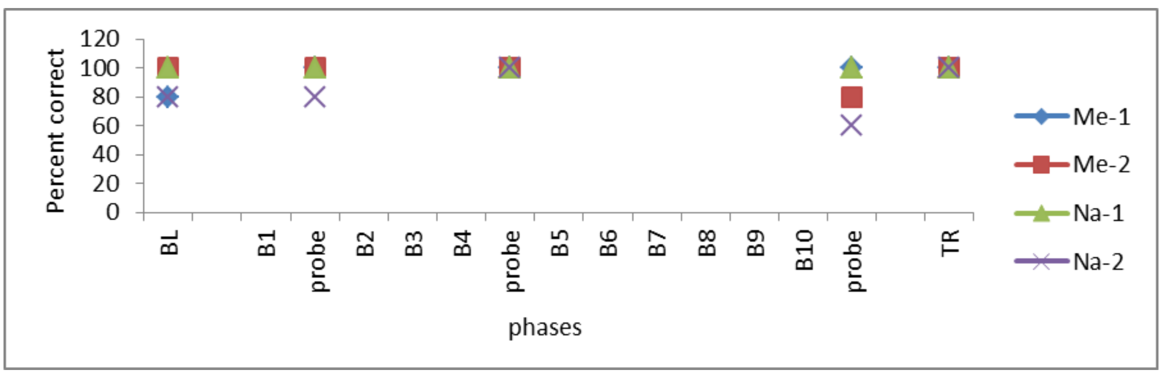


Figure 120 Control-group-Participant-30 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

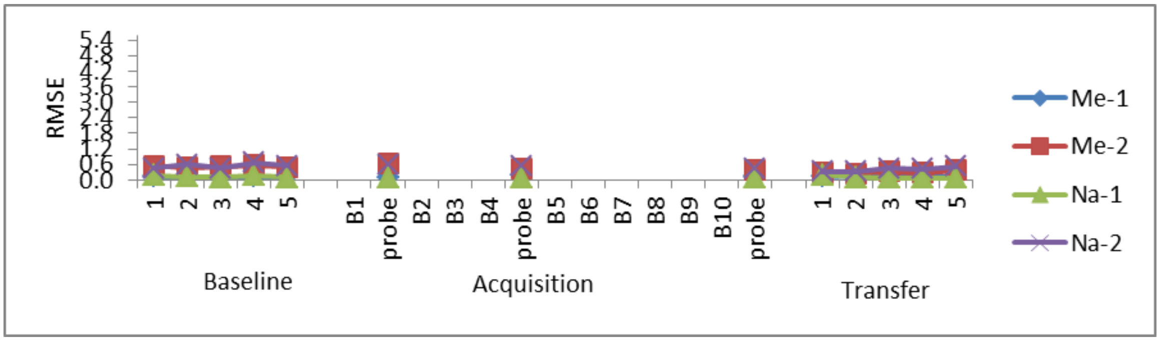


Figure 121 Control-group-Participant-31 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

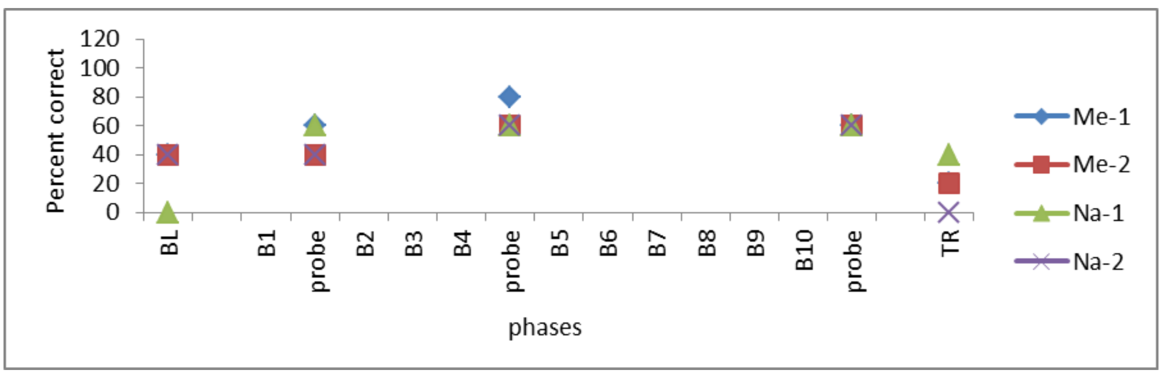


Figure 122 Control-group-Participant-31 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

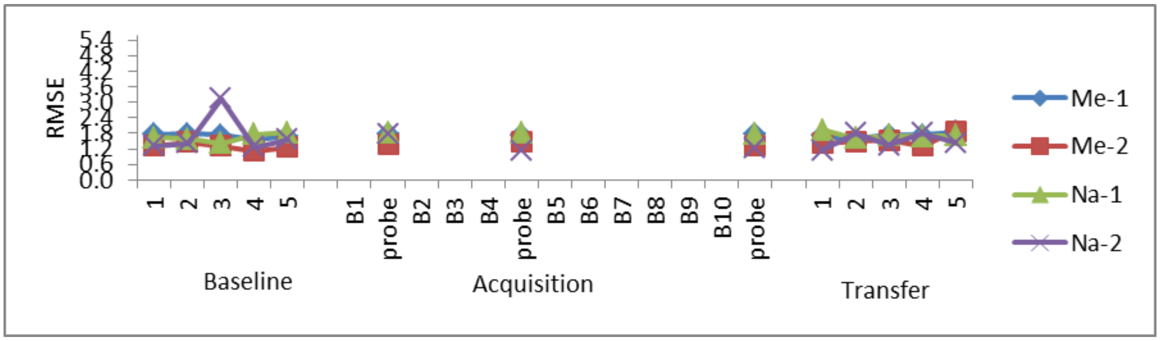


Figure 123 Control-group Participant-37 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

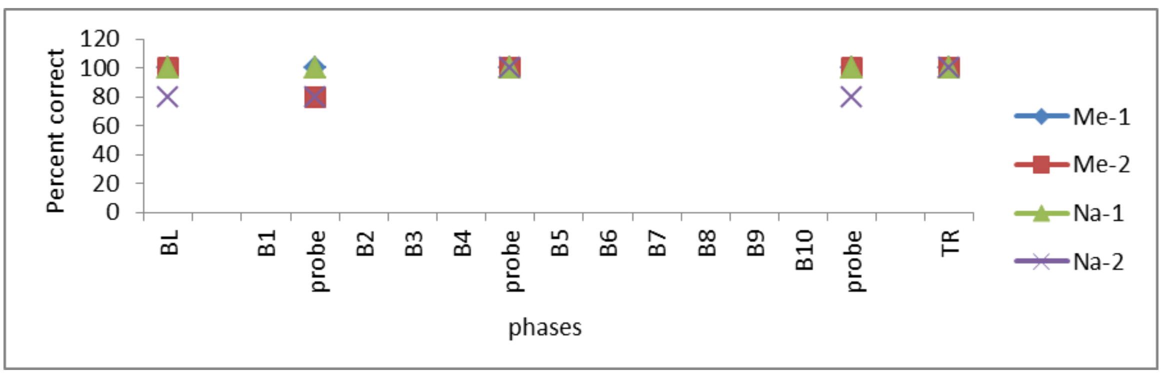


Figure 124 Control-group-Participant-37 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T1 and T2

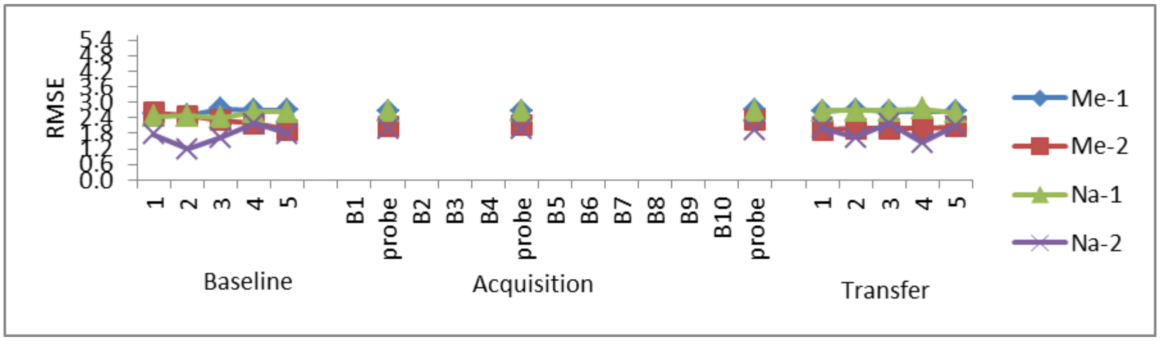


Figure 125 Control-group-Participant-48 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

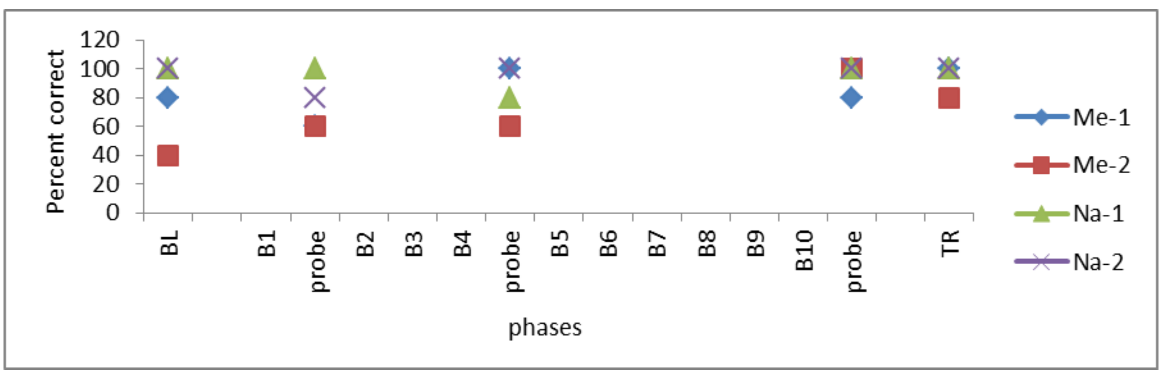


Figure 126 Control-group-Participant-48 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2)during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

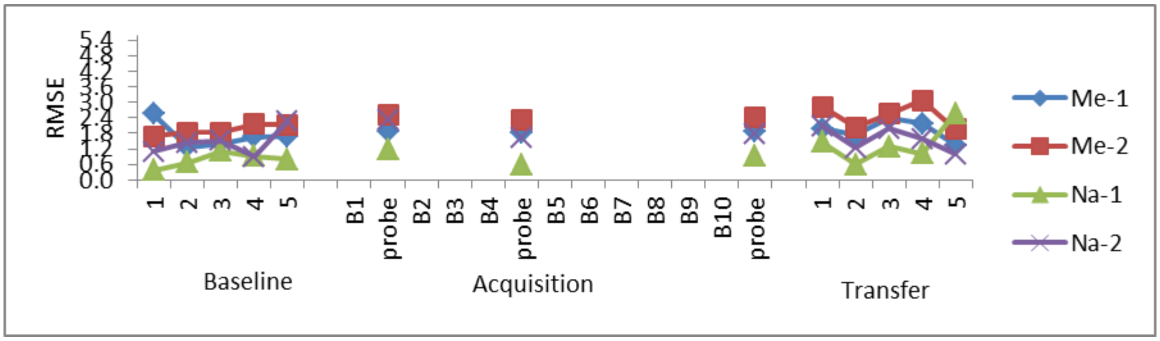


Figure 127   Control-group-Participant-12 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

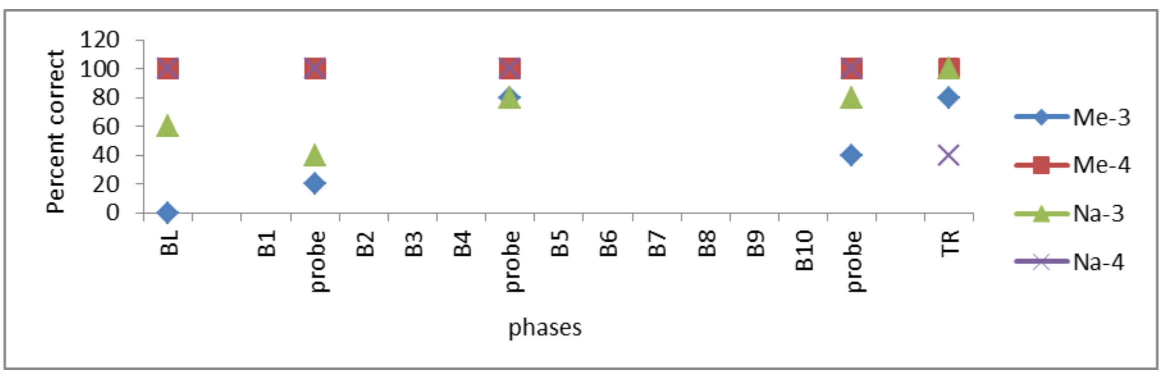


Figure 128   Contro-group-Participant-12 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

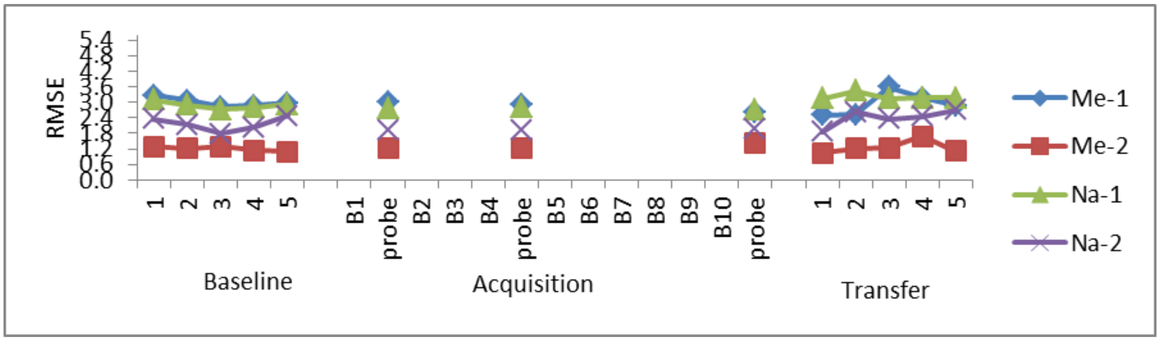


Figure 129 Contro-group-Participant-16 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

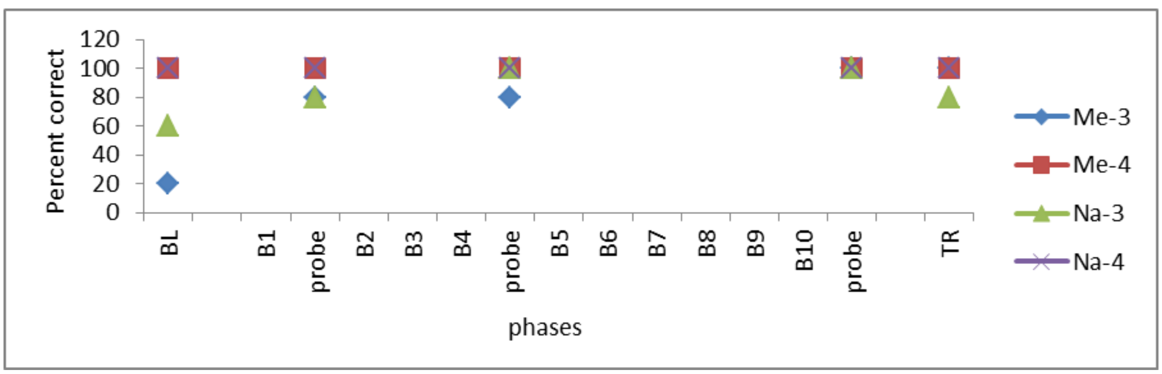


Figure 130 Control-group-Participant-16 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

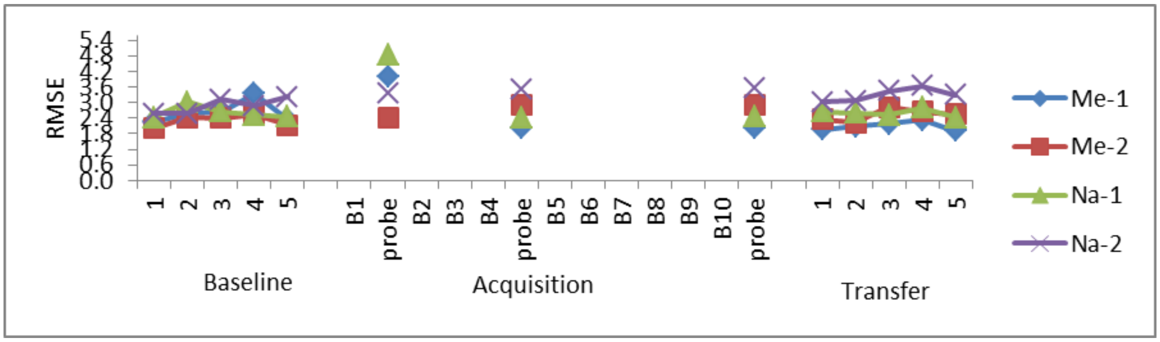


Figure 131 Control-group-Participant-18 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

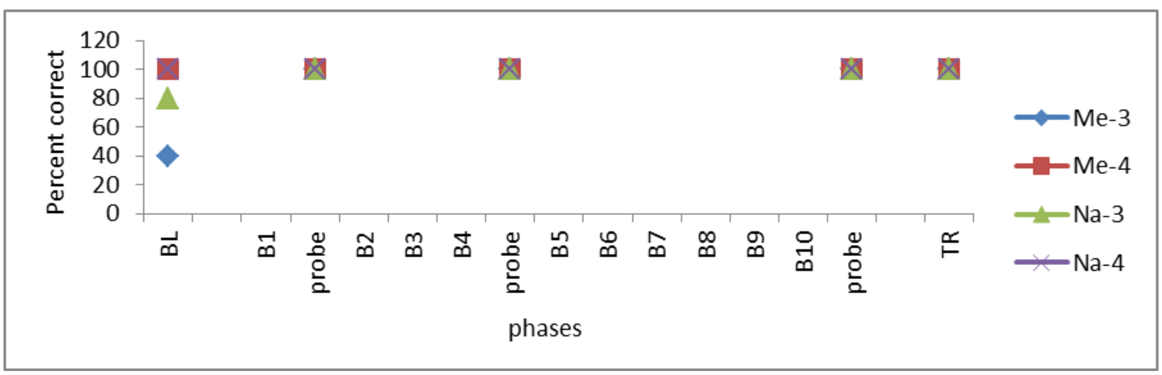


Figure 132 Control-group-Participant-18 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4:

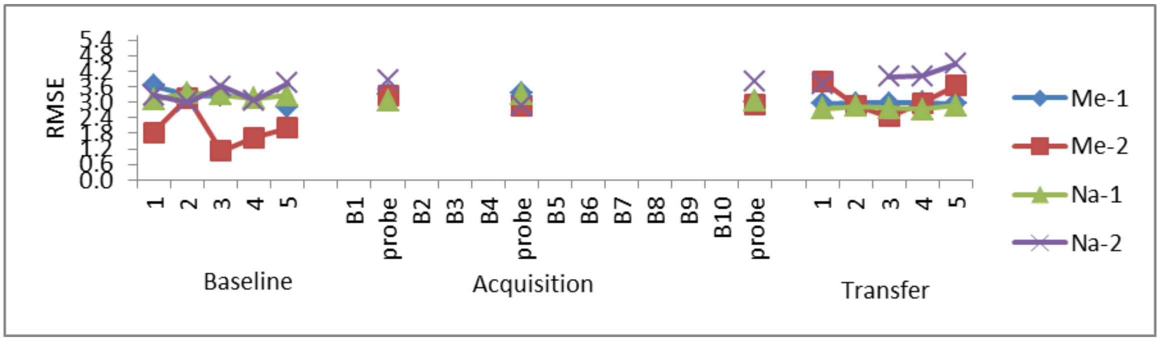


Figure 133 Control-group-Participant-19 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2)during baseline, acquisition phase, and transfer test.

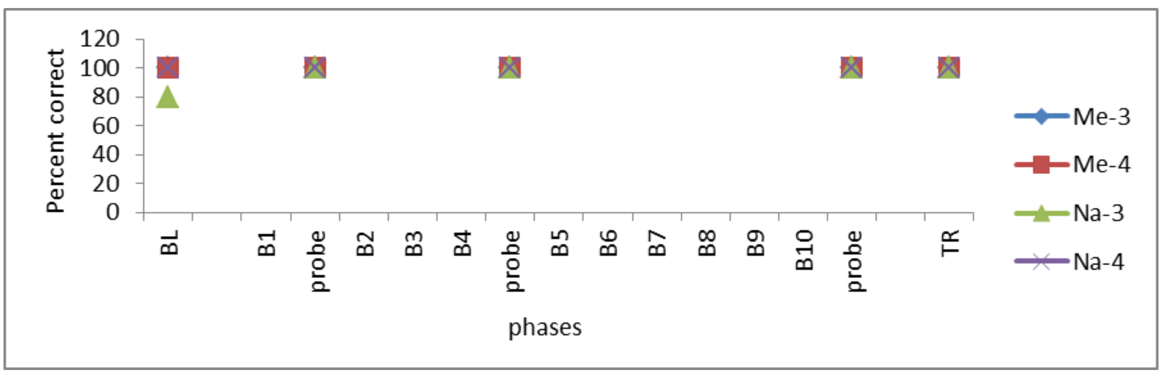


Figure 134 Control-group-Participant-19 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

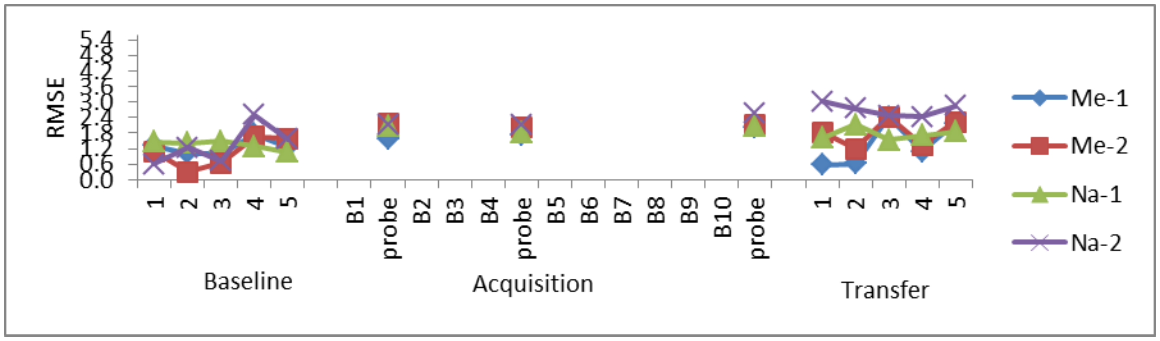


Figure 135 Control-group-Participant-21 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

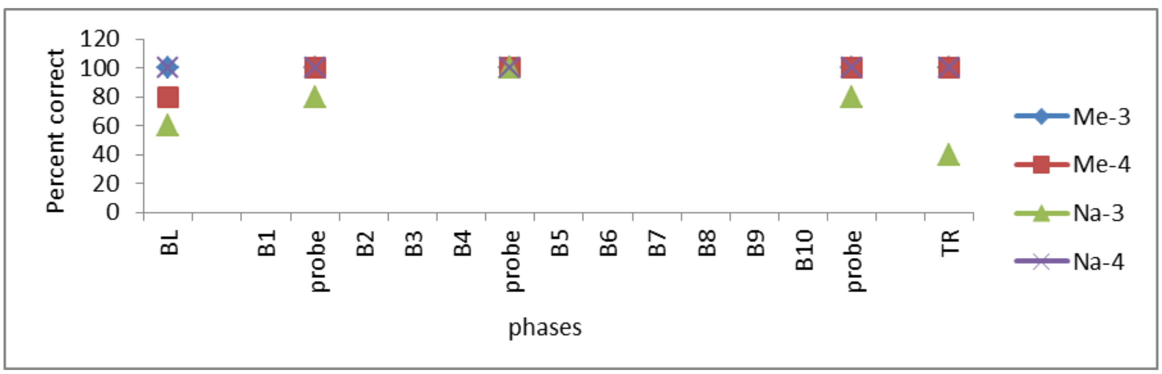


Figure 136 Control-group-Participant-21 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

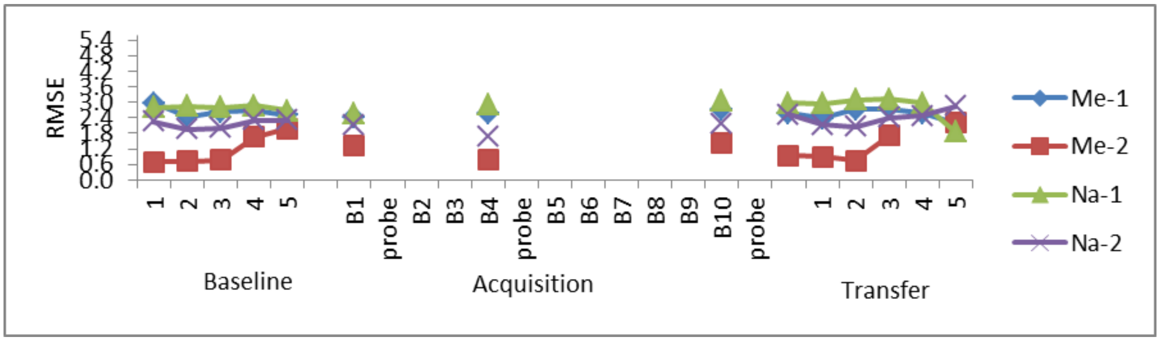


Figure 137 Control-group-Participant-23 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

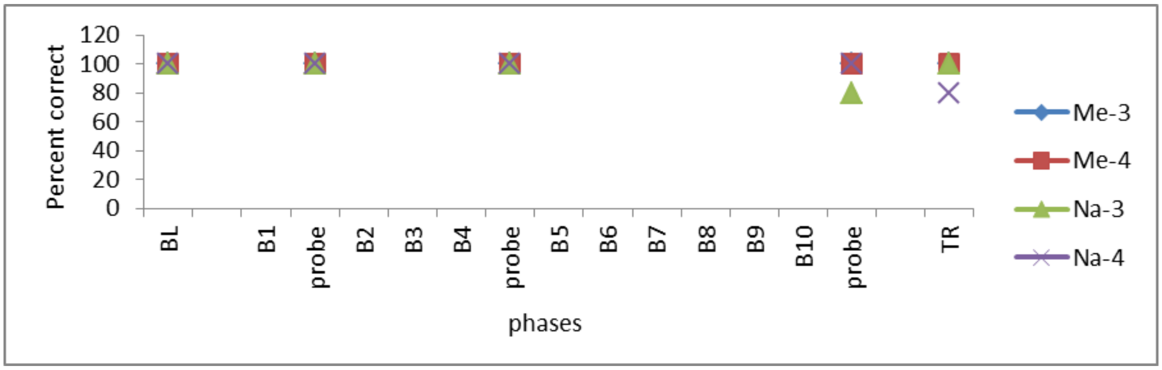


Figure 138 Control-group-Participant-23 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

Probes T3 and T4

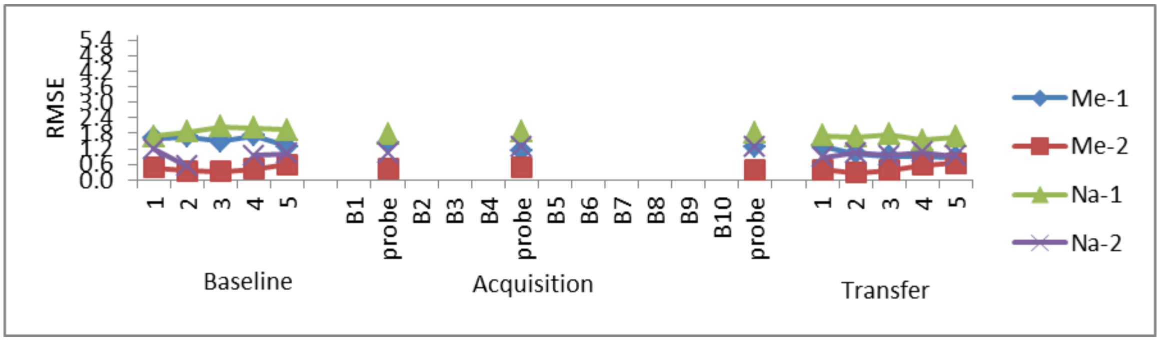


Figure 139 Control-group-Participant-35 The RMSE for the probed words (Me-1, Me-2, Na-1, Na-2) during baseline, acquisition phase, and transfer test.

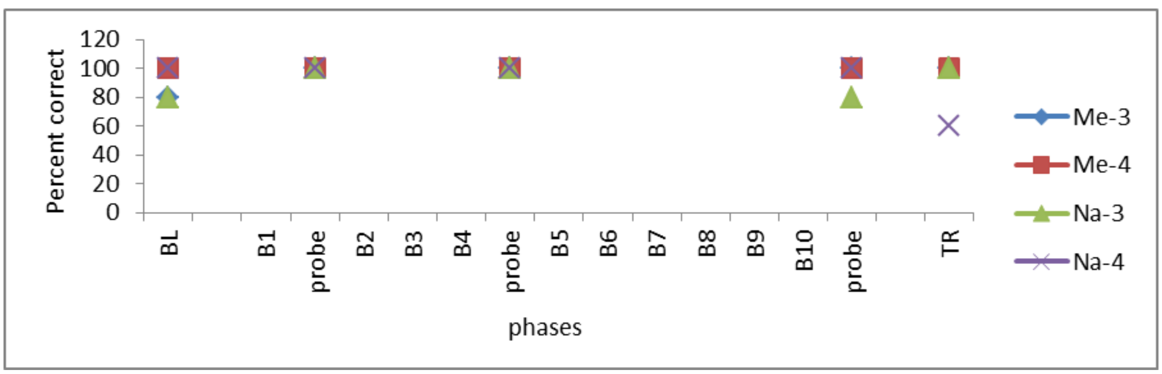


Figure 140 Control-group-Participant-35 The average of the percentage of correctly perceived probes productions (Me-1, Me-2, Na-1, Na-2) during baseline phase, acquisition phase, and transfer test. BL = baseline, B1-B10 = blocks 1-10, TR = transfer.

BIBLIOGRAPHY

Abberton, E., & Fourcin, A. (1975). Visual feedback and acquisition of intonation. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of Language Development (2nd ed.) (157–165).* New York Academic Press.

Abbs, J. H., Gracco, V. L., & Cole, K. J. (1984). Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *Journal of Motor Behavior, 16*, 195-231.

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3*, 111-150.

Adams, S., & Page, A. (2000). Effects of selected practice and feedback variables on speech motor learning. *Journal of Medical Speech-Language Pathology, 8*, 215-220.

Adams, S., Page, A., & Jog, M. (2002). Summary feedback schedules and speech motor learning in Parkinson's disease. *Journal of Medical Speech-Language Pathology, 10*, 215-220.

Adelman, Brown, & Quesada (2006) SUBTLEXWF. Retrived from http://expsy.ugent.be/subtlexus.

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88*, 148-159.

Albertson, K. (1982). Teaching pronunciation with visual feedback. *National Association of Language Laboratory Directors journal 17*, 18-33.

Anderson, D. I., Magill, R. A., Sekiya, H., & Ryan, G. (2005). Support for an explanation of the guidance effect in motor skill learning. *Journal of Motor Behavior, 37*, 231-238.

Anderson, F. (1960). An Experimental Pitch Indicator for Training Deaf Scholars. *Journal of the Acoustical Society of America, 32*, 1065-1074.

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*, 1088–1113.

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21*, 627–635.

Baguley, T. (2004). Understanding statistical power in the context of applied research. *Applied Ergonomics 35*, 73–80

Ballard, K., Granier., & Robin, D. (2000). Understanding the nature of apraxia of speech: Theory, analysis, and treatment. *Aphasiology, 14*, 969-995.

Ballard, K., Maas, E., & Robin, D. (2007). Treating control of voicing in apraxia of speech with variable practice. *Aphasiology, 14*, 969–995

Ballard, K., Robin, D., & Folkins, J. (2003). An integrative model of speech motor control: a response to Ziegler. *Aphasiology, 17*, 37-48.

Ballard, K., Smith, H., Paramatmuni, D., McCabe, P., Theodoros D & Murdoch B (2012). Amount of Kinematic feedback Affects learning of speech motor skills. *Motor Conrol*. 16, 106-119.

Barlow, D., & Hersen, M. (1984). *Single case experimental design: strategies for studying behavior change (2nd. ed)* pergamon press.

Bradshaw, E., & McHenry, M. (2005). Pitch Discrimination and Pitch Matching Abilities of Adults who Sing Inaccurately. *Journal of Voice, 19*, 431-439.

Brownman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetics, 49*, 155-180.

Brunner, J., Hoole, P., & Perrier, P. (2007). Speech motor learning as a process between acoustic target achievement and movement optimisation [Electronic Version]. Retrieved Retrieved July 10,2010, from http://summerschool2007.danielpape.info/abstract/bru.pdf.

Brydges, R., Dubrowski, A., & Carnahan, H. (2007). Attentional focus and motor learning: possible applications to health professions education. *E-Journal Bewegung and Training, 1*, 15-16. Retrieved October 30, 2009, from http://www.ejournal-but.de/doks/wulf_2007.pdf.

Bund, A., Wiemeyer, J., & Angert, R. (2007). Attentional focus and motor learning: notes on some problems of a research paradigm. *E-Journal Bewegung and Training, 1*, 17-18. Retrieved October 30, 2009, from http://www.ejournal-but.de/doks/wulf_2007.pdf

Busk, P. L., & Serlin, R. (1992). Meta-analysis for single case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single Case Research Design and Analysis: New Directions for Psychology and Education*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chen, G. T. (1974). The pitch range of English and Chinese speakers. *Journal of Chinese Linguistics, 2*, 159-171.

Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *The Journal of the Chinese Language Teachers Association, 32*, 21-39.

Chiviacowsky S, Wulf G, & R., W. (2010). An external focus of attention enhances balance learning in older adults. *Gait Posture, 32*, 572-575.

Chun, D. (1989). Teaching Tone and Intonation with Microcomputers. *CALICO Journal, 7*, 21-46.

Clark, J., Yallop, C., & Fletcher, J. (2007). *An Introduction to Phonetics and Phonology (3rd. ed)*. Oxford:: Wiley-Blackwell.

Cohen, E. (2010). *The effect of manipulation of attentional focus on temporospatial parameters of gait (PhD thesis).* Trident University International, Cypress, California.

De Bot, K. (1983). Visual feedback of intonation I: effectiveness and induced practice behavior. *Language and Speech 26*, 331-350.

Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *96*, 283-321.

Ding, H., Jokisch, O., & Rudiger, H. (2010). *Perception and Production of Mandarin Tones by German Speakers.* Retrieved from http://speechprosody2010.illinois.edu/papers/100153.pdf.

Duffy, J. R. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management (2nd. ed.).* St. Louis, MO.: Elsevier Mosby.

Durham, K., Van Vliet, P. M., Badger, F., & Sackley, C. (2009). Use of information feedback and attentional focus of feedback in treating a person with a hemiplegic arm. *Physiotherapy Research International*, *14*, 77-90.

Ehrlenspiel, F. (2007). Constraining action through attentional focusing happens at points in time *E-Journal Bewegung and Training*, *1*, 19-20. Retrieved October 30, 2009, from http://www.ejournal-but.de/doks/wulf_2007.pdf.

Fasoli, S. E., Trombly, C. A., Tickle-Degnen, L., & Verfaellie, M. H. (2002). Effect of Instructions on Functional Reach in Persons With and Without Cerebrovascular Accident. *American Journal of Occupational Therapy 56*, 380-390.

Ferrand, C. T. (2007). *Speech science : an integrated approach to theory and clinical practice (2nd ed.).* Boston : Pearson / Allyn and Bacon.

Fox, C. M., Morrison, C. E., Ramig, L. O., & Sapir, S. (2002). Current perspectives on the Lee Silverman Voice Treatment (LSVT) for individuals with idiopathic Parkinson disease. *American Journal of Speech-Language Pathology*, *11*, 111–123.

Freedman, S. E., Maas, E., Caligiuri, M. P., Wulf, G., & Robin, D. A. (2007). Internal versus external: oral-motor performance as a function of attentional focus. *Journal of Speech, Language and Hearing Research, 50*, 131-136.

Fry, D. B. (1979). *The physics of speech*: New York : Cambridge University Press.

Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*, 149-175.

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production.* (Vol. 1, pp. 177-220). London: Academic Press.

Garrett, M. (1982). Production of speech: Observations from normal and pathological use. In A. W. Ellis (Ed.), *Normality and Pathology in Cognitive Functions*. New York, NY: Academic Press, Inc.

Garrett, M. (1984). The organization of processing structure for language production: Application to aphasic speech. . In D. Caplan, A. R. Lecours & A. Smith (Eds.), *Biological perspective on Language* (pp. 172-193). Cambridge, MA: MIT Press.

Giuffrida, C. G., Shea, J. B., & Fairbrother, J. T. (2002). Differential transfer benefits of increased practice for constant, blocked, and serial practice schedules. *Journal of Motor Behavior, 34*, 353–365.

Gracco, V. L., & Abbs, J. H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research, 65*, 156-166.

Grigos, M (2009). Changes in articulator movement variability during phonemic development: A longitudinal study. *Journal of Speech, Language, and Hearing Research. 52*, 164–177.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior, 36*, 212-224.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*, 594-621.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*, 280-301.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*, 280-301.

Hall, K., & Magill, R. A. (1995). Variability of Practice and Contextual Interference in Motor Skill Learning. *Journal of Motor Behavior, 27*, 299-309.

Hermes, D. J. (1998). Auditory and visual similarity of pitch contours. *Journal of Speech, Language, and Hearing Research*, *41*, 63-72.

Hodges, N. J., & Franks, I. M. (2000). Attention focusing instructions and coordination bias: implications for learning a novel bimanual task. *Human Movement Science, 19*, 843-867.

Howie, J. M. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge: Cambridge University Press.

Katz, W. F., Bharadwaj, S. V., & Carstens, B. (1999). Electromagnetic Articulography Treatment for an Adult With Broca's Aphasia and Apraxia of Speech. *J Speech Lang Hear Res*, *42*(6), 1355-1366.

Katz, W. F., Carter, G. C., & Levitt, J. S. (2007). Treating buccofacial apraxia using augmented kinematic feedback. *Aphasiology, 21*, 1230–1247.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin, 70*, 387-403.

Keetch, K. M., Schmidt, R. A., Lee, T. D., & Young, D. E. (2005). Especial skills: Their emergence with massive amounts of practice. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 970–978.

Kelso, J., Saltzman, E., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics, 14*, 29-59.

Kelso, J. A. S., & Tuller, B. (1981). Toward a theory of apractic syndromes. *Brain and Language, 12*, 224-245.

Kent, R., Adams, S., & Turner, G. (1996). Models of speech production. In N. J. Lass (Ed.), *Principles of Experimental Phonetics* (pp. 3-45). St. Louis, MO: Mosby.

Kent, R., & Read, C. (2002). *Acoustic analysis of speech (2nd. ed.)*: Singular,Thompson Learning.

Kent, R. D. (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of Communication Disorders*, *33*, 391-428.

Kim, I. (2007). Effect of feedback schedules and number of practice trials on motor retention of novel speech behaviors *(Doctoral dissertation).* The Florida State University, Florida.

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*, 135-150.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*, 820-857.

Klugman, T., & Ross, E. (2002). Perceptions of the Impact of Speech, Language, Swallowing, and Hearing Difficulties on Quality of Life of a Group of South African Persons with Multiple Sclerosis. *Folia Phoniatrica et Logopaedica*, *54*, 201-221.

Knock, T. R., 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*, 653-668.

Lai, Q., & Shea, C. H. (1998). Generalized motor program (GMP) learning: Effects of reduced frequency of knowledge of results and practice variability. *Journal of Motor Behavior, 30*, 51–59.

Lai, Q., Shea, C. H., Wulf, G., & Wright, D. L. (2000). Optimizing generalized motor program and parameter learning. *Research Quarterly for Exercise and Sport, 71*, 10–24.

Laufer, Y., Rotem-Lehrer, N., Ronen, Z., Khayutin, G., & Rozenberg, I. (2007). Effect of Attention Focus on Acquisition and Retention of Postural Control Following Ankle Sprain. *Archives of Physical Medicine and Rehabilitation, 88*, 105-108.

Leathers, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech (p.305-341)*. Amsterdam: University of Amsterdam

Lee, T., Wulf, G., & Schmidt, R. (1992). Contextual Interference in Motor Learning: Dissociated Effects Due to the Nature of Task Variations. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 44*, 627-644.

Lee, T. D., & Genovese, E. D. (1989). Distribution of practice in motor skill acquisition: Different effects for discrete and continuous tasks. *Research Quarterly for Exercise and Sport, 60*, 59–65.

Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 730–746.

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control* (pp. 3–22). Amsterdam: Elsevier.

Lee, T. D., Magill, R. A., & Weeks, D. J. (1985). Influence of practice schedule on testing schema theory predictions in adults. *Journal of Motor Behavior, 17*, 283–299.

Levelt, W. (1989). *Speaking: From intention to articulation. Cambridge. MA: MIT Press*.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Li, C. H., & Thompson, S. A. (1977). The acquisition of tone in Mandarin speaking children. *Journal of Child Language*, *4*, 185-199.

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, *4*, 185-199.

Lin, M. C. (1965). The pitch indicator and the pitch characteristics of tones in Standard Chinese. *Acta Acoustica (China), 2*, 8-15.

Lindqvist, J. (1972). Laryngeal Articulation Studied on Swedish Subjects. *Speech Transmission Laboratory Quarterly Progress and Status Report, Royal Institute of Technology, Stockholm, Sweden, 2*, 10-27.

Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin Infant-directed speech. *Developmental Psychology, 43*, 912-917.

Lohse, K. R., Sherwood, D. E., & Healy, A. F. (2010). How changing the focus of attention affects performance, kinematics, and electromyography in dart throwing. *Human Movement Science, 29*, 542-555

Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S., Wulf, G., Ballard, K. J., & Schmidt, R. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*, 277-298.

McAlister, R. B. (2006). The effects of attentional focus instructions on simulated upper extremity amputees' movement kinematics when learning a novel function task (Doctoral dissertation.) Auburn University.

McNeil, M. R., Katz, W. F., Fossett, T. R. D., Garst, D. M., Szuminsky, N. J., Carter, G., et al. (2010). Effects of Online Augmented Kinematic and Perceptual Feedback on Treatment of Speech Movements in Apraxia of Speech. *Folia Phoniatrica et Logopaedica 62*, 127-133.

McNeil, M. R., Pratt, S. R., & Fossett, T. R. D. (2004). The Differential Diagnosis of Apraxia of Speech. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. M. Van Lieshout & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech*. Oxford, UK: Oxford Medical Publications.

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (1997). Apraxia of speech: Definition, differentation, and treatment. In M. R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders (pp. 311-344)*. New York: Thieme.

McNevin, H., & Wulf, G. (2002). Attentional focus on supra-postural tasks affects postural control. *Human movement science*, *21*, 187-202.

McNevin, N. H., Shea, C. H., & Wulf, G. (2003). Increasing the distance of an external focus of attention enhances learning. *Psychological Research, 67*, 22-29.

McNevin, N. H., Wulf, G., & Carlson, C. (2000). Effects of attentional focus, self-control and dyad training on motor learning: implications for physical rehabilitation. *Physical Therapy, 80*, 373-387.

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of Chinese Language Teachers Association, 24*, 49-65.

Molholt, G. (1988). Computer-assisted instruction in pronunciation for Chinese speakers of American English. *TESOL Quarterly, 22*, 91-111.

Mornell, A. (2007). Opening musicians' ears to attentional focus: Commentaries on "Attentional focus and motor learning". *E-Journal Bewegung und Training, 1*, 35-36. Retrieved October 30, 2009, from http://www.ejournal-but.de/doks/wulf_2007.pdf.

Newell, K. M., Carlton, M. J., & Antoniou, A. (1990). The interaction of criterion and feedback information in learning a drawing task. *Journal of Motor Behavior, 22*, 536-552.

Ng, M., Lerman, J., & Gilbert, H. (1998). Perceptions of tonal changes in normal laryngeal, esophageal, and artificial laryngeal male Cantonese speakers. *Folia Phoniatrica et Logopaedica, 50*, 64–70.

Ohala, J. J. (1978). Production of Tone. In V. A. Fromkin (Ed.), *Tone: a Linguistic Survey (p.5-39)*. New York: Academic Press.

Ohala, J. J., & Ewan, W. G. (1973). Speed of Pitch Change. *Journal of the Acoustical Society of America, 53*, 345.

Öster, A.-M. (1998). *Spoken L2 teaching with contrastive visual and auditory feedback*. Paper presented at the Fifth International Conference on Spoken Language.

Park, J. H., & Shea, C. H. (2003). Effect of practice on effector independence. *Journal of Motor Behavior, 35*, 33–40.

Park, J. H., & Shea, C. H. (2005). Sequence learning: Response structure and effector transfer. . *Quarterly Journal of Experimental Psychology, 58A*, 387–419.

Park, J. H., Shea, C. H., & Wright, D. L. (2000). Reduced-frequency concurrent and terminal feedback: A test of the guidance hypothesis. *Journal of Motor Behavior, 32*, 287-296.

Parker R., Vannest K., Davis J. & Sauber, S. (2011) Combining Nonoverlap and trend for single-case research. *Behavior Therapy, 42*, 284-299.

Perkell, J., Matthies, M. L., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J., et al. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*, 227–250.

Perkins-Ceccato, N., Passmore, S. R., & Lee, T. D. (2003). Effects of focus of attention depend of golfers' skill. *Journal of Sports Sciences, 21*, 593-600.

Porter, J. M., Nolan, R. P., Ostrowski, E. J., & Wulf, G. (2010). Directing attention externally enhances agility performance: a qualitative and quantitative analysis of the efficacy of using verbal instructions to focus attention. *Frontiers in Psychology, 1*, 1-7.

Ramig, L., Sapir, S., Fox, C., & Countryman, S. (2001). Changes in vocal loudness following intensive voice treatment (LSVT) in individuals with Parkinson's disease: A comparison with untreated patients and normal aged-matched controls. *Movement Disorders*, *16*, 79-83.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*, 460-499.

Rapp, B., & Goldrick, M. (2004). Feedback by any other name is still interactivity: A reply to Roelofs (2002). *Psychological Review, 111*, 573-578.

Robert, C. (1997). Concerns and issues in studying and assessing motor learning. *Measurement in physical education and exercise science, 1*, 19-38.

Robertson, E. M., Pacual-Leone, A., & Miall, R. C. (2004). Current concepts in procedural consolidation. *Nature Neuroscience Reviews, 5*, 1-7.

Rogers, M. A., & Spencer, K. A. (2001). Spoken word production without assembly: Is it possible? *Aphasiology, 15*, 68-73.

Ross, C. (2001). Evaluating intermediate Chinese textbooks. *Journal of the Chinese Language Teachers Association, 36*, 1-22.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical appraisal. *Psychological Bulletin*, *95*, 355-386.

Saltzman, E., & Kelso, J. A. S. (1987). Skilled actions: A task dynamic approach. *Psychological Review*, *94*, 84-106.

Saltzman, E., & Munhall, K. (1989). A dynamical approach to gestural patterning in speech Production. *Ecological Psychology, 1*, 333-382.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225-260.

Schmidt, R. A. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. In J. Requin & G. E. Stelmach (Eds.), *Tutorials in Motor Neuroscience (pp. 59-75)*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, R. A. (2003). Motor schema theory after 27 years: reflections and implications for a new theory. *Research Quarterly for Exercise and Sport, 74,* 366-375.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.

Schmidt, R. A., & Lee, T. D. (2005). *Motor Control and Learning: A Behavioral Emphasis (4th ed)*. Champaign, IL: Human Kinetics.

Schmidt, R. A., & Wulf, G. (1997). Continuous concurrent feedback degrades skill learning: Implications for training and simulation. *Human Factors, 39,* 509-525

Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: detection, search and attention. *Psychological review, 84*, 1-66.

Scruggs, T, Mastropieri, M, & Casto, G. (1987). The quantitative synthesis of single subject research methodology: Methodology and validation. *Remedial and Special Education*, 8, 24-33.

Sekiya, H., Magill, R. A., Sidaway, B., & Anderson, D. I. (1994). The contextual interference effect for skill variations from the same and different generalized motor programs. *Research Quarterly for Exercise and Sport, 65,* 330–338.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial order mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence Processing: Psycholinguistic studies presented to Merill Garrett* (pp. 295-342). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shattuck-Hufnagel, S. (1987). The role of word-onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller & M. Gopnik's (Eds.), *Motor and sensory processes of language* (pp. 17-51). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition, 42,* 213-259.

Shea, C., & Wulf, G. (1999). Enhancing motor learning through external-focus instructions and feedback. *Human movement science, 18,* 553-571.

Shea, C. H., & Kohl, R. M. (1991). Composition of practice: Influence on the retention of motor skills. *Research Quarterly for Exercise and Sport, 62,* 187–195.

Shea, C. H., Kohl, R. M., & Indermill, C. (1990). Contextual interference: Contributions of practice. *Acta Psychologica, 73,* 145–157.

Shea, C. H., Lai, Q., Black, C., & Park, J. H. (2000). Spacing practice sessions across days benefits the learning of motor skills. *Human Movement Science*, *19*, 737-760.

Shea, C. H., Lai, Q., Wright, D. L., Immink, M., & Black, C. (2001). Consistent and variable practice conditions: Effects on relative and absolute timing. *Journal of Motor Behavior, 33*, 139–152.

Shea, C. H., & Wulf, G. (2005). Schema Theory: A critical appraisal and reevaluation. *Journal of Motor Behavior, 37,* 85-101.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning & Memory, 5,* 179–187.

Shea, J. B., & Titzer, R. C. (1993). The Influence of Reminder Trials on Contextual Interference Effects. *Journal of Motor Behavior, 25,* 264 - 274.

Shea, J. B., & Zimny, S. T. (1983). Context effects in memory and learning movement information. In R. A. Magill (Ed.), *Memory and Control of Action* (pp. 345-366). Amsterdam: Elsevier

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association, 24,* 27-47.

Shriberg, L. D., & Kent, R. D. (2003). *Clinical Phonetics (3rd ed.)* Boston, MA: Allyn & Bacon.

Sims, B (2010). Focus of Attention Influences elite athletes sprinting performance. (Doctoral dissertation) Southern Illinois University Carbondale.

Spaai, G. W., & Hermes, D. J. (1993). A visual display for the teaching of intonation. *CALICO Journal, 10*, 19-30.

Spielman, J., Ramig, L. O., Mahler, L., Halpern, A., & Gavin, W. J. (2007). Effects of an extended version of the Lee Silverman Voice Treatment on voice and speech in Parkinson's disease. *American Journal of Speech-Language Pathology, 16*, 95–107.

Steinhauer, K., & Grayhack, J. P. (2000). The role of knowledge of results in performance and learning of a voice motor task. *Journal of Voice, 14*, 137–145.

Sterenberg, S., Knoll, R., Monsell, S., & Wright, C. (1988). Motor programs and hierarchical organization in the control of rapid speech *Phonetica, 45*, 175-197.

Sternberg, S., Monsell, S., Knoll, R. R., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.)*, Information processing in motor control and learning (pp. 117-152)*. New York: Academic Press.

Stetson, R. H. (1951). *Motor Phonetics (2nd ed)*. North-Holland Publishing Company: Amsterdam

Stibbard, R. (1996). Teaching English intonation with a visual display of fundamental frequency. *TESOL Quarterly, 2*.

Sundberg, J. (1979). Maximum Speed of Pitch Changes in Singers and Untrained Subjects. *Journal of Phonetics 7,* 71-79.

Swinnen, S., Schmidt, R. A., Nicholson, D. E., & Shapiro, D. C. (1990). Information feedback for skill acquisition: Instantaneous knowledge of results degrades learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 706–716.

Swinnen, S., Walter, C. B., Lee, T. D., & Serrien, D. J. (1993). Acquiring bimanual skills: Contrasting forms of information feedback for interlimb decoupling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1328-1344.

Swinnen, S. P. (1996). Information feedback for motor learning: A review. In H. N. Zelaznik (Ed.), *Advances in motor learning and control (pp. 37-66)*. Champaign, IL: Human Kinetics.

Swinnen, S. P., Lee, T. D., Verschueren, S., Serrien, D. J., & Bogaerds, H. (1997). Interlimb coordination: Learning and transfer under different feedback conditions. *Human Movement Science, 16*, 749-785.

Teitelbaum, P. (2012). Some useful insights for graduate students beginning their research in physiological psychology: Anecdotes and attitudes. *Behavioural Brain Research, 231*, 234-249.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press/Bradford.

Titze, I. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America, 85*, 1699-1707.

Van der Merwe, A. (2008). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders (2nd ed). (pp. 1-25)*. New York: Thieme Medical Publishers.

Van Rossum, J. H. A. (1990). Schmidt’s schema theory: The empirical base of the variability of practice hypothesis. A critical analysis. *Human Movement Science, 9*, 387–435.

Vance, J., Wulf, G., McNevin, N., Tollner, T., & Mercer, J. (2004). EMG activity as a function of the performer’s focus of attention. *Journal of Motor Behavior, 36*, 450-459.

Vander Linden, D. W., Cauraugh, J. H., & Greene, T. A. (1993). The effect of frequency of kinetic feedback on learning an isometric force production task in nondisabled subjects. *Physical Therapy, 73*, 79-87.

Vannest, K.J., Parker, R.I., & Gonen, O. (2011). Single Case Research: web based calculators for SCR analysis. (Version 1.0) [Web-based application]. College Station, TX: Texas A&M University. Retrieved from singlecaseresearch.org.

Vardanian, R. M. (1964). Teaching English through oscilloscope displays. *Language Learning, 3*, 109-118.

Varley, R., Whiteside, S., Windsor, F., & Fisher, H. (2006). Moving up from the segment: A comment on Aichert and Ziegler's Syllable frequency and syllable structure in apraxia of speech. *Brain and Language, 96*, 235-239.

Verdugo, D. R. (2006). A study of intonation awareness and learning in non-native speakers of English *language awareness, 15*, 141-159.

Vuillerme, N., & Nafati, G. (2007). How attentional focus on body sway affects postural control during quiet standing. *Psychological Research, 71*, 192-200.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). L2 acquisition and processing of Mandarin Tone. In L. P. Tan, E. Bates & O. J. L. Tzeng (Eds.), *Handbook of Chinese Psycholinguistics*. Cambridge: University Press.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Acoustical Society of America, 113*, 1033-1043.

Wang, Y., Ye tan, M., Xuexi, H., Yuyan, J., & Yu, Y. (1995). On American learners' tone acquisition. *Language Teaching and Research 2*, 126-140.

Weltens, B., & De Bot, K. (1984). Visual feedback of intonation II: Feedback delay and quality of feedback. *Language and speech 27*, 79-88.

White, C. M. (1981). Tonal perception errors and interference from English intonation. *Journal of Chinese Language Teachers Association 16*, 27-56.

Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition,, 16*, 677-691.

Wright, D. L., Black, C. B., Immink, M. A., Brueckner, S., & Magnuson, C. (2004). Long-term motor programming improvements occur via concatenating movement sequences during random but not blocked practice. *Journal of Motor Behavior, 36*, 39–50.

Wulf, G. (2007a). *Attention and motor skill learning*: Chamoaign, IL: Human Kinetics Inc.

Wulf, G. (2007b). Attentional focus and motor learning: a review of ten years of research. *E-Journal Bewegung und Training,, 1*, 4-14 Retrieved October 30, 2009, from http://www.ejournal-but.de/doks/wulf_2007.pdf.

Wulf, G., Chiviacowsky, S., Schiller, E., & Avilla, L. T. G. (2010). Frequent external-focus enhances motor learning. *Frontiers in Psychology, 1*, 1-7.

Wulf, G., HoB, M., & Prinz, W. (1998). Instructions for motor learning: Differential effects of internal versus external focus of attention. *Journal of Motor Behavior, 30*, 169-179.

Wulf, G., Landers, M., Lewthwaite, R., & Töllner, T. (2009). External focus instructions reduce postural instability in individuals with Parkinson disease. *Physical Therapy, 89*, 162-168.

Wulf, G., & Lee, T. D. (1993). Contextual interference in movements of the same class: Differential effects on program and parameter learning. *Journal of Motor Behavior, 25*, 254–263.

Wulf, G., Lee, T. D., & Schmidt, R. A. (1994). Reducing knowledge of results about relative versus absolute timing: Differential effects on learning. *Journal of Motor Behavior, 26*, 362-369.

Wulf, G., McConnel, N., Gartner, M., & Schwarz, A. (2002). Enhancing the learning of sports skills through external-focus feedback. *Journal of Motor Behavior, 34*, 171-182.

Wulf, G., McNevin, N., & Shea, C. H. (2001). The automaticity of complex motor skill learning as a function of attentional focus. *The Quarterly Journal of Experimental Psychology, 54A*, 1143-1154.

Wulf, G., Mercer, J., McNevin, N., & Guadagnoli, M. A. (2004). Reciprocal influences on attentional focus on postural and suprapostural task performance. *Journal of Motor Behavior, 36*, 189-199.

Wulf, G., & Prinz, W. (2001). Directing attention to movement effects enhances learning: a review. *Psychonomic Bulletin & Review, 8*, 648-660.

Wulf, G., & Schmidt, R. A. (1989). The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 748-757.

Wulf, G., Schmidt, R. A., & Deubel, H. (1993). Reduced feedback frequency enhances generalized motor program learning but not parameterization learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1134-1150.

Wulf, G., & Shea, C. H. (2002). Principles derived form the study of simple motor skills do not generalize to complex skill learning. *Psychonomic Bulletin and Review, 9*, 185-211.

Wulf, G., & Shea, C. H. (2004). Understanding the role of augmented feedback: The good, the bad, and the ugly. In A. M. Williams, & Hodges, N. J. (Ed.), *Skill acquisition in sport: Research, theory and practice* (pp. 121–144). London: Routledge.

Wulf, G., & Su., J. (2007). An external focus of attention enhances golf shot accuracy in beginners and experts. *Research Quarterly for Exercise and Sport, 78*, 384-389.

Wulf, G., Toller, T., & Shea, C. (2007). Attentional focus effects as a function of task difficulty. *Research Quarterly for Exercise and Sport, 78*, 257-264.

Wulf, G., & Weigelt, C. (1997). Instructions about physical principles in learning a complex motor skill: to tell or not to tell. *Research Quarterly for Exercise and Sport, 68*(4), 362-367.

Wulf, G., Weigelt, M., Poulter, D., & McNevin, H. (2003). Attentional focus on suprapostural tasks affects balance learning. *Quarterly Journal of Experimental Psychology, 56*, 1191-1211.

Wulf, G., & Weight, C. (1997). Instructions about physical principles in learning a complex motor skill: To tell or not to tell. *Quarterly Journal of Experimental Psychology, 56*, 1191-1211.

Xu, Y. (2004). Understanding tone from the perspective of production and perception. *language and linguistics, 5*, 757-797.

Yaruss, S. (2010). Assessing quality of life in stuttering treatment outcomes research. *Journal of fluency Disorders, 35*, 190-202.

Yip, M. (2002). *Tone*. Cambridge: The University Press.

Yiu, E. M., van Hasselt, C. A., Williams, S. R., & Woo, J. K. S. (1994). Speech intelligibility in tone language Chinese laryngectomy speakers. *European Journal of Disorders of Communication, 29*, 339–347.

Zachry, T., Wulf, G., Mercer, J., & Bezodis, N. (2005). Increased movement accuracy and reduced EMG activity as the result of adopting an external focus of attention. *Brain Bulletin, 67*, 304-309.

Zentgraf, K., Lorey, B., Bischoff, M., Zimmermann, K., Stark, R., & Munzert, J. (2009). Neural correlates of attentional focusing during finger movements: a fMRI study. *Journal of Motor Behavior, 41*, 535-541.

Ziegler, W. (2002). Psycholinguistic and motor theories of apraxia of speech. *Seminars in Speech and Langauge, 23*, 231-243.

Ziegler, W. (2003a). Speech motor control is task-specific. Evidence from dysarthria and apraxia of speech. *Aphasiology, 17*, 3-36.

Ziegler, W. (2003b). To speak or not to speak: distinctions between speech and nonspeech motor control. A reply to Ballard, Robin & Folkins. *Aphasiology, 17*, 99-105.