RULE- VERSUS INSTANCE-BASED LEARNING IN SPEECH-LIKE BEHAVIOR: AN EVALUATION OF TRANSFER AND MOTOR CLASS EFFECTS

by

Kimberly M. Meigh, M.A. CCC-SLP

B.A., Michigan State University, 2002

M.A., University of Pittsburgh, 2004

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

2014

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REAHBILITATION SCIENCES

This dissertation was presented

by

Kimberly M. Meigh, M.A. CCC-SLP

It was defended on

July 14, 2014

and approved by

Susan Shaiman, PhD, Associate Professor, Communication Science and Disorders

Connie Tompkins, PhD, Professor, Communication Science and Disorders

Katherine Verdolini Abbott, PhD, Professor, Communication Science and Disorders

Timothy Nokes-Malach, PhD, Associate Professor, Psychology

Dissertation Advisor: Susan Shaiman, PhD, Associate Professor,

Communication Science and Disorders

Copyright © by Kimberly M. Meigh

2014

RULE- VERSUS INSTANCE-BASED LEARNING IN SPEECH-LIKE BEHAVIOR: AN EVALUATION OF TRANSFER AND MOTOR CLASS EFFECTS

Kimberly M. Meigh, PhD, CCC-SLP

University of Pittsburgh, 2014

Two information-processing theories of motor control have been postulated for motor learning. Rule-based learning theory predicts transfer when new, untrained stimuli or behaviors share the same set of rules. Instance-based learning theory predicts transfer when new, untrained stimuli are similar in a specific way to the trained stimuli. The purpose of this study was to provide insight into the learning theory operating during nonword acquisition and transfer by evaluating reaction times during an old-new judgment task. Nonword stimuli were constructed to bias familiarity judgments by systematically varying two parameters associated with each theory: phonetic similarity (instance-based) and syllable stress pattern (rule-based).

Twenty-four participants (18-35 years of age) with normal hearing and speech production participated in a syllable stress training task and an old-new judgment task. During training, participants articulated a series of nonword stimuli while producing a specific syllable stress pattern. Syllable stress accuracy was monitored by the examiner via perceptual judgments and custom software evaluating acoustic intensity of the articulated stressed syllable. Accurate articulation of nonwords was monitored with recognition probes throughout training. Participants

met pre-established accuracy criteria for syllable stress and phonetic production of each experimental nonword. Once criterion was met, participants were assumed to have a highly-accurate baseline memory representation of the trained items that was judged against a variety of untrained transfer stimuli varying in phonetic similarity and syllable stress pattern. Following training, an old-new judgment task was administered in which participants made familiarity judgments upon hearing a trained or untrained nonword; reaction times were collected via a response box.

Reaction time results indicated participants responded faster to untrained nonwords with different phonemes than to untrained nonwords with similar phonemes. Syllable stress pattern did not affect reaction time. These results are consistent with instance-based learning. However, the direction of the similarity effect was in the opposite direction as originally predicted for this theory, i.e., positive transfer occurred when stimuli were dissimilar to one another. Future studies should evaluate what parameters need to be manipulated along a similarity index, and how the variable of dissimilarity may affect overall transfer patterns.

TABLE OF CONTENTS

1.0 INTRODUCTION 1

2.0 RULE-BASED LEARNING 3

2.1 INTRODUCTION 3

2.2 HISTORICAL OVERVIEW 4

2.2.1 Categorization 4

2.2.2 Artificial Grammar 6

2.3 REPRESENTATIONAL FORMS 7

2.3.1 Cognitive Forms 7

2.3.2 Motor Forms 9

2.3.3 Speech production 12

2.4 TRANSFER 15

2.4.1 Cognitive Tasks 15

2.4.2 Motor Tasks 17

2.5 FUNDAMENTAL FLAWS 21

2.5.1 Abstraction 21

2.5.2 Defining Rules 22

2.6 SUMMARY 23

3.0 INSTANCE-BASED LEARNING 25

3.1 INTRODUCTION 25
3.2 HISTORICAL OVERVIEW 26
3.2.1 Frequency 26
3.2.2 Categorization 27
3.3 REPRESENTATIONAL FORM 29
3.3.1 Cognitive 29
3.3.2 General Motor 32
3.3.3 Speech Production 35
3.4 TRANSFER 38
3.5 FUNDAMENTAL FLAWS 42
3.5.1 Storage 42
3.5.2 Forgetting 43
3.6 SUMMARY 44
4.0 COMPARISONS OF LEARNING THEORIESS IN MOTOR BEHAVIOR 46
4.1 INTRODUCTION 46
4.2 EVALUATING THE REPRESENTATION 47
4.2.1 Models of motor performance 47
4.2.2 Assessing the representation 52
4.3 DIRECT COMPARISONS IN MOTOR BEHAVIOR 56
4.3.1 Transfer Predictions 56
4.3.2 Experiments 58
4.3.2.1 Chamberlin and Magill, 1992b 58
4.3.2.2 Crump and Logan, 2010a 59

4.3.3 Critical Review 60
4.3.3.1 Assessing Memory Representation 61
4.3.3.2 Assessing Different Models of Learning 65
4.4 SUMMARY 66
5.0 OVERALL SUMMARY AND DISCUSSION 68
6.0 SPECIFIC AIMS AND RESEARCH QUESTIONS 73
7.0 METHOD 84
7.1 STIMULI VALIDATION STUDY 84
7.1.1 Participant Criterion 87
7.1.2 Participant Recruitment 90
7.1.3 Participant Sample Size 90
7.1.4 Participant Characteristics 91
7.1.5 Experimental Procedure 91
7.1.5.1 Training Block 92
7.1.5.2 Experimental Blocks 93
7.1.6 Stimuli 94
7.1.6.1 Kendall et al. (2005) Nonword Stimuli 94
7.1.6.2 Kendall et al. (2005) Real Word Stimuli 95
7.1.6.3 Different-Phonemes-Different-Motor-Class (DPDC) Stimuli 96
7.1.6.4 Stimuli Preparation 97
7.2 RULE VERSUS INSTANCE-BASED LEARNING 99
7.2.1 Participant Criterion 101
7.2.2 Participant Sample Size and Power Analysis 102

7.2.3 Participant Recruitment 103

7.2.4 Participant Characteristics 103

7.2.5 Experimental Procedures and Equipment Configuration 104

7.2.5.1 Procedures 104

7.2.5.2 Equipment Configuration 106

7.2.6 Task Descriptions 107

7.2.6.1 Syllable Stress Training: General Overview 107

7.2.6.2 Training Task: Perception-Production Training 110

7.2.6.3 Training Task: Syllable Stress Training 111

7.2.6.4 Old-New Judgment Task: Rationale 118

7.2.6.5 Old-New Judgment Task: General 119

7.2.6.6 Old-New Judgment Task: Instructions 121

7.2.7 Stimuli 123

7.2.7.1 Trained Stimuli: Experimental 123

7.2.7.2 Trained Stimuli: Filler 125

7.2.7.3 Untrained Stimuli: General Overview 126

7.2.7.4 Untrained Stimuli: SPSC Stimuli 127

7.2.7.5 Untrained Stimuli: DPSC Stimuli 128

7.2.7.6 Untrained Stimuli: DPDC Stimuli 129

7.2.8 Stimuli Preparation 131

8.0 RESULTS 132

8.1 STIMULI VALIDATION 132

8.1.1 Data Preparation 133

8.1.1.1 Accuracy Analyses 133

8.1.1.2 Reaction Time Data 134

8.1.2 PRIMARY OUTCOMES..... 135

8.1.2.1 Research Question 1 (RQ1): Nonword Frequency 135

8.1.2.2 Research Question 2 (RQ2): Real Word Frequency..... 137

8.1.2.3 Research Question 3 (RQ3): Stimuli Type 137

8.2 RULE- VERSUS INSTANCE-BASED LEARNING 138

8.2.1 Data Preparation 139

8.2.1.1 Accuracy Data: Overview of *d'* Statistic Calculations..... 140

8.2.1.2 Accuracy: *d'* Statistic and Error Analyses 141

8.2.2 PRIMARY OUTCOMES..... 144

8.2.2.1 Research Question 4 (RQ4): Phonetic Similarity..... 145

8.2.2.2 Research Question 5 (RQ5): Motor Class..... 146

8.2.3 SECONDARY OUTCOMES 147

8.2.3.1 Effect of Response Button Position..... 147

8.2.3.2 Effect of Training Amounts 149

8.2.3.3 Effect of Stimuli Type..... 152

9.0 DISCUSSION 160

9.1 STIMULI VALIDATION STUDY 160

9.1.1 RQ1: Nonword Frequency..... 160

9.1.2 RQ2: Real Word Frequency..... 160

9.1.3 RQ3: Stimuli Type..... 161

9.1.4 Summary 162

9.2 RULE- VERSUS INSTANCE-BASED LEARNING STUDY 162

9.2.1 Primary Outcomes 162

9.2.2 RQ4: Phonetic Similarity 164

9.2.3 RQ5: Motor Class 166

9.2.4 Secondary Outcomes 168

9.2.4.1 Response Button Position 168

9.2.4.2 Training Amount 168

9.2.4.3 Stimuli Item Analysis 169

9.2.5 Summary 179

9.3 LIMITATIONS 181

9.3.1 Construction of the Stimuli: Syllable versus adjacent phonemes 181

9.3.2 Instructional Bias 182

9.3.3 Measuring Implicit Memory with an Explicit Task 184

9.3.4 Measuring Syllable Stress During Training 186

9.3.5 Training Program *Stimulate* 188

9.4 FUTURE DIRECTIONS 189

APPENDIX A: MODELS OF MOTOR PERFORMANCE 193

APPENDIX B: LANGUAGE SCREENING SCRIPT 196

APPENDIX C: SCREENING FORM STIMULI VALIDATION STUDY 198

APPENDIX D: SCREENING FORM MAIN STUDY 202

APPENDIX E: DATA FOR PREDICTED ACCURACY RESULTS 205

APPENDIX F: POST HOC ANALYSIS OF PRIMARY OUTCOMES 207

BIBLIOGRAPHY 210

LIST OF TABLES

Table 1: Shared characteristics between GMPs and Stressed Syllables 15

Table 2: Rule- and instance-based learning contrasts 69

Table 3: Predictions for high vs. moderate vs. low frequency nonword stimuli 75

Table 4: Predictions for high vs. low frequency real words 76

Table 5: Predictions for Kendall et al. (2005) stimuli vs. DPDC stimuli 77

Table 6: Predictions for transfer pattern based on similarity 80

Table 7: Predictions for motor class when similarity is held constant 83

Table 8: Parameters controlled for in Kendall et al.'s (2005) nonword stimuli 86

Table 9: Training Block Stimuli 93

Table 10: Kendall et al. (2005) nonword stimuli 95

Table 11: Kendall et al. (2005) real word stimuli 96

Table 12: DPDC Stimuli 97

Table 13: Means and SDs for re-recorded Kendall et al. (2005) stimuli 99

Table 14: Overview of the experimental tasks, stimuli, and learning theories 101

Table 15: Training blocks required for accuracy criterion to be met 112

Table 16: Syllable stress makers produced by participants 115

Table 17: Trained stimuli 124

Table 18: Filler stimuli 125

Table 19: SPSC Stimuli 127

Table 20: DPSC Stimuli 128

Table 21: Syllable stress proportions of Trained and DPSC stimuli 128

Table 22: Intergestural phonemic constraints for DPDC stimuli 130

Table 23: DPDC Stimuli 130

Table 24: Error analysis for real words (across participants) 133

Table 25: Error analysis for nonwords (across participants) 133

Table 26: Total number and percentage of No Responses and Breath Triggers for real word stimuli 134

Table 27: Total number and percentage of No Responses and Breath Triggers for nonword stimuli 135

Table 28: Test of normality of nonword gestural frequency for each frequency category 135

Table 29: Mean and SD for nonword gestural frequency categories 136

Table 30: Transfer variables associated with each learning theory 139

Table 31: Total number of No Responses and Outliers greater than >3SD from the median 139

Table 32: Anticipated and summarized data to calculate *d'* statistic 140

Table 33: *H*, *F*, *d'* and *Br* statistics for all participants 142

Table 34: Error analysis by stimuli type (across participants) 143

Table 35: *Br* statistics in relation to *F* rates and accuracy scores on SPSC stimuli 143

Table 36: Test of normality across stimuli type 144

Table 37: Mean and SD for stimuli with varying phonetic similarity 146

Table 38: Test of equality of error variances by stimuli type 147

Table 39: Test of normality of stimuli type by button group 148

Table 40: Mean and SDs for stimuli type by button push group 149

Table 41: Test of equality of error variances by stimuli type 150

Table 42: Test of normality of stimuli type by training amount 151

Table 43: Mean and SDs for stimuli type by training amount 152

Table 44: Significant pairwise comparison results for SPSC stimuli 153

Table 45: Median reaction times of SPSC stimuli items 155

Table 46: Significant pairwise comparison results for DPSC stimuli comparisons 156

Table 47: Median reaction times of DPSC stimuli items 158

Table 48: Median reaction times of DPDC stimuli 159

Table 49: Mean and SDs of reaction time across stimuli type 167

Table 50: Fastest versus slowest SPSC stimuli 170

Table 51: Fastest versus slowest DPSC stimuli 176

Table 52: Median reaction times of DPDC stimuli 179

Table 53: High hit rate 206

Table 54: Low hit rate 206

Table 55: Median reaction times of Trained, SPSC, and DPSC stimuli 208

LIST OF FIGURES

Figure 1: Prototype and similarity 5

Figure 2: Example of uniform pattern of generalization in a GMP study 19

Figure 3: Effect of motor class on different theories of learning..... 20

Figure 4: Instance representation 31

Figure 5:Visual depiction of an instance-based explanation of Keetch et al.'s (2005) data..... 34

Figure 6: Instance- versus- rule based learning in hypothetical stimuli 39

Figure 7: Potential response alternatives 49

Figure 8: Overview of response time..... 53

Figure 9: Rule-based learning predictions for transfer based on phonetic similarity 78

Figure 10: Instance-based learning predictions for transfer based on phonetic similarity 79

Figure 11: Rule-based learning motor class predictions..... 81

Figure 12: Instance-based learning motor class predictions 82

Figure 13: Experimental Protocol..... 106

Figure 14: Visual display of a model depicting first syllable stress 108

Figure 15: Stimulate feedback displays 110

Figure 16: Recognition Probe Schedule 118

Figure 17: Response box display 121

Figure 18: Legend for untrained stimuli construction 126

Figure 19: Mean reaction times (ms) and standard errors of interphonemic transitional gestural frequency categories. 136

Figure 20: Mean reaction times (ms) and standard errors of real word stimuli across frequency categories 137

Figure 21: Mean reaction times (ms) and standard error of nonword stimuli type 138

Figure 22: Mean reaction times (ms) and standard errors of stimuli varying by phonetic similarity 145

Figure 23: Mean reaction times (ms) and standard errors of stimuli varying by motor class 146

Figure 24: Mean reaction times (ms) and standard errors of stimuli type across button position groups (2 = OLD, 4 = OLD) 149

Figure 25: Mean reaction times (ms) and standard errors of stimuli type across training block groups (1 training block, 2+ training blocks) 152

Figure 26: Mean reaction times (ms) and standard errors of individual SPSC stimuli. 154

Figure 27: Mean reaction time (ms) and standard error of individual DPSC stimuli. 157

Figure 28: Mean reaction times (ms) and standard errors of individual DPDC stimuli 159

Figure 29: Transfer predictions for each learning theory 163

Figure 30: Information-processing model of general motor behavior 194

Figure 31: Information-processing model of speech production 195

Figure 32: Mean reaction times (ms) and standard errors of Trained, SPSC, and DPSC stimuli (inaccurate responses removed) 208

Figure 33: Mean reaction times (ms) and standard errors of stimuli varying by motor class 209

PREFACE

This study was supported in part through funding received from the SHRS Research Development Fund, School of Health and Rehabilitation Sciences, University of Pittsburgh, as well as the Audrey Holland Endowed Research Award, Department of Communication Sciences and Disorders, University of Pittsburgh.

1.0 INTRODUCTION

Investigators in skill acquisition evaluate learning patterns to extrapolate the underlying memory representations directing skilled behavior (Chamberlin & Magill, 1992a; Crump & Logan, 2010; Kantak & Winstein, 2012; Neal & Hesketh, 1997; Shanks, 1995). Transfer, or generalization of the trained skill, is important as it describes how acquired knowledge can be applied in novel contexts (Nokes, 2004, 2009; Schmidt & Wrisberg, 2004), which may suggest how stored knowledge is represented in memory (Chamberlin & Magill, 1992a; Shanks, 1995). Currently, there are two main forms of learning proposed for cognitive and motor behaviors: rule- and instance-based learning (Mathews et al., 1989; Neal & Hesketh, 1997; Shanks, 1995).

Rule-based learning abstracts relevant information during skill training (Dopkins & Gleason, 1997; Neal & Hesketh, 1997; Posner & Keele, 1970; Shanks, 1995). The resultant representation is a context-limited set of rules, which are refined with practice (Doody & Zelaznik, 1988; Neal & Hesketh, 1997; Shanks, 1995). Transfer occurs when trained behaviors share the same set of rules as untrained behaviors (Medin & Ross, 1989; Neal & Hesketh, 1997; Shanks, 1995). Instance-based learning, conversely, encodes all information (relevant and irrelevant) into an exemplar representation during skill acquisition (Hintzman, 1976; Jacoby & Brooks, 1984; Shanks, 1995). Instead of relying on rules to direct transfer effects, generalization is based on similarity of the stored, trained exemplars to untrained behaviors (Brooks, 1978; Hintzman, 1986; Palmeri, 1997; Shanks, 1995).

While both forms of learning are addressed in motor control theory, the application of rule-based learning is historically more prevalent. However, transfer results are not always consistent with rule-based transfer (Beek, 2000; Logan, 1985; Magill, 2001; Rosenbaum, Carlson, & Gilmore, 2001). Inconsistencies in rule-based learning are also evident in the speech production literature (e.g., Austermann-Hula, Robin, Maas, Ballard, & Schmidt, 2008; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000), and instance-based learning has never been empirically evaluated as a possible alternative explanation. Thus, it is unclear if instance-based learning would provide a more accurate description of speech production acquisition and transfer effects.

The main goal of the proposed investigation is to evaluate the pattern of learning occurring in a novel speech task. Specifically, the current examination will evaluate transfer performance to determine if the pattern of results is more consistent with rule- or instance-based learning. In the following text, a detailed background will be developed of both learning paradigms. A summary and discussion of the literature review will be provided, which leads to the specific aims, research questions, and proposed methods for this dissertation. Finally, the results and discussion will be presented.

2.0 RULE-BASED LEARNING

2.1 INTRODUCTION

Rule-based learning is based on the concept of abstraction, a process in which central information about a stimulus is summarized into an averaged representation (Dopkins & Gleason, 1997; Neal & Hesketh, 1997; Posner & Keele, 1970; Shanks, 1995). This provides a "compact, efficient representation" to direct behaviors that "transcends the training stimuli" from which the rule-based representation was generated (Shanks, 1995, p. 167). The summarized representation then promotes generalization to untrained stimuli sharing similar underlying characteristics (Medin & Ross, 1989; Neal & Hesketh, 1997; Shanks, 1995).

Rule-based learning has dominated theories of general motor control since the late 19th century (e.g., James, 1890; see J. A. Adams, 1987 for a historical review), and continues to be the most prevalent learning theory in describing speech acquisition and transfer effects. An operational definition of "rule" is needed to better understand the variety of rule-based systems in cognitive and motor theory, and the type of knowledge representation transferred during learning. There is considerable debate as to what constitutes a rule in cognitive psychology; however, as mental operations are not visible, rule-based learning and representation must be inferred from behaviors (Shanks, 1995). For this dissertation, the following definition will be adopted: "…behaviour is based on a rule if no difference is observable between performance to trained (old) and untrained (new) stimuli that fall into the same category" (Shanks, 1995, p. 153;

also: Herrnstein, 1990; E. E. Smith, Langston, & Nisbett, 1992). This definition is used both directly (e.g., Chamberlin & Magill, 1992a) and indirectly (e.g., Schmidt, 1975) in theories of motor learning and transfer. This chapter will provide a historical review of rule-based behavior, representational form, transfer and generalization effects, and flaws in this theory.

2.2 HISTORICAL OVERVIEW

2.2.1 Categorization

Rule-based learning is inferred from studies of visual categorization. Studies are designed so an implicit rule, i.e., one not overtly obvious to the participant, is learned through repeated exposures to stimuli (Ward & Churchill, 1998). When presented with novel test items following training, participants would pick the item which instantiated the rule over items in which the rule did not apply (Bright & Burton, 1994; Ward & Churchill, 1998). This effect is also noted in recognition paradigms in which stimuli are categorized as “old” or “new” (Nosofsky, Clark, & Shin, 1989; Shanks, 1995), and novel stimuli sharing the same underlying rule are considered “old” despite never being encountered in training (Nosofsky et al., 1989; Shanks, 1995).

Prototype memory representations are postulated to account for the variability and similarity of stimuli that are easily categorized (Shanks, 1995). These two issues are intimately linked as repeated encounters with a stimulus, e.g., a robin, will lead to multiple, similar memory representations being encoded into memory. However, each encounter with the stimulus may be slightly different from a previous encounter (e.g., different robins have unique features), which causes variability in the overall sum of the exposures. Prototype theory suggests abstraction of

the relevant features of a stimulus creates a mean representation (a prototype) that other exposures can be compared to and then categorized (Ashby & Maddox, 2005; Dopkins & Gleason, 1997; Homa, Sterling, & Trepel, 1981; Olsson, Wennerholm, & Lyxzèn, 2004; Rosch, Simpson, & Miller, 1976; Shanks & St. John, 1994; Shanks, 1995). The closer in similarity a stimulus is to a prototype the more quickly the novel stimulus is categorized based on response time, see Figure 1 (Chamberlin & Magill, 1992a; Jacoby & Brooks, 1984; Medin & Ross, 1989; Rosch et al., 1976; Shanks, 1995).

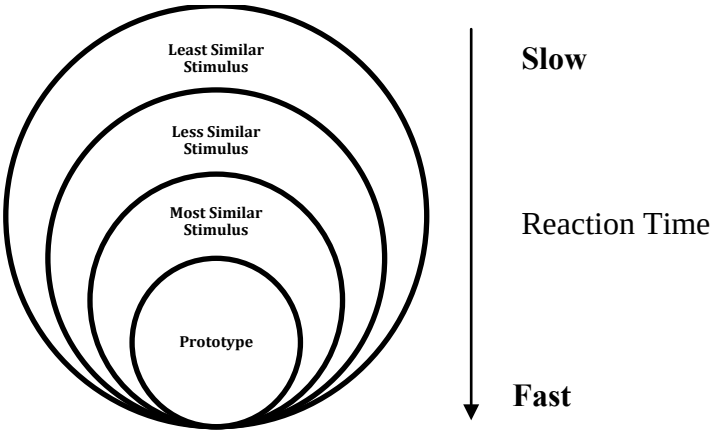

Figure 1: Prototype and similarity

Note: Adapted from *The Psychology of Learning and Motivation: Advances in Research Theory*, Vol. 18, Jacoby & Brooks, Nonanalytic cognition: Memory, perception, and concept learning. p.25, 1984, with permission from Elsevier.

This effect is justified by different rules directing categorization of the prototype. Narrow rules are limited to shared features between the stimuli, whereas broad rules apply to less similar features (Altmann, Dienes, & Goode, 1995; Shanks & St. John, 1994; Shanks, 1995; Shepard,

Hovland, & Jenkins, 1961). This variety of rules, and their abstraction process, will be evaluated in Section 2.4 in regard to transfer effects.

Empirical results demonstrating accurate and rapid prototype classification is consistent with rule-based learning, even when the prototype is not encountered in training (Homa et al., 1981; Posner & Keele, 1968, 1970; Shanks, 1995). Homa et al. (1981) reported prototypes were classified more accurately than trained items following a one week follow up. Evidence consistent with prototype theory has also been reported in research using complex stimuli, as well as in categories in which within-category variability is large (Minda & Smith, 2001; Olsson et al., 2004; Shanks & St. John, 1994; Shepard et al., 1961; J. D. Smith & Minda, 1998).

2.2.2 Artificial Grammar

Artificial grammar learning is traditionally explained by rule-based learning. Participants are trained on strings of letters representing specific syntactic rules for how the novel strings are combined (McAndrews & Moscovitch, 1985). Like natural languages, artificial languages can have rule violations in syntactic sequencing, e.g., nongrammatical "sentences," that can be learned with training (McAndrews & Moscovitch, 1985). The rules of the artificial grammar may not be explicitly obvious to the participants (Reber, 1967, 1989); however, after training, participants can classify grammatical strings more accurately and rapidly than ungrammatical strings (McAndrews & Moscovitch, 1985; Reber, 1967, 1989). Narrow and broad transfer effects are also noted in artificial grammar studies though these are often referred to as learning the "surface" (narrow transfer) versus "deep" (broad transfer) structures (Altmann et al., 1995; Brooks & Vokey, 1991; Shanks, 1995; Witt & Vinter, 2012). Altmann et al.'s (1995) transfer data on artificial grammar strings and associated tone sequences is consistent with deep structure

transfer, in which training on either rule-based grammar strings or tone sequences yielded transfer to both types of novel stimuli sharing the same underlying rules.

2.3 REPRESENTATIONAL FORMS

Rule-based learning representations take many forms. The specific form is determined by the amount of abstraction occurring during encounters with stimuli (Ashby & Maddox, 2005; Shanks, 1995). Abstraction of prototype representations consists of the summary features of a group of similar stimuli (e.g., wings, feathers, and a beak may be the features associated with the category "bird"). These features may be abstracted in a single, brief encounter with a stimulus. Abstraction of other information, such as a list of rules or production statements (J. R. Anderson, 1993), may only occur over several exposures to a stimulus as noted in learning or training paradigms. A variety of representational rule-based forms will be described in the following section; however, the main emphasis will be placed on schema theory in cognitive and motor research.

2.3.1 Cognitive Forms

Regardless of the form, all rule-based learning representations characterize the mean, or summary representation, of information gathered during an encounter (or several encounters) with a stimulus (Dopkins & Gleason, 1997; Neal & Hesketh, 1997; Posner & Keele, 1970; Shanks, 1995). These summary representations maintain only the relevant features of the encounter – all other information is considered irrelevant to the representation and discarded

(Neal & Hesketh, 1997; Ohlsson, 1993; Shanks, 1995). Feature relevancy is determined by the rule-based learning system, but often includes salient or prominent features associated with a given type of stimuli (Shanks, 1995). For example, the salient features of wings, beak, and song production may be encoded to form a prototype representation for "bird." Less relevant features, such as size, type of feet, etc. would be discarded and not used to build the prototype representation for "bird." As each encounter encodes only partial information, earlier encounters with a stimulus are more significant in determining the rule-based representation than later exposures to a stimulus (Crump & Logan, 2010; Posner & Keele, 1970; Ross, Taylor, Middleton, & Nokes, 2008).

Rule-based learning is an umbrella term in cognitive psychology that covers a heterogeneous set of theories and representations, with applications including (but not limited to) analogies (e.g., D. Gentner, 1983; Holyoak & Thagard, 1989; Nokes & Ohlsson, 2001), rule training strategies (e.g., Fong, Krantz, & Nisbett, 1986; Larrick, Morgan, & Nisbett, 1990; Singley & Anderson, 1989), declarative to procedural transfer rules (e.g., J. R. Anderson, 1987, 1993; Ohlsson, 1996), algorithms (e.g., Domingos, 1996), reasoning rules (e.g., E. E. Smith et al., 1992), phonological rules (e.g., Chomsky & Halle, 1968; Chomsky, 1963; Levelt, Roelofs, & Meyers, 1999), judgment heuristics (e.g., Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974), and inferencing strategies (e.g., Persson & Rieskamp, 2009). Neural network models also describe and replicate rule-based behaviors (e.g., Li & Nara, 2012; Sweegers, Takashima, Fernández, & Talamini, 2013; Yap, Lim, & Au, 2011). Patterns of neural activation may be likened to other rule-based representations (e.g., prototypes); however, the feedback received by the neural network may only allow for narrow transfer effects (Shanks, 1995).

Schemas are a popular rule-based representation within cognitive psychology, and have been adopted from this literature into motor theory. Whereas prototypes can be considered a collection of features directing categorization, schemas are a collection of features linked together in a useful way to direct complex behavior (Marshall, 1995; Rumelhart & Ortony, 1977). Schema representations are often denoted in higher-order, top-down processing cognitive activities, such as text comprehension (e.g., R. C. Anderson & Pearson, 1984; Bartlett, 1932; Kintsch, 1998), problem solving (e.g., Marshall, 1995), mathematics (Skemp, 1987), and general theoretical knowledge not represented by particular environmental exemplars (e.g., the span of the universe; Ohlsson, 1993). Despite the wide usage of schema theory, the origins and development of schema representations are often poorly described (Marshall, 1995; Nokes & Ohlsson, 2001). Further specification of schema representations in motor behavior will be provided below.

2.3.2 Motor Forms

Rule-based learning is prevalent in limb and speech motor control theories. Complex motor acts rely on multiple levels of processing (Cooper & Shallice, 2000; Norman & Shallice, 1986). Higher levels of control, such as scripts (Schank & Abelson, 1977), direct well-learned, goal-oriented motor activities (e.g., driving to work). These scripts are composed of sequenced schemas that break down the larger script into smaller, component parts (e.g., separate schemas for applying pressure to the gas pedal versus braking). The schemas direct the lowest level of motor control, the biomechanical gestures, required for the movement (Cooper & Shallice, 2000; Rosenbaum, Inhoff, & Gordon, 1984). For the purposes of this dissertation, focus will be placed

on the intermediary level of motor control, i.e., schema theory, which is the most published and empirically studied rule-based theory of motor control.

Schema theories provide explanations for two problems in motor control: 1.) How do smooth, rapid transitions occur in ballistic movements? 2.) What mechanism explains motor equivalence (Bernstein, 1967) in which a given skilled movement can be achieved through a variety of different muscle configurations? Schema representations provide a programming mechanism (or set of rules) which controls the deployment and timing of movements (Schmidt & Wrisberg, 2004). Programming also provides goal-directed guidelines for what motor transitions need to occur, given particular environmental constraints (Hall, 1989; Lee & Swinnen, 1993; Magill, 2001). Thus, the schema directs specification of the movement based on the overall task goal.

Schema theory takes the form of a motor program in which practice refines the program by some secondary mechanism until the desired performance outcome is realized over time (Coker, 2004; Keele, 1968; Lee & Swinnen, 1993; Schmidt, 1975). Refinement can be equated to the abstraction process in which a centralized set of features or rules is derived over the course of practice (Hall, 1989; Magill, 2001). What is left following refinement, i.e., what the motor program consists of, is dependent upon the motor program theory. Rosenbaum (1980) proposed "the information in a motor program can be assumed to consist of prescriptions for the *values* that a forthcoming movement should have on *dimensions* that are under the program's control" (p. 446).

Schmidt (1975) specified "values" and "dimensions" in his generalized motor program (GMP) theory by distinguishing invariant from variant features of movement. Invariant features, symbolized in a motor program, are generalized so that flexible parameters can direct a range of

movements within a given class of behaviors (Coker, 2004; Lee & Swinnen, 1993; Magill, 2001; Schmidt & Wrisberg, 2004). Invariant features include the relative timing structure between component parts, sequencing of parts, and relative force in performing the action (Coker, 2004; Lee & Swinnen, 1993; Magill, 2001; Schmidt & Wrisberg, 2004; Schmidt, 1975, 1983). Variant features, i.e., parameters of the GMP, customize the movement by directing the general dimensions of the movement. Parameters include overall time (duration), force, and muscle selection (Coker, 2004; Hall, 1989; Magill, 2001; Schmidt, 1975).

During exposure to a stimulus, information about the initial conditions of the environment, the parameters modulating the program, the outcome of the action, and the sensory feedback are abstracted to refine and update the recall and recognition schema (the motor program; Hall, 1989; Schmidt, 1975). The recall schema abstracts information about the parameters and their associated outcomes, which directs future selection of parameters in similar circumstances (Hall, 1989; Schmidt, 1975). The recognition schema utilizes information about the initial conditions, outcomes, and feedback to provide an estimate of what sensory consequences should be expected in subsequent trials (i.e., generates an error signal; Hall, 1989; Schmidt, 1975). Motor learning occurs when the recognition schema aids the recall schema in selecting parameters that will reduce the overall error between the expected outcome and the actual outcome (Coker, 2004; Cummings & Caprarola, 1986; Schmidt, 1975).

Historically, response latency studies in motor behavior have yielded evidence consistent with motor programs. Typical results indicate a relationship between increased response latency and increased complexity of movement, which is attributed to an increase in the number of motor programs sequenced together (Henry & Rogers, 1960; Rosenbaum, 1980; Schmidt & Wrisberg, 2004; Sternberg, Knoll, Monsell, & Wright, 1988; Sternberg, Monsell, Knoll, & Wright, 1978).

Anticipatory errors in movement productions (e.g., spoonerisms; Fromkin, 1980; Garrett, 1982; Norman, 1981) and anticipated pre-motor responses during obstruction of movement patterns (Wadman, Denier van der Gon, Geuze, & Mol, 1979) yield additional evidence consistent with motor program theory. Schmidt (1975) noted, "The most impressive kind of evidence that could be generated in support of the schema is that subjects can produce movements of a given class that they had, strictly speaking, never performed previously" (p. 245). This transfer is attributed to the variability-of-practice hypothesis, which claims the more variable the practice conditions in refining the schema, the stronger the representation (Chamberlin & Magill, 1992b; Lee & Swinnen, 1993). This hypothesis is based on how the two schemas (recall and recognition) interact with one another. Variable practice conditions allow for a wide array of parameters and associated outcomes to be generated, which provides a more flexible rule to apply in novel transfer situations (Lee & Swinnen, 1993; Schmidt & Wrisberg, 2004). However, overall evidence consistent with the variability-of-practice hypothesis is equivocal (cf. reviews by Shapiro & Schmidt, 1982; van Rossum, 1987).

2.3.3 Speech production

There are many parallels between information-processing models of general motor behavior and speech production. Speech is a complex act involving hierarchal control and programming of multiple components in a time-sensitive manner (Sternberg et al., 1978). There are multiple levels of processing occurring: programming and sequencing at a variety of linguistic levels (e.g., syntax, lemma), mapping the message onto phonetic targets, and further mapping onto the systems of the vocal tract (D. E. Meyer & Gordon, 1985; Munhall, 1993; Van der Merwe, 1997). Articulations at the lower level of the hierarchy require ballistic, accurate movements to result in

specific acoustic goals. These factors supported the adoption of generalized motor program theory (Schmidt, 1975) in the motor speech research community.

The application of GMP theory (Schmidt, 1975) in speech production requires identification of invariant features (i.e., motor programs). Both invariant timing structures (e.g., Ballard et al., 2007; Caruso, Abbs, & Gracco, 1988; Gracco & Abbs, 1986; Martin, 1972; Sternberg et al., 1988) and force parameters (e.g., Gracco & Abbs, 1986; Knock et al., 2000; Meigh & Shaiman, 2010; Shaiman, McNeil, Szuminsky, Meigh, & Kotler, 2006) have been reported for speech production. Irrespective of this evidence, the motor program unit of speech has been difficult to determine as invariant structures of timing and force are distributed across multiple programming levels during speech production (Gracco, 1994; Martin, 1972). Many units have been proposed, including phonemes (e.g., Ballard, Robin, Knock, & Schmidt, 1999; Goffman & Smith, 1999; Gracco, 1990, 1991; Guenther, 1994), syllables (e.g., Aichert & Ziegler, 2004; Cholin, Levelt, & Schiller, 2006; Cholin & Levelt, 2009; Levelt et al., 1999; Sevald, Dell, & Cole, 1995), words (e.g., Klapp, 2003), and phrases (e.g., A. Smith, Goffman, Zelaznik, Ying, & McGillem, 1995; Varley, Whiteside, Windsor, & Fisher, 2006; Youmans, Youmans, & Hancock, 2011).

Smaller units, e.g., the phoneme, have been proposed as viable motor programs for treatment in apraxia of speech (AOS; e.g., Austermann-Hula et al., 2008; Ballard et al., 2007, 1999; Knock et al., 2000). These studies manipulated practice (Ballard et al., 2007, 1999; Knock et al., 2000) and feedback (Austermann-Hula et al., 2008) conditions, which are variables associated with enhanced motor program development. All of the studies investigated the place and manner of articulation (parameters) of specific phoneme units (motor programs). Positive acquisition and transfer of treatment targets varied depending on the complexity of the phoneme

and severity of the apraxia. Method issues (e.g., counterbalancing of treatments, small sample sizes) limit the generalization of these results. It is unclear from these studies if the motor program, i.e., the targeted phonemes, was the incorrect motor program to target in treatment.

As noted earlier, other investigators have postulated larger motor program units, e.g., syllable, for speech production. Syllable stress placement in words and nonwords is dependent on a variety of factors, including syllable weight (i.e., vowel length and number of coda consonants; e.g., Chomsky & Halle, 1968; Guion, Clark, Harada, & Wayland, 2003; Hayes, 1982), linguistic foot structure (e.g., Guion et al., 2003), and frequency of occurrence (i.e., high- versus low-frequency; Aichert & Ziegler, 2004; Cholin et al., 2006; Cholin & Levelt, 2009; Laganaro, 2005, 2008). These variables provide a predictable set of patterns (or rules) for placement of syllable stress in English words (Chomsky & Halle, 1968; Guion et al., 2003; Hayes, 1982), which has provided some evidence of the syllable unit as a motor program (Aichert & Ziegler, 2004; Cholin et al., 2006; Levelt et al., 1999; Maas & Mailend, 2012). Additional comparisons between Schmidt's (1975) GMP theory and the perceptual characteristics of syllable stress are also consistent with the hypothesis of stressed syllables as motor programs, see Table 1. Levelt (1999) postulated a similar relationship where stored syllable programs could vary along free parameters (e.g., force, duration, pitch) to meet the variety of articulatory needs during motor programming (as reviewed in Cholin et al., 2006).

Table 1: Shared characteristics between GMPs and Stressed Syllables

Schmidt's (1975) motor program invariant features	**Characteristics of Stressed Syllables**
Relative timing	Increase in duration
Relative force	Increase in pitch and intensity
Motor program is the proposed programming unit in complex, sequential actions	Proposed programming unit by some speech investigators (e.g., Aichert & Ziegler, 2004; Cholin et al., 2006; Levelt et al., 1999; Sevald et al., 1995)

Other rule-based theories outside of GMP theory (Schmidt, 1975) may also provide an explanation of normal and disordered speech production. Speech production theories applying a more comprehensive view of schema, i.e., not limiting the rule-based behavior to parameters and motor programs, have also been proposed in normal speech motor control theory (e.g., Cholin & Levelt, 2009; Guenther, Ghosh, & Tourville, 2006; Sevald et al., 1995; A. Smith et al., 1995). These theories depend on a unit of analysis, such as a syllable, to denote the schema being trained; however, there is little agreement about the rules, unit size, or representational form of these schemes. Thus, currently there is no unifying framework for rule-based learning in the speech motor control literature.

2.4 TRANSFER

2.4.1 Cognitive Tasks

The generalized transfer effect noted with rule-based learning is explained by the abstraction process: the resultant centralized representation is *context-limited*, which allows it to generalize

to a wide array of new stimuli sharing the same abstracted features (Doody & Zelaznik, 1988; Logan, 1988; Medin & Ross, 1989; Neal & Hesketh, 1997; Nokes & Ohlsson, 2003; Shanks, 1995). However, transfer effects vary depending on how similar the abstracted features are between two tasks (Shanks, 1995). As discussed in section 2.2.1, narrow transfer effects are observed with novel items very similar to the prototype, whereas broader transfer effects are observed with novel items less similar to the prototype. The variability of transfer effects observed may be due to the task and its influence on abstraction. As stated previously, rule-based learning is an umbrella theory representing a variety of memory forms and describing a variety cognitive and motor tasks. For each type of rule-based representation, the level of abstraction will influence the degree of transfer (Shanks, 1995). Consider the brief encounters required for abstraction of a prototype – transfer is often very narrow in scope. However, abstraction occurring over multiple exposures may produce production rules that generalize to a broader, or diverse, set of experiences. This variety in transfer effects is noted in the work of Nokes and Ohlsson (2003), who advocate for a continuum of transfer spanning near (within-domain) to far (cross-domain) tasks. These authors advocate there may be several strategies for transfer, and it is the match between the different processes and task demands directing the generalization to novel items across this continuum (Nokes, 2005, 2004, 2009; Nokes & Ohlsson, 2003).

For this dissertation, a more restricted view of transfer effects will be adopted, which is consistent with transfer predictions in motor theory. Motor control theorists using general schema models (e.g., Chamberlin & Magill, 1992a, 1992b; Crump & Logan, 2010) have adopted the definition of "rule" posed in the introduction. This definition proposes generalization will be the *same*, or uniform, across a class of behaviors in which all novel behaviors are governed by the same abstracted schema (Chamberlin & Magill, 1992a, 1992b; Crump & Logan, 2010;

Logan, 1988; Shanks, 1995). The class of behaviors is determined by the underlying invariant features of the motor program (e.g., relative timing and force parameters, sequencing of movements; Magill, 2001; Schmidt, 1975; C. H. Shea & Wulf, 2005; Gabriele Wulf & Schmidt, 1988). As long as two motor behaviors share the same invariant features, i.e., motor program, transfer will occur in the same uniform pattern (Chamberlin & Magill, 1992a, 1992b). Motor class assignment is important to rule-based theory as it details when generalization will occur (i.e., across within-class behaviors) and when it will cease (i.e., for any outside-class behaviors; Gabriele Wulf & Schmidt, 1988). Uniform transfer patterns and class effects in rule-based learning provide a theoretical divergence between rule- and instance-based learning theories. As will be described in Chapter 3, instance-based learning predicts transfer between trained and untrained stimuli sharing "relevant" and "irrelevant" features. Transfer predicated by similarity between trained and untrained stimuli is difficult in rule-based learning due to the abstraction process. The relevant features, or "rules," may not have enough individualized detail for a comparison to novel behaviors based on similarity alone (Chamberlin & Magill, 1992b; Shanks, 1995; E. E. Smith et al., 1992).

2.4.2 Motor Tasks

The topic of transfer in motor theory is often constrained to the direction of transfer, i.e., positive (facilitation of a skill from training) or negative (decrement in skill; Kleinman, 1983; Magill, 2001). This description does not dictate how or why the direction of transfer may be occurring. Often investigators attribute transfer effects to Thorndike's (1903) theory of identical elements (e.g., J. A. Adams, 1987; Coker, 2004; Kleinman, 1983; Magill, 2001). However, there is little consensus what elements need to be identical between trained and untrained motor behaviors for

transfer to occur (Rosalie & Muller, 2012). Theories of transfer-appropriate processing (C. D. Morris, Bransford, & Franks, 1977) provide an alternate explanation of transfer in motor behavior (e.g. Coker, 2004; Lee, 1988; Magill, 2001); however, identifying which cognitive processes are shared between trained and untrained motor behaviors is also problematic (Rosalie & Muller, 2012).

The pattern of transfer described in general schema theory (Chamberlin & Magill, 1992a, 1992b; Crump & Logan, 2010) is not explicitly described in Schmidt's (1975) GMP theory. The original studies investigating GMPs, and much of the validating investigations following, focused on dissociating the parameters from the program (e.g., Lai, Shea, Wulf, & Wright, 2000; Lai & Shea, 1999; Whitacre & Shea, 2000; D. L. Wright & Shea, 2001; G. Wulf, Schmidt, & Deubel, 1993; G. Wulf et al., 1993; G. Wulf & Schmidt, 1989). However, further inspection of this literature reveals uniform transfer patterns are observed with invariant features of the motor program, see Figure 2 (adapted from D. L. Wright & Shea, 2001, fig. 1).

Wright and Shea trained participants on a speeded sequential key-press task with three invariant timing structures in either a blocked or random training session. Participants' performance was evaluated using a retention and transfer task following a 24-hour period. As noted in Figure 2, the relative timing errors associated with the trained motor program (i.e., AE prop) are uniform in their direction with minimal variability during retention and transfer phases of the experiment. Invariant timing data in other motor tasks have similar patterns of uniform transfer (e.g., Lai et al., 2000; Magnuson & Wright, 2004; G. Wulf et al., 1993; G. Wulf & Schmidt, 1989).

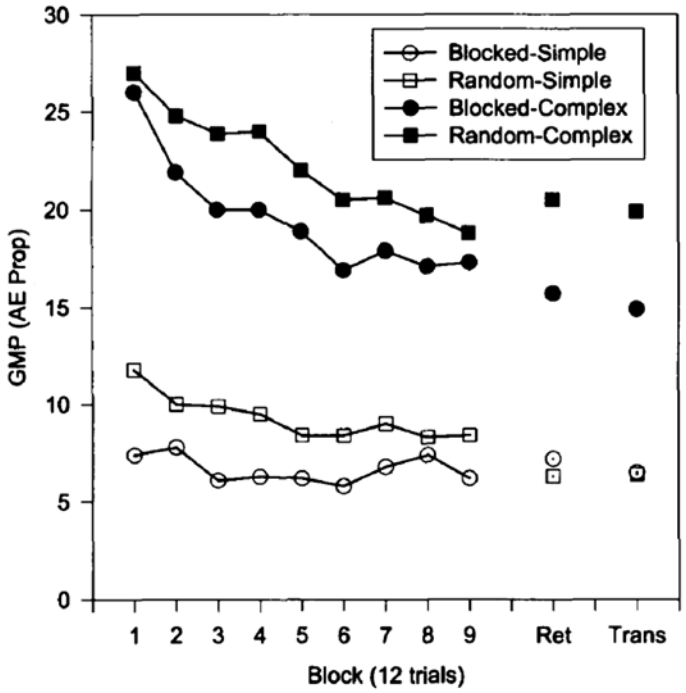

Figure 2: Example of uniform pattern of generalization in a GMP study

Note: 1 Reprinted from *Research Quarterly for Exercise & Sport*, *72*(1), D. L. Wright & Shea, Manipulating generalized motor program difficulty during blocked and random practice does not affect parameter learning, p. 35, 2001, with permission from Taylor and Francis.

Schmidt (1975) and Chamberlin and Magill (1992a) postulated all untrained behaviors within a motor class are directed by the same underlying abstracted rules (or invariant features) of the trained GMP. Uniform transfer would *not* be anticipated with outside class motor behaviors as these behaviors do not share the same invariant features as the trained GMP. Thus, motor class provides a boundary for when transfer effects are no longer predicted to occur for a trained movement (Chamberlin & Magill, 1992b; Gabriele Wulf & Schmidt, 1988).

Class effects are unique to rule-based learning as instance-based learning relies on similarity (not rules) to predict transfer. Although motor behaviors within a class may be similar to one another (Palmeri, 1997), the variables of motor class and similarity may be dissociated, see Fig. 3. In instance-based learning, decreases in similarity between trained and untrained behaviors results in poor transfer regardless of motor class. However, rule-based learning maintains transfer effects as long as the class boundary is maintained. This divergence in transfer pattern is noted when similarity of the untrained behavior decreases within the same motor class.

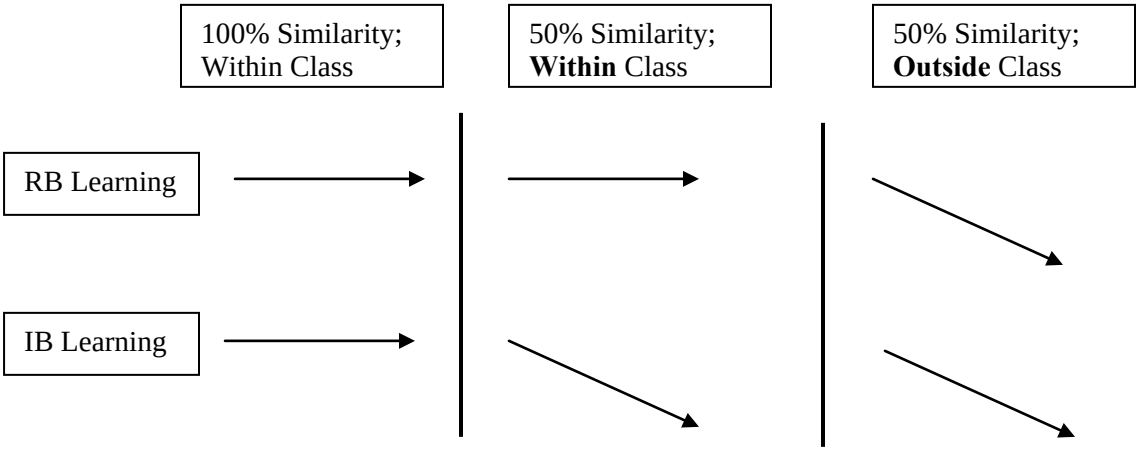

Figure 3: Effect of motor class on different theories of learning

Note: 2 RB = rule-based; IB = instance-based

2.5 FUNDAMENTAL FLAWS

There are several flaws with rule-based learning that need to be addressed despite its popularity in cognitive psychology and motor learning theory. Given the breadth of interpretations in rule-based learning theory, it is not possible to provide a detailed description of the flaws and inconsistent results associated with specific rule-based models. Instead, a general overview of flaws inherent to *all* rule-based theories will be described.

2.5.1 Abstraction

The role of abstraction has been described in depth in this chapter as it signifies the main difference between rule- and instance-based learning theories. Abstraction is the process of identifying and refining the relevant features of a stimulus (or behavior) during a learning encounter. Incoming sensory information is either preserved, forming a summarized representation, or discarded from memory. However, evidence from categorization studies suggests information is not lost and irrelevant features can be recalled (e.g., Jacoby & Brooks, 1984; Jacoby, 1978; Medin & Ross, 1989; Medin, 1986; Neal & Hesketh, 1997; Ross, 1984, 1987; Shanks, 1995; Witt & Vinter, 2012). This finding has also been reported in perceptual categorization in which perception of non-native language contrasts can be perceived in adults with various retrieval cues (e.g., Best, McRoberts, & Sithole, 1988; Werker & Logan, 1985; Werker & Tees, 1984b). This is in contrast to well-documented, replicated research in rule-based perception in infants, in which language contrasts are summarized into categorical boundaries by the end of the first year of life (Best et al., 1988; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Lalonde, 1988; Werker & Tees, 1984a). These exceptions may not mean that the

process of abstraction is invalid but theories utilizing this construct need to specify what is abstracted, as well as the cues enabling retrieval of relevant ("maintained") and irrelevant ("discarded") information.

2.5.2 Defining Rules

A consistent argument against rule-based learning is the difficulty in defining the set of rules for a given class of behaviors (Shanks & St. John, 1994). GMP theory provides some specification of the invariant features inherent in a motor program; yet, how much variability is allowed with the invariant features has been contested (Lee & Swinnen, 1993). Coarticulation, motor equivalence, and general motor dynamics in speech production impose variable mapping conditions between an articulatory gesture and an acoustic goal. Thus, even with invariant features, such as timing and force, there must be a continuum of movement patterns to allow for these variable conditions. In line with this difficulty is the challenge of defining the class of actions or behaviors delineating the limits of generalization (Hall, 1989; Lee & Swinnen, 1993; Zelaznik, 1977). The class of actions should all share the same GMP, or invariant features; however, without a firm description of the invariant features it becomes even more difficult to discern the class of actions shared by a program. The problem of under-specification and lack of detail in defining the representations in rule-based behavior is not unique to this form of learning. Instance-based learning, too, must provide detailed information about what features are encoded in an instance representation. Further specification of the features inherent in any representation will only strengthen that theory of learning.

2.6 SUMMARY

This chapter provided an overview of the history of rule-based learning in cognitive psychology and motor learning theory. Schema theories, in particular Schmidt's (1975) GMP theory, were highlighted in the representation and transfer sections. Expectations about learning (via abstraction) and transfer effects were also outlined and contrasted with those from instance-based learning. Chapter 3 emphasizes additional contrasts between these theories in terms of representational form and transfer effects.

Rule-based learning is inferred from a wide variety of experimental tasks in cognitive psychology and motor theory. Extrapolation of rule-based learning outside the laboratory has been reported in cognitive psychology (e.g., Fong et al., 1986; Larrick et al., 1990; Nisbett, Fong, Lehman, & Cheng, 1987). Many research participants, when asked to describe how they learned a specific experimental task, respond with rules they generated during the research experiment (E. E. Smith et al., 1992). These accounts do not seem surprising as the idea of rule-learning and storage is cognitively "economical." Less cognitive space and/or capacity should be needed for smaller, more finite content compared to retaining entire sensory packets of information about a stimulus encounter (Shanks, 1995). Indeed, it is this parsimonious account of processing and memory representation that has made rule-based learning so attractive to motor theorists.

Despite its popularity and wide-spread application, however, the rule-based learning approach has serious flaws. Speech production theories lack evidence for GMPs and struggle to define motor programs. Other representational forms of rule-based learning may fare better in speech production theory (e.g., neural networks); however, defining the underlying rules for these models is also likely to be an issue. The overt flaws of rule-based learning, in speech

production and in general, are most obvious when contrasted with instance-based learning, which will be highlighted in the next chapter.

3.0 INSTANCE-BASED LEARNING

3.1 INTRODUCTION

Instance-based theories of memory came into vogue in the late 1970s and early 1980s as an alternative to rule-based abstraction theories (Jacoby & Brooks, 1984; Shanks & St. John, 1994). As noted in Chapter 2, rule-based learning predicts abstraction of relevant features will occur during an encounter with a stimulus, which results in a summarized, context-limited memory representation (Jacoby & Brooks, 1984). In contrast, instance-based theories (also known as episodic trace or exemplar memory theories) predict all features of a stimulus will be encoded, resulting in a context-specific memory representation (Hintzman, 1976, 1986; Jacoby & Brooks, 1984; Jacoby, Marriot, & Collins, 1990; Neal & Hesketh, 1997; Shanks, 1995). Multiple encounters with the same stimulus result in multiple memory traces being encoded (Hintzman, 1976, 1986; Jacoby & Brooks, 1984).

Advocates of instance-based theory differ in their opinion about the permanence of the memory trace. Proponents of permanent trace models assume all memory traces are encoded as permanent records that are never revised. When new, but similar, instances are encountered they are encoded along with the original memory trace (Barsalou, 1990; Hintzman, 1986; Medin & Schaffer, 1978). Others advocate for a revisable trace model in which original traces are modified, or updated, as new similar instances are encoded (Barsalou, 1990). Revisable trace

theories are difficult to distinguish from rule-based theories as revising a memory trace appears similar to the process of abstraction (Barsalou, 1990; Logan, 1988). This may be why motor theories do not describe instance-based learning using revisable trace theory, but instead rely on other rule-based terminology, e.g., schema. Due to the difficulty in distinguishing rule-based learning from revisable trace theory, and the exclusive application of permanent trace models in motor control theory, only permanent trace models will be reviewed. This chapter will provide a historical review of instance-based behavior, representational form, transfer and generalization effects, and flaws of instance based theories.

3.2 HISTORICAL OVERVIEW

3.2.1 Frequency

Frequency judgments during memory tests are not easily explained by rule-based learning, and instead were attributed to early accounts of instance-based learning. During these tasks, participants are asked to estimate how many times a stimulus is presented during training (Hintzman, 1976). Rule-based theories (and revisable trace theories) predict a single memory trace is strengthened with an increase in stimuli presentation frequency. The information not encoded in the strengthened trace is lost and no longer accessible from memory (Barsalou, 1990; Hintzman, 1976). Permanent trace (or instance-based learning) theories predict each encounter with a stimulus is encoded separately even if a similar, or exact, encounter has been previously encoded (Barsalou, 1990; Hintzman, 1986, 1988; Jacoby & Brooks, 1984). Loss of information within a permanent trace theory can only occur if an entire instance, or exemplar, degenerates

(Barsalou, 1990). Evidence in frequency judgment experiments suggests information about frequency is not lost following exposure to the stimuli. For example, subjects in frequency judgment experiments can report the time stamp of when a particular stimulus occurred relative to other presentations of the same stimulus (Flexser & Bower, 1974; Hintzman & Block, 1971; Hintzman, 1976). Revised trace theories would predict a convergence of time-stamp (or other item-specific) information over the course of exposures, rendering the subject unable to specify presentation, or stimulus-specific, information at the end of the experiment (Hintzman & Block, 1971; Hintzman, 1976; Wells, 1974).

3.2.2 Categorization

In Chapter 2, prototype development in categorization was described. During analytic generalization relevant stimulus features are encoded and irrelevant features are discarded (Jacoby & Brooks, 1984; see Medin & Ross, 1989 for similar proposal). In contrast, instance-based theories rely on nonanalytic generalization, in which all of the features (relevant and irrelevant) are encoded together (cf. Jacoby & Brooks, 1984; Mathews et al., 1989). Retrieval of any feature associated with the memory recalls the entire instance representation; thus, categorization requires a level of specificity and experience (Jacoby & Brooks, 1984; Medin & Ross, 1989; Medin, 1986). Both specificity and experience are operationally defined within instance-based theories. Specificity is the encoding and retrieval of *all* features of the stimulus into an exemplar representation. Experience is the accumulation of exemplars over time into a knowledge base (Barsalou, 1990; Logan, 1988; Medin & Ross, 1989). Both specificity and experience work together to generate categorization effects. New, unidentified stimuli are encoded and similar instances are retrieved directing categorization of the item. With each

stimulus encounter the exemplar knowledge base expands making it easier, and faster, to categorize the stimulus (Ashby & Maddox, 2005; Dopkins & Gleason, 1997; Masson, 1986; Mathews et al., 1989; Shanks, 1995). Categorization may occur by comparing similarity between stimuli (Brooks, 1978; Cock, Berry, & Gaffan, 1994; Masson, 1986; Medin, Altom, Edelson, & Freko, 1982; Witt & Vinter, 2012) or dissimilarity between stimuli (Ward & Churchill, 1998; R. L. Wright & Burton, 1995); however, it is the variable of similarity of the new stimulus to a stored representation driving categorization (not the application of a rule; Ashby & Maddox, 2005; Masson, 1986; Shanks, 1995; Ward & Churchill, 1998; Witt & Vinter, 2012).

As noted in Chapter 2, empirical evidence in categorization suggests participants can retrieve relevant and irrelevant features pertaining to a stimulus (e.g., Jacoby & Brooks, 1984; Jacoby, 1978; Medin & Ross, 1989; Medin, 1986; Neal & Hesketh, 1997; Ross, 1984, 1987; Shanks, 1995; Witt & Vinter, 2012). This finding is inconsistent with rule-based categorization because irrelevant information would have been discarded during analytic generalization. Instance-based learning in categorization also has ecological validity as individuals tend to look for examples within their environment that are similar to the new stimuli they are encountering (Brooks, 1978, 1987; Medin & Ross, 1989). Finally, instance-based categorization is simplified from prototype categorization in which the relevant features must first be gleaned, then compared to a stored prototype, and then classified (Chamberlin & Magill, 1992a; Jacoby & Brooks, 1984; Medin & Ross, 1989; Medin & Schaffer, 1978).

3.3 REPRESENTATIONAL FORM

Defining an instance (or exemplar) varies depending on a variety of factors: task demands (e.g., perceptual learning, categorization, motor learning), level of modeling (e.g., computational, networking, theoretical), as well as the permanence of the instance representation. The majority of instance-based representations have been formulated in cognitive psychology, with limited extrapolation to motor theory. The next section will introduce the various instance-based learning forms in cognitive psychology first, followed by those formulated in motor theory.

3.3.1 Cognitive

Contrasted with prototypes, schemas, or other rule-based learning representations, instance representations are comprehensive in their depiction of an encounter with a stimulus. Information is not lost through the process of abstraction. Instead, information is directly encoded as a whole instance unit, which is stored in memory with similar other comprehensive instances (Jacoby & Brooks, 1984; Logan, 1988; Medin & Ross, 1989; Shanks, 1995). During encoding, similar instance representations are encoded close to one another and dissimilar instances are encoded farther apart in memory (Nosofsky, Little, Donkin, & Fific, 2011; Nosofsky, 1986, 1992; Shanks, 1995). Two variables, attention processing and environmental context, determine what is encoded into an instance representation (Jacoby et al., 1990; Shanks, 1995). Attention directs memory encoding and retrieval processes, which determine what co-occurrences or events will make up an instance representation (Boronat & Logan, 1997; Logan & Etherton, 1994; Neal & Hesketh, 1997). Environmental context provides further specification of the instance by directing attention processing to specific task demands or “irrelevant”

features.[1] Instance representation encoding has been likened to taking a picture with a camera in which the representation of the memory is a "snapshot" of the encounter with the stimulus (Shanks, 1995). In this metaphor, both attention and environmental context interact with one another to create a comprehensive and highly specific memory representation.

Logan (1988) referred to instance representations as processing episodes, which "consist of the goal the subject was trying to attain, the stimuli encountered in pursuit of that goal, the [subject's] interpretation [of] the stimuli with respect to the goal, and the response made to the stimulus" (p. 495). Structural features (i.e., physical attributes), environmental conditions, and mental operations occurring at the time of the stimulus encounter are encoded in a instance representation (Gonzalez, 2012; Medin & Schaffer, 1978; Neal & Hesketh, 1997; Shanks, 1995; Whittlesea & Dorken, 1993). As will be described in Section 3.4, transfer effects are determined by the similarity between these varying components of the instance representation and the novel transfer stimuli (Neal & Hesketh, 1997; Shanks, 1995). See Figure 4 for a depiction of the potential co-occurrences that may be encoded or retrieved for a given instance.

[1] "Irrelevant" features are those features discarded in rule-based learning that provide specification to the instance representation (Jacoby & Brooks, 1984)

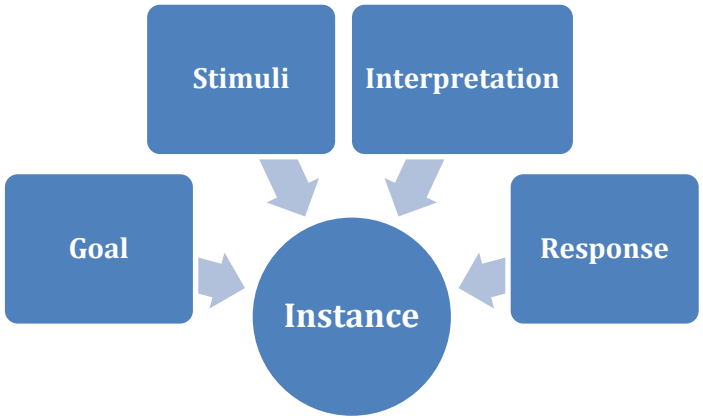

Figure 4: Instance representation

Note: Adapted from *Psychological Review*, *85*(3), Medin & Schaffer, Context theory of classification learning, p. 211, 1978, with permission from American Psychology Association.

Instance representations within cognitive psychology have been conceptualized as theoretical memory units (e.g., Brooks & Vokey, 1991; Dopkins & Gleason, 1997; Hintzman, 1984; Neal & Hesketh, 1997), neural networks[2] (e.g., McClelland & Rumelhart, 1985; Nosofsky, Kruschke, & McKinley, 1992; Rodrigues & Murre, 2007; Sakamoto, Matsuka, & Love, 2004; Shanks, 1991), production rules or propositions (e.g., Gonzalez & Dutt, 2011; Gonzalez, Lerch, & Lebiere, 2003; Logan, 2005), as well as algorithms (e.g., Aha, Kibler, & Albert, 1991; Aha, 1997; Cost &

[2] McClelland & Rumelhart (1985) reproduced the instance-based results of Whittlesea (1987) using a neural network model. However, their simulations were also consistent with activated prototype representations. McClelland and Rumelhart argued that connectionist models were not simply a realization of rule- or instance-based representations but a separate entity entirely (cf.,Broadbent, 1985; reviewed in Shanks, 1995).

Salzberg, 1993; Domingos, 1996; Nosofsky, 1988). All of these varied representations, however, incorporate the same idea: the memory representation is comprehensive. All memory representations encode the context and environmental cues, general goals, features, and interpretations associated with the stimuli.

3.3.2 General Motor

Chamberlin and Magill (1992a) noted instance-based models of motor control are attractive in their explanation of item-specific transfer effects, their focus on environmental and contextual influences affecting motor performance, and their efficient computation. Yet, instance-based learning in motor theory is limited to models of arm reaching (Chamberlin & Magill, 1992b; Rosenbaum, Loukopoulos, Meulenbroek, Vaughan, & Engelbrecht, 1995; Rosenbaum, Meulenbroek, Vaughan, & Jansen, 2001) and typing (Crump & Logan, 2010). Item-specific transfer effects are common in the motor control literature, but are rarely attributed to instance-based learning. This may be due to the popularity of schema theory proliferating the motor learning literature and influencing the interpretation of experimental results (e.g., Keetch, Schmidt, Lee, & Young, 2005).

Keetch et al. (2005) evaluated experienced basketball players' ability to perform set shots versus jump shots from a variety of distances on the basketball court. Schema theory predicts variable practice will enhance the motor program being trained (e.g., basketball shots; Doody & Zelaznik, 1988), and predicts uniform transfer for behaviors within the same motor class (e.g., set and jump shots). This predicted pattern of transfer was observed with jump shots, where practice at a variety of distances on the basketball court resulted in uniform performance accuracy. However, set shots were performed more accurately for a specific distance (15-feet

from the foul line) despite variable training. Keetch et al. (2005) theorized the item-specific transfer effects resulted from a new parameter being generated for the generalized motor program *or* from a new generalized motor program being created through training. These theories are difficult to distinguish from one another as parameters and motor programs are difficult to define (K. M. Newell, 2003; C. H. Shea & Wulf, 2005). As there is no specification for the original motor program, it is difficult to determine how it was updated or evolved with practice. Instance-based learning was not posed as a theoretical alternative by these authors.

An advocate of instance-based learning may have viewed Keetch et al.'s (2005) results as individual exemplars being encoded for each basketball shot. Within each shot the distance to the basket, the weight of the basketball, the intention to throw the ball, other variables in the environment (e.g., audience, shouting from team mates) and from within the player (e.g., fatigue, motivation) may have been encoded. Each subsequent shot would retrieve any information similar to the given circumstances, which would direct performance of future shots (Chamberlin & Magill, 1992b). Item-specific transfer with set shots may be attributed to the environmental context. The environment for set shots is inherently less variable than those of jump shots (e.g., no interactions from other players, occurs from the foul line), which increases the similarity between encoded instance representations and restricts retrieval for these particular shots to a select knowledge base of exemplars. The generalized transfer results noted with jump shots may be due to the increased variability of information being encoded, which would result in a variety of instance representations being retrieved from a more varied database. See Figure 5 for a visual depiction of how specificity and experience may generate different knowledge bases. Thus, an instance-based account of learning could describes both types of transfer effects noted in Keetch et al.'s (2005) data.

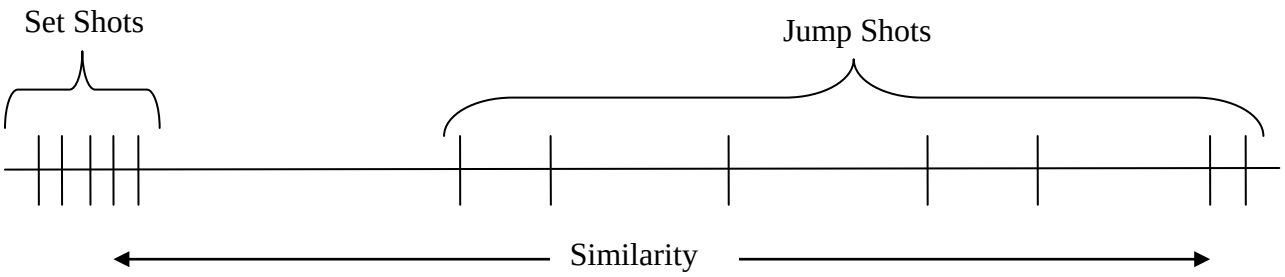

Figure 5:Visual depiction of an instance-based explanation of Keetch et al.'s (2005) data

Keetch et al. (2005) are not unique in attributing item-specific transfer effects to rule-based learning (as reviewed in: Chamberlin & Magill, 1992a, 1992b; Crump & Logan, 2010; Magill, 2001). Currently, there are only two models of motor control that attribute item-specific transfer to instance-based learning. Rosenbaum and colleagues' (Rosenbaum et al., 1995; Rosenbaum, Meulenbroek, et al., 2001) computational model, Knowledge II, is a dynamical model of hand reaching. Knowledge II was created, in part, to computationally model the motor equivalence problem first described by Bernstein (1967), which indicates there are more variables for a motor system than are required to complete a given task. Knowledge II overcomes the motor equivalence problem by storing smaller units (movement trajectories) versus larger movement units (Rosenbaum et al., 1995).

Knowledge II encodes and retrieves motor postures (composed of a variety of vectors) as individual exemplars in memory (Rosenbaum et al., 1995). When an individual encounters a motor goal, an assignment of values (from 0-1) is applied to each stored exemplar posture (and its associated vectors). The vectors are then combined into a single mathematical trajectory directing movement (Rosenbaum et al., 1995). Later versions of Knowledge II (Rosenbaum, Meulenbroek, et al., 2001) have applied different aggregation principles to compute the trajectory, but the overall idea is the same: the retrieved instances are weighted based on

similarity to the task demands, aggregated, and performed with the overall goal of meeting the task demands.

The Knowledge II model has been evaluated in handwriting tasks (Meulenbroek, Thomassen, Rosenbaum, Loukopoulos, & Vaughan, 1996) and arm reaching using discrete variables (Rosenbaum, Meulenbroek, et al., 2001). Although it has not been applied to the head and neck musculature, this approach has the potential to provide both an instance, as well as a dynamic, model of speech production. As a computational model, it may coincide with current neural network models of speech (e.g., DIVA, Guenther, 1994); however, the overall task goals between the DIVA model and Knowledge II are different in scope (cf. invariant acoustic targets, Guenther, Hampson, & Johnson, 1998; dynamic postures, Rosenbaum et al., 2001). Knowledge II disregards larger (potentially higher-order) units of analysis and interactions with language (like other dynamical models of speech, e.g., Saltzman & Munhall, 1989) and would need to be updated to reflect these higher-order units.

Chamberlin and Magill (1992b) and Crump and Logan (2010) evaluated instance representation in motor behavior through an information-processing model more akin to the studies presented in Section 3.3.1. The description of these studies, as well as critiques of the theory and methods, are presented in Chapter 4. They are noted here as evidence of the limited extension of instance-based learning and memory representation in general motor theory.

3.3.3 Speech Production

Instance-based learning has not been empirically investigated in speech motor control theory; however, reference to instance-based representations is documented in one study. Youmans, Youmans, and Hancock (2011) utilized Logan's (1988) interpretation of instance representation

to justify training phrase-length scripts with individuals with AOS. Phrase-length scripts were considered context-specific memory units retrieved during adult speech production. These larger memory units were considered “instance-based” due to the large amount of information encoded during training, which was contrasted to smaller memory representations (phonemes) evaluated in earlier AOS treatment studies. During training, the authors applied motor learning rules designed to refine motor programs (S. G. Adams, Page, & Jog, 2002; Clark, 2003; Knock et al., 2000; G. Wulf et al., 1993), and the reported uniform transfer effects were consistent with rule-based learning. Thus, this study is really an example of rule-based learning involving larger rule-based representations. As noted in Chapter 2, other larger units may exist for rule-based representations in speech motor control (e.g., A. Smith et al., 1995; Varley et al., 2006) despite only smaller phoneme units being evaluated in AOS treatments (e.g., Austermann-Hula et al., 2008; Ballard et al., 2007, 1999; Knock et al., 2000).

Currently, there are no formalized models of instance-based learning in speech motor control theory. However, predictions regarding the parameters encoded in a speech instance memory representation may be derived from descriptions of instance-representations in cognitive tasks, see Section 3.3.1. Motor instance representations would involve environmental/contextual cues, action plans activated in previous contexts, the current action plan, and the success of the action plan to meet the current goals (Chamberlin & Magill, 1992b). Instance representation in speech production would vary from general motor instance representations due to the differences in goals (i.e., acoustic goal, Guenther et al., 1998), and may include more contextual information related to communication success (e.g., communication partners’ verbal and nonverbal cues).

Instance-based learning in speech motor control would require flexible aggregation of instance representations to meet the dynamic and kinematic sub-goals supporting acoustic

targets, and these memories would need to represent a variety of articulatory and vocal tract configurations (Gracco, 1994; Perkell, Matthies, Svirsky, & Jordan, 1995). To be specific, instance representations of speech need to overcome the problem of motor equivalence (Bernstein, 1967). Rosenbaum (1995; 2001) represented the instance memory as a vector in his Knowledge II model to decrease the effects of potential memory storage problems (also, Perruchet, 1994; Rosenbaum, 1980). However, other investigators propose larger movement sequences may be encoded as instance representations (e.g., Brooks & Vokey, 1991; Brooks, 1978; Logan, 1988). The debate regarding the size of the instance memory representation is similar to the debate regarding the motor program unit described in Chapter 2; however, the idea of a fundamental instance representation may not be valid in instance-based learning (Jacoby & Brooks, 1984). Instance-based learning is directed by the task environment and demands (Jacoby et al., 1990; Shanks, 1995), which directs encoding and retrieval of the instance memory representation (Jacoby & Brooks, 1984; Logan, 1988). During speech production, the communicative demands, biomechanical constraints of the vocal tract, etc would all be encoded as features within the instance memory. The retrieval of these features may occur at a variety of processing levels (e.g., articulatory gesture, syllable, word, phrase levels), which provides a flexible motor system capable of ballistic, fluid articulatory movements. Within such a model, the motor memory representation does not need to be defined (e.g., motor program for phonemes) because the aggregation of instances will be determined by the environment and task demands. Similar complex, hierarchal models of instance-based learning have been postulated for reading comprehension (Rawson, 2004) and word recognition (Nielsen, 2011; Pierrehumbert, 2001, 2002; M. Walsh, Möbius, Wade, & Schütze, 2010).

3.4 TRANSFER

During instance-based learning, encoding and retrieval of specific, comprehensive memory representations occur with each encounter with a stimulus (Jacoby & Brooks, 1984; Jacoby, 1978; Medin & Ross, 1989; Medin & Schaffer, 1978; Medin, 1986; Ross, 1984, 1987). Over time a knowledge base is constructed in which similar experiences are retrieved and direct transfer effects (Barsalou, 1990; Jacoby & Brooks, 1984; Mathews et al., 1989). Retrieval is based on the contexts and/or variables encoded at the time of the encounter; portions of the stimuli not encoded are not available as retrieval cues to aid transfer (Medin & Schaffer, 1978).

Abstraction is absent from instance-based learning; consequently, a different mechanism is required to drive transfer effects. A gradient of transfer effects can be observed along a similarity index within a stimuli set in which increased similarity equates to increased transfer (Ashby & Maddox, 2005; Brooks, 1978; Chamberlin & Magill, 1992b; Hintzman, 1986; Masson, 1986; Mathews et al., 1989; Palmeri, 1997; Shanks, 1995). Transfer effects to new, un-similar stimuli are poor as there are few exemplars residing in memory to aid generalization (Chamberlin & Magill, 1992b; Logan, 1988; Masson, 1986; Shanks, 1995). Strong versions of instance-based theory state transfer between trained and untrained stimuli may only occur if the trained stimuli have been previously encoded (e.g., “item-specific” transfer, Logan, 1988, p. 494). Weaker models of instance-based transfer evaluate the similarity between trained and untrained stimuli using a mathematical computation to determine the potential degree of transfer (e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997).

The differences in transfer predictions between rule-and instance-based learning are illustrated in the same hypothetical data set in Figure 6. A category boundary distinguishing dimension 1 from dimension 2 is represented by the dotted line in the data set. During instance-

based learning, item-specific transfer of similar stimuli across a given dimension (e.g., dimension 1) will result in data being close to the dotted line. Dissimilar stimuli would be represented farther away on a given dimension (e.g., the diamonds farther along the x-axis along dimension 2), indicating poor transfer. Alternatively during rule-based learning, rules are abstracted and direct transfer. This data set has two rule sets, one for each dimension. Any stimuli adhering to the rules of dimension 1 would be classified in the dimension 1 space on the figure; other stimuli using dimension 2 rules result in transfer in the dimension 2 space underneath the dotted line.

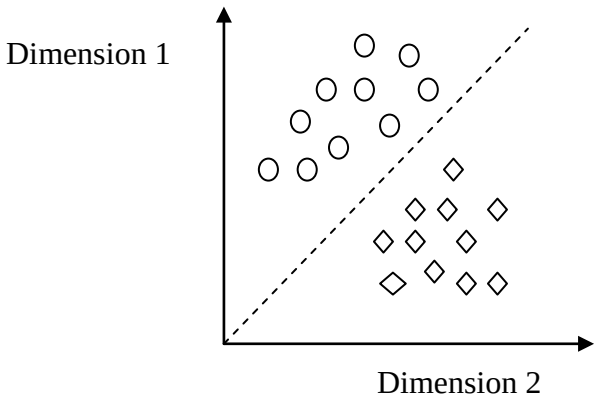

Figure 6: Instance- versus- rule based learning in hypothetical stimuli

Note: Adapted from *The psychology of associative learning*, Shanks, p. 153, 1995, with permission from Cambridge University Press

Distinctions in class of behavior can also be hypothetically observed in Figure 6 in which each dimension would consist of its own class of behaviors (e.g., each circle in dimension 1 resides within the dimension 1 class or circle class). Transfer based on rules would adhere to class boundaries to determine when one rule (e.g., relating to circle classification) no longer applies (e.g., in the dimension 2 space). However, class distinctions are not relevant for instance-

based learning because transfer is based on similarity (Shanks, 1995). Items within-class are typically more similar to one another than items between-class, which increases transfer effects for within-category learning (Palmeri, 1997). However, these variables could be experimentally manipulated to evaluate transfer effects separate from class membership as noted in Chapter 2 (Figure 3). Untrained stimuli with characteristics of both dimension 1 and 2 (e.g., a rhombic triacontahedron) would result in positive transfer in an instance-based learning system even though the untrained shape crosses class boundaries. As this hypothetical data set illustrates, each learning theory predicts unique transfer patterns according to similarity and class of the trained and untrained stimuli.

The hypothetical data set in Figure 6 is based on simple one-to-one stimulus matching between trained and untrained shapes. However, most advocates of instance-based learning argue pooling, or aggregation, across instances occurs for complex behaviors to transfer (Jacoby & Brooks, 1984; Logan, 1988; Mathews et al., 1989; Medin & Ross, 1989; Neal & Hesketh, 1997). From a motor perspective, aggregation of instance representations seems inevitable as even relatively simple motor activities require coordination across a number of motor and sensory systems (e.g., articulating a single phoneme requires recruitment of several muscles in the vocal tract). Different models of instance-based learning aggregate differently (cf. Rosenbaum's (1995) trajectory of vectors to Palmeri's (1997) similarity algorithm) but the overall idea is stability of performance occurs during memory retrieval as similar instances are aggregated together (Crump & Logan, 2010; Jacoby & Brooks, 1984). Jacoby and Brooks (1984) indicated aggregation consists of a "…very local average at the *moment* of a test, a local chorus of instances" (p. 24; italics added). This online aggregation may be differentiated from a schema, or other rule-based behavior, as it occurs during *retrieval* versus encoding when abstraction

occurs (Crump & Logan, 2010; Neal & Hesketh, 1997). Consequently, aggregation of exemplars during memory retrieval may retrieve all components of the stored exemplar not just the mean, abstracted portion of the memory representation (Medin & Ross, 1989).

Evidence consistent with instance-based learning and item-specific transfer has been reported in a variety of work: frequency judgments (e.g., Flexser & Bower, 1974; Hintzman & Block, 1971; Hintzman, 1976, 2010; Jacoby & Brooks, 1984; Jacoby, 1978), categorization (e.g., Homa et al., 1981; Jacoby & Brooks, 1984; Malt, 1989; Medin & Ross, 1989; Medin & Schaffer, 1978; Nosofsky, 1986), semantic memory (e.g., Landauer, 1975), social judgment (e.g., Kahneman & Miller, 1986; E. R. Smith & Zárate, 1992), negative priming (e.g., Neill, Valdes, Terry, & Gorfein, 1992; Neill & Valdes, 1992; Neill, 1997), counting (e.g., Lassaline & Logan, 1993), problem solving (e.g., Gonzalez et al., 2003; Lejarraga, Dutt, & Gonzalez, 2012; Medin & Ross, 1989; Ross, 1984), artificial grammar learning (e.g., Jamieson & Mewhort, 2009a, 2010), sequential learning in serial reaction time (e.g., Jamieson & Mewhort, 2009b), neuroscience modeling (e.g., Helie & Ashby, 2009; Helie, Waldschmidt, & Ashby, 2010), cognitive skill learning (e.g., Touron, Hooyer, & Cerella, 2001), text comprehension (e.g., Rawson & Middleton, 2009), typing (e.g., Crump & Logan, 2010), and reverse mirror reading (e.g., Masson, 1986). Additionally, instance-based transfer effects align well with other memory retrieval phenomena, such as encoding specificity (Tulving & Thomson, 1973) and transfer-appropriate processing (C. D. Morris et al., 1977), in which encoding and retrieval of a given memory rely on similarity for transfer to occur (Neal & Hesketh, 1997).

3.5 FUNDAMENTAL FLAWS

Although there is evidence consistent with instance-based learning across the domains of cognitive psychology, and increased potential for application of this theory in motor control, there are several concerns with this theory of learning.

3.5.1 Storage

One potential fatal flaw for instance-based learning is the issue of memory storage. If every instance representation is permanently encoded and stored for future retrieval is there enough capacity to withstand a lifetime of instance-based learning? This problem is what makes rule-based learning so attractive: irrelevant information is discarded and rules take the place of a knowledge base of instances. In motor control theory, the storage problem is of great concern because of motor equivalence. Motor goals can be obtained by a variety of different movement trajectories and structures – are all potential movements stored for all structures?

Depending on the representational form, the issue of storage may not be catastrophic. As noted in Section 3.3, instances may be manifested in neural network models in which the representation is inherent within the pattern of activation or weights of network activity. These types of models do not rely on a “black box” in which theoretical memory units are stored but instead let the representation emerge during activity (Masson, 1990; Reisberg, 2001; Rumelhart, 1989; Shanks, 1995). It seems storage problems may only be an issue if the level of analysis is more theoretical in which memory representations reflect a “unit” physically stored within the brain.

The storage problem is not unique to instance-based learning, as rule-based learning may also suffer from this issue (Jacoby & Brooks, 1984). It is not clear how much storage is allocated for rule-based representation and if that limit is reached throughout an individual's life. Instance-based theories employ aggregation techniques to "save space," which condenses multiple instances into a new single instance representation; however, the question persists of how much space is saved (Barsalou, 1990). Though the storage problem is not unique to instance-based learning it is more overt; thus, any accounts of instance-based learning in speech motor control must account for this issue.

3.5.2 Forgetting

Forgetting may be a fatal flaw for instance-based theories depending on how the phenomenon of forgetting is characterized. Forgetting, as defined as a loss of information from memory, could be devastating for any behavior relying on instance-based representation (J. R. Anderson, 1992; Logan, 1995; Rawson & Middleton, 2009; Shanks, 1995; Wilberg & Guay, 1985). Entire behaviors could be lost without some portion of the instance being available for retrieval. Yet, there is little evidence to suggest genuine forgetting, based on this definition, actually occurs in cognition (Galotti, 1994; Shanks, 1995; Shiffrin & Schneider, 1977) or motor behavior (Jahnke & Duncan, 1956; Wilberg & Guay, 1985).[3] Instead, individuals who appear to "forget" information employ inappropriate retrieval cues to recover the memory representation (Chandler,

[3] Many motor theorists attribute performance decrements to a "forgotten" motor program (e.g., J. A. Adams, 1987; Lee & Magill, 1985; J. B. Shea & Wright, 1991). It is unclear if these motor investigators truly believe that information is lost or if it is just irretrievable given cueing.

1991, 1993; Galotti, 1994; D. Lieberman, 1990; Shanks, 1995; Shiffrin & Schneider, 1977). Interference from learning new information (Chandler, 1991, 1993), ineffective rehearsal and maintenance strategies (Carlson, Sullivan, & Schneider, 1989; Terry, 2000), premature termination of the memory search (Shiffrin & Schneider, 1977; Shiffrin, 1970, 1976), as well as poor matching between encoding and retrieval environments (Jacoby & Brooks, 1984; D. Lieberman, 1990; C. D. Morris et al., 1977; Shiffrin & Schneider, 1977; Terry, 2000) can impede retrieval of specific stored instance representations. Thus, the instance representations are not lost or decayed; they have become inaccessible given the retrieval cue employed by the individual (Shanks, 1995). Consequently, forgetting, as defined as a loss of information, is not truly a problem for instance-based learning theories. However, understanding how information is retrieved from memory would be essential in implementing this type of learning theory in a practical way (e.g., motor rehabilitation).

3.6 SUMMARY

This chapter described instance-based learning through a historical framework, as well as defined instance representation forms in the cognitive and motor domains. Predicted transfer patterns were described, and were based on encoding and retrieval characteristics of instance representations that share similar features. Contrasts between instance- and rule-based learning theories were also described to further distinguish these two frameworks of learning and memory representation.

Generally, instance-based learning has been more readily adopted by investigators in cognitive psychology than by investigators in motor control theory. Though there are still

proponents of rule-based learning in cognitive psychology, there is evidence consistent with instance-based theory in the areas of frequency and category judgments (as well as many other phenomena, as outlined in section 3.4). There are disagreements as to the representational form, as well as the permanence of the memory representation, within instance-based learning. However, all investigators agree encoding and retrieval occur each time a stimulus is encountered. This type of learning generates a highly-specific knowledge base, promoting generalization to untrained stimuli sharing features to stored instances, i.e., item-specific transfer.

The limited application of instance-based learning in motor control theory is curious given motor learning effects are often consistent with item-specific transfer. The application of rule-based theory to describe these effects (as noted in Keetch et al., 2005) may be problematic if instance-based representations are directing motor behavior. As learning and transfer effects vary greatly between these two learning theories, strategies advancing transfer in one theory of learning may not promote transfer in the other theory. For instance, training with an instance-based learning framework may rely on strategies to focus attention to enable retrieval of instance representations. This is in contrast to current training regimens using a rule-based learning in which feedback schedules are manipulated to refine motor programs. Further empirical investigation is warranted to work toward distinguishing these two learning theories in motor behavior given the evidence consistent with both types of learning. Two studies in general motor control have attempted this comparison and will be presented and critiqued in the next chapter.

4.0 COMPARISONS OF LEARNING THEORIESS IN MOTOR BEHAVIOR

4.1 INTRODUCTION

The overviews provided in Chapter 2, rule-based learning, and Chapter 3, instance-based learning, provide a basis for understanding these two learning frameworks and how they may be realized within a model of motor control. Direct comparisons of these two theories are nonexistent in the speech production literature and limited to two studies in the general motor literature. The first study to contrast these frameworks, conducted by Chamberlin and Magill (1992b), reported evidence for rule-based learning only. No further empirical evaluation of these learning theories was undertaken again until Crump and Logan (2010), who concluded instance-based learning characterized performance in two typewriting tasks.

Though the conflicting nature of these results is of interest, the underlying methods of both studies may render them ineffective in critically examining rule- versus instance-based learning in motor behavior. To preface the discussion of validity problems in these studies, hierarchical information-processing models of motor control will first be described. Following this, assessments of each motor control level will be presented. This will inform the critical review and evaluation of Chamberlin and Magill's (1992b) and Crump and Logan (2010)'s experiments, which will lead to the rationale for the variables and methods in the proposed dissertation study.

4.2 EVALUATING THE REPRESENTATION

4.2.1 Models of motor performance

Both rule- and instance-based learning rely on information-processing models of motor control in which a memory representation of the motor behavior drives the execution, or production, of that motor act. These motor control models rely on a hierarchy in which upper levels control lower levels by organizing and fine-tuning the representation as processing travels down the levels (Hulstijn & Van Galen, 1983; D. E. Meyer & Gordon, 1985; Rogers & Storkel, 1998; Rosenbaum, 1990; Schmidt & Wrisberg, 2004; Schmidt, 1988). Generally, information-processing models of motor control have four levels: cognitive input, response selection, response programming, and execution (Schmidt & Wrisberg, 2004; Schmidt, 1988).

During the cognitive input stage, stimulus properties are identified to determine whether the stimulus is auditory, visual, kinesthetic, proprioceptive, or some combination of these properties (Miller & Ulrich, 1998; Schmidt & Wrisberg, 2004). From a communication standpoint, the input signal may derive from a communication partner (e.g., a spoken auditory signal) or an individual's own intent to communicate. Factors influencing the cognitive input stage include: attention (Ono, 1990; Rhodes, Bullock, Verwey, Averbeck, & Page, 2004), working memory (Rhodes et al., 2004), and stimulus characteristics (e.g., clarity, complexity, intensity; Schmidt, 1988). Most speech production models would also place linguistic processing at this level (e.g., Van der Merwe, 1997). Once the stimulus input is received, a decision is made to act on the information, and a motor response is selected.

Response selection involves selecting the memory representation that will direct the intention to act put forth by the cognitive input stage (Curtis, Rao, & D'Esposito, 2004). Verwey

(2001) considered this memory representation a "goal structure of action" (p. 71). The overall goal of the movement is defined during this stage, as well as the general movement parameters for achieving the goal (Curtis et al., 2004). This memory representation is typically depicted as a motor program[4] (Klapp, 1995; McCann & Johnston, 1992); however, an instance-based interpretation is also feasible if the memory representation is conceptualized as a knowledge base of instances. These potential response alternatives are depicted in Figure 7 in which the responses selected can be viewed as a collection of individual exemplars *or* a schema.

[4] The term "motor program" as used here is more general in its scope than as used in Chapter 2 (e.g., GMP). The term here refers to a general plan of action for achieving a goal state. More finite "motor programs" are established in the response programming stage, in which the sequence of movements and other factors contribute to a specific movement pattern.

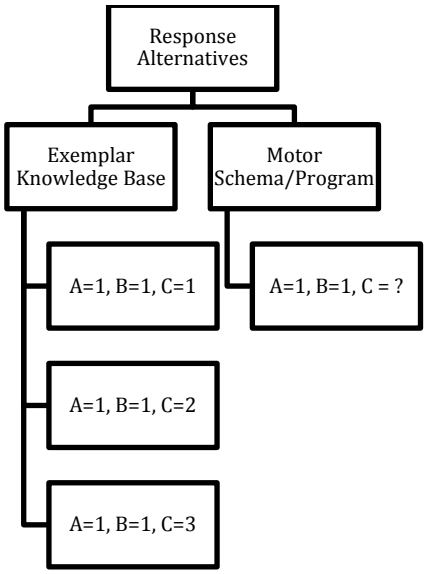

Figure 7: Potential response alternatives

Note: Adapted from *Journal of Experimental Psychology: General*, *109*(4), Rosenbaum, Human movement initiation: Specification of arm, direction, and extent, p. 447, 1980, with permission from American Psychology Association.

Response programming organizes and specifies the parameters of the selected representation (Klapp, 1995). Programming of the selected representation may involve specifying the spatial and temporal parameters of the movement, as well as stringing together several representations, or chunks, into larger, more complex movement patterns (Hulstijn & Van Galen, 1983; Klapp, 1995, 1996; Schmidt & Wrisberg, 2004). Evidence of dissociation between programming specification and serial ordering has been reported within the response programming stage (Deger & Ziegler, 2002; Klapp, 1995; Maas, Robin, Wright, & Ballard, 2008). The majority of reaction time studies in the speech production literature are focused on response programming (e.g., Deger & Ziegler, 2002; Maas & Mailend, 2012; Maas, Robin, Wright, et al., 2008; D. E. Meyer & Gordon, 1985; Rogers & Storkel, 1998).

The content of the response selection and programming stages changes over the course of learning. Early in practice (or when performing novel movements) specific temporal and spatial targets, as well as motor commands to specific muscle groups, are not yet coded. These specifications are determined during the programming stage of motor performance (Klapp, 1995). With practice, chunking occurs and the specifications of the movement parameters become associated with the representation itself (Schmidt & Wrisberg, 2004). Thus, later in practice (or when performing a well-learned task) chunks, or pre-loaded motor representations, may be selected and sent on for serial ordering only in the programming stage (Hulstijn & Van Galen, 1983; Maas & Mailend, 2012; Verwey, 2001).

At the lowest level of the model is motor execution, in which muscles contract and the behavior is realized physiologically (Hulstijn & Van Galen, 1983; Schmidt & Wrisberg, 2004). Motor output can be described by how well the movement performance met the desired goal (i.e., accuracy or error values), the kinematics of the movement (e.g., force, velocity, acceleration), or the muscular activity during the movement (e.g., electromyography; Schmidt & Lee, 2005). Analysis of an acoustic signal is an important variable of motor execution in speech production, as the movement goal results in an acoustic signal a listener, or communication partner, can understand (Guenther et al., 1998; Perkell et al., 1995).

Different models of motor control employ different terminology to describe these four levels, and may only incorporate certain levels depending on the specificity of the model. However, generally the four main components of motor performance are assumed within each model. Additionally, it is debated whether the transmission of information between stages within each model is serial (e.g., Henry & Rogers, 1960; Hulstijn & Van Galen, 1983; D. E. Meyer & Gordon, 1985) or parallel (e.g., Miller & Ulrich, 1998; Verwey, 1995, 1999, 2001). For ease of

explanation, I have presented the information-processing model of motor control as a serial model, in which each stage of the model processes information in a step-wise fashion. In this way, only one level of the model is operating at a single moment in time. However, there is a variety of research suggesting parallel processing can occur between stages, especially between the levels of response selection and programming (e.g., Rhodes et al., 2004; Verwey, 1995, 1999, 2001). For example, while end-sequence movements are finalized in the response programming buffer (i.e., ready for execution), upcoming motor responses are being selected (Verwey, 1999, 2001). The purpose of presenting these models is not to debate whether models of motor control have serial or parallel processing. The purpose, instead, is to outline a general approach to models of motor control and how their components are assessed (see Section 4.2.2) to inform an evaluation of rule- versus instance-based learning in the motor literature.

Appendix A depicts two information-processing motor control models. Limb motor control is illustrated in Figure 30, in which the "motor program" is situated at the general level of execution. Input from the higher levels of response selection and programming develop the motor program. Thus, in this context, the motor program encapsulates the memory representation, the kinematic parameters, muscle specification, and serial order of the movement (Schmidt & Wrisberg, 2004). Van der Merwe's (1997) model of speech production is depicted in Figure 31. The highest level within the model is composed of cognitive-linguistic input, followed by a large motor "planning" stage, then programming, and finally an execution stage for speech output. The motor planning stage portrays several selection processes, including phonemic and phonetic selections, which are sequenced together. The "programming" stage of this model specifies the muscles required to complete the phonetic commands. The parsing of speech and language in this model distinguishes cognitive/linguistic planning (language) and motor

planning/programming (speech). This division in the planning stage is not noted in other speech production models in which the speech components are maintained in the programming or execution stages (e.g., Levelt et al., 1999).

4.2.2 Assessing the representation

Historically, motor performance has been described at the level of motor execution to examine accuracy (e.g., absolute error, root-mean-square error), kinematic (e.g., velocity profiles), and physiologic (e.g., electromyography) performance (Schmidt & Lee, 2005). These variables are useful in providing insight into general performance curves (e.g., A. Newell & Rosenbloom, 1981), as well understanding environmental and training variables affecting learning (Schmidt & Lee, 2005). Yet, performance data derived from the execution level of the motor hierarchy cannot provide direct information about the memory representation directing the motor response. Based on the general model of motor performance presented in Section 4.2.1, this memory representation is stored and selected farther up the hierarchy. The representation is influenced by a variety of factors operating at each level of the hierarchy (D. E. Meyer & Gordon, 1985; Rogers & Storkel, 1998; Rosenbaum, 1990; Schmidt & Wrisberg, 2004; Schmidt, 1988), such as variable mapping conditions between the response and stimuli (affecting response selection), accuracy demands (affecting response programming), and biomechanical constraints (affecting execution). Historically, these factors have been experimentally manipulated and the effects examined using reaction time (Klapp, 1996; Magill, 2001).

In the motor literature, reaction time has been described as the interval of time between the offset of stimulus presentation and the beginning of the motor response (Klapp, 1996; Magill, 2001; Rosenbaum, 1980; Schmidt & Wrisberg, 2004). Note reaction time does not include the

movement or execution of the response, which is a separate measurment. The time from initiation of the movement until its completion is the movement time of the response (Magill, 2001; Rosenbaum, 1980). Thus, the overall response time is the reaction time plus the movement time (Magill, 2001), see Figure 8.

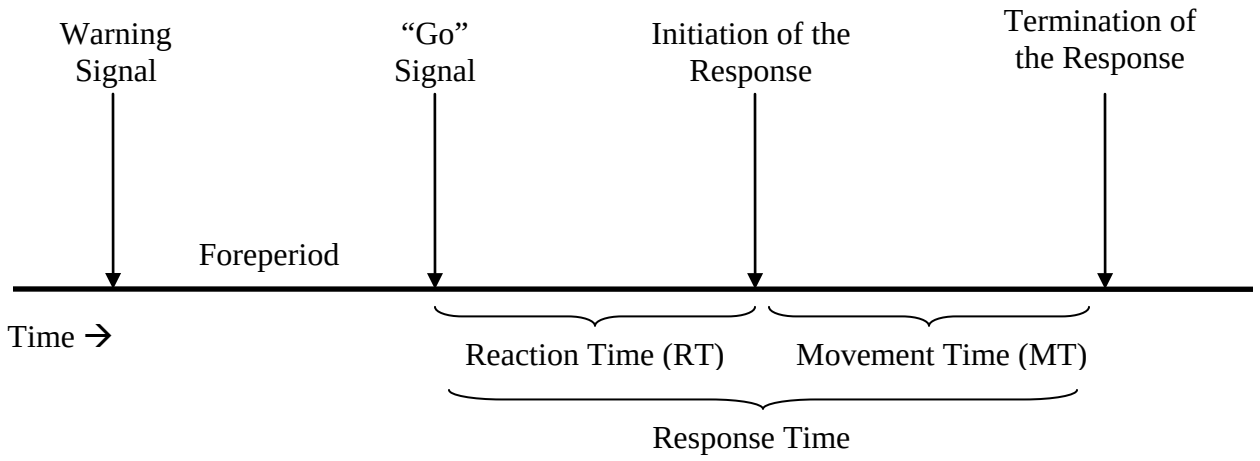

Figure 8: Overview of response time

Note: Adapted from Motor Learning: Concepts and Applications (6th ed.), Magill, pg. 25, 2001, with permission from McGraw-Hill Companies, Inc.

Reaction time and movement time are independent measures (Henry, 1961; Magill, 2001), and each can be mapped onto hierarchical information-processing models of motor control. Response selection and programming can be probed by reaction time, while execution of the movement can be assessed by movement time (Klapp, 1995, 1996; Magill, 2001). Movement time can be analyzed further using a fractionated reaction time. Premotor reaction time is defined from the initiation of electromyographic (EMG) activity to movement onset, and motor reaction time extends from the movement onset to its completion (Klapp, 1996; Maas &

Mailend, 2012; Ono, 1990). Assessments of movement time have been employed to study the effects of response dynamics on goal-directed movement (Carlton, Carlton, & Newell, 1987; Klapp, 1996), as well as attention strategies on movement efficiency (Ono, 1990).

Although assessment of reaction time may describe both response selection and response programming, different variables interact with each stage (Klapp, 1995, 2003). Increasing response uncertainty, i.e., the difficulty in selecting the appropriate response, will increase reaction time during the selection stage of motor performance (Klapp, 1995, 1996, 2003). Response uncertainty can be manipulated by decreasing the compatibility, or learned association, between the stimulus and a given response (e.g., Fitts & Seeger, 1953). Other variables increasing response uncertainty during selection include: increasing the number of possible alternative responses (e.g., Brainard, Irby, Fitts, & Alluisi, 1962; Hick, 1952; Miller & Ulrich, 1998), varying the probability of competing responses (e.g., Hyman, 1953; Kornblum, 1969, 1973), and varying the number of parameters or distinctive markers that would distinguish between two alternative responses (e.g., Heuer, 1982; Rosenbaum, 1980, 1990).

Historically, motor theorists have only examined reaction time at the level of response programming (e.g., Henry & Rogers, 1960; Klapp, 1995; Osman, Kornblum, & Meyer, 1990; Sternberg et al., 1978), which assumes selection of the memory representation has already occurred. Slowed reaction times during programming are associated with increased complexity of the motor response. Complexity may be manipulated in a variety of ways, including the number of elements within a sequence (e.g., Anson, 1982; Canic & Franks, 1989; Christina & Rose, 1985; Deger & Ziegler, 2002; Henry & Rogers, 1960; Sternberg et al., 1978), the number of effectors (e.g., Schmidt & Wrisberg, 2004), accuracy demands (e.g., Sidaway, Sekiya, & Fairweather, 1995), and movement durations (e.g., Schmidt & Wrisberg, 2004). Each of these

factors requires an increase in organization and processing to program the sequence of movements, which results in longer reaction times (Klapp, 1996; Schmidt & Wrisberg, 2004). Complex motor behaviors may involve multiple motor representations, which are sequenced in a "buffer," or cognitive space (Klapp, 1996; Osman et al., 1990; Rogers & Storkel, 1999; Verwey, 1995). The nature of the buffers varies based on the motor performance model, with some models preferring capacity limitations constraining their size (e.g., Rogers & Storkel, 1999) and other placing no restrictions on the capacity of the buffer (e.g., Deger & Ziegler, 2002). Thus, reaction times assessed at the programming level are influenced by the memory representation, interactions with complexity variables, and potential buffer capacity limitations.

As noted earlier, the information-processing models emphasized here are hierarchical, i.e., processing at one level will have an impact on subsequent stages (D. E. Meyer & Gordon, 1985; Rogers & Storkel, 1998; Rosenbaum, 1990; Schmidt & Wrisberg, 2004; Schmidt, 1988). Therefore, reaction times assessed at the level of programming may have been influenced by higher-level processing, e.g., at the levels of selection and/or cognitive input stages. Many motor theorists have attempted to assess memory representations by manipulating programming variables and recording reaction times (e.g., Chamberlin & Magill, 1992b; Crump & Logan, 2010; Gordon & Meyer, 1984; Ludlow, Connor, & Bassich, 1987); yet, a true assessment of memory representation cannot be evaluated at the level of programming. The memory representation for the motor behavior would have already been selected (i.e., processed at the response selection stage) and manipulated during the programming stage. Investigators attempting to study the memory representation of motor control via reaction time, therefore, need to access the response selection stage without the cumulative confounds of the programming or execution stages influencing the reaction time.

4.3 DIRECT COMPARISONS IN MOTOR BEHAVIOR

Chamberlin and Magill (1992b) and Crump and Logan (2010) were pioneers in introducing instance-based learning to the motor research community. Chamberlin and Magill (1992a) published a review of instance-based learning effects in frequency and categorization tasks (similar to the review provided in Chapter 3), and challenged motor investigators to systematically evaluate schema and instance-based theories using motor learning paradigms. In the same year, Chamberlin and Magill (1992b) initiated the first motor learning investigation to experimentally contrast variables of rule- and instance-based learning during a motor learning task. Almost a decade later, Crump and Logan (2010) evaluated instance-based learning in a typing paradigm. A brief review of the transfer predictions and experiments conducted by these investigators will be presented, followed by a critical analysis of the experimental methods and interpretation of the results.

4.3.1 Transfer Predictions

Chamberlin and Magill (1992b) and Crump and Logan (2010) hypothesized comparable memory representations and transfer patterns for rule- and instance-based learning. Each set of investigators evaluated these memory representations through a two-phase motor learning paradigm. During Phase One, the training phase, specific movements related to distance (Chamberlin & Magill, 1992b) or stimuli type (Crump & Logan, 2010) were practiced and encoded into memory. During Phase Two of the experiment, performance on the trained movements were compared to performance on untrained movements. The underlying memory representation encoded during training was inferred from this transfer pattern.

The rule-based memory representations encoded during training were hypothesized to be summarized representations of all the trained movements, and were postulated to transfer to untrained movements sharing similar relevant features (Crump & Logan, 2010). Transfer patterns were predicted to be uniform across trained and untrained movements within the same motor class, as all within-class stimuli were directed by the same generalized motor program and recall schema (Chamberlin & Magill, 1992b). Chamberlin and Magill explicitly stated all trained and untrained movements within their two experiments were within the same motor class, which was defined as the range of arm movements required to meet the predetermined distances in the experimental task. Rule-based memory representations were not explicitly defined in Crump and Logan's experiment. However, Crump and Logan hypothesized transfer results with a uniform pattern may be explained by schema formation resulting from training.

Instance-based representations were hypothesized by both sets of investigators as being comprehensive memory representations in which all features of the movement were encoded (i.e., relevant and irrelevant features; Chamberlin & Magill, 1992b; Crump & Logan, 2010). Transfer predictions were based on the similarity between the trained and untrained movements. Specifically, Chamberlin and Magill (1992b) defined similarity as "the congruence between the environmental conditions immediately before and during the performance of the movement and the response characteristics" (p. 311), which specifically related to the relative distances participants needed to move their arm during the training and transfer tasks. Crump and Logan defined similarity based on the features of the trained typed words, specifically whether trained letters or bigrams were present in the untrained stimuli. For both investigators, a linear relationship between similarity and transfer performance was hypothesized. Maximum transfer performance occurred when similarity was greatest between trained and untrained movements;

as the similarity decreased between trained and untrained movements, transfer performance also decreased.

4.3.2 Experiments

4.3.2.1 Chamberlin and Magill, 1992b

Chamberlin and Magill (1992b) conducted two experiments in which subjects were required to push a lever with their arm over four specified areas and three distances (short = 15cm, medium = 45cm, long = 135cm). Transfer tasks were constructed based on similarity to the training task, defined as "the relative distance over which the response was performed" (Chamberlin & Magill, 1992b, p. 311); thus, participants were required to push the arm lever to untrained distances. In both experiments, participants trained for a significant amount of time (over 1000 trials across five days of training) prior to the transfer task. However, the transfer tasks for each experiment were different. In Experiment One, all of the transfer tasks contained novel, untrained distances; Experiment Two included two novel, untrained and two trained distances. Absolute error (AE) and variable error (VE) of averaged data (blocks consisting of 50-trials) for both training and transfer phases of the two experiment were compared.

The transfer pattern associated with both experiments was uniform, i.e., not statistically different, when comparing AE values between trained and untrained movements. This effect was noted for novel, untrained transfer distances (Experiment One) and trained transfer distances (Experiment Two). Chamberlin and Magill (1992b) concluded these results were consistent with rule-based learning. Specifically, these investigators postulated training encoded a schema for the arm distances practiced during the training phase of the experiment. The untrained distances were hypothesized to be within the same motor class, which resulted in uniform transfer.

4.3.2.2 Crump and Logan, 2010a

Crump and Logan (2010) conducted two experiments in which participants practiced typing words during a training phase, and then were required to type trained and untrained words during the transfer phase of the experiment. Untrained stimuli were classified into three sets: words from the training phase (old/old set), new words constructed from trained letters (old/new set), and new words constructed from untrained letters (new/new set). The design for these untrained stimuli sets was based on Masson's (1986) seminal experiment on reading mirror-reversed words. During both experiments, participants practiced typing the trained stimuli (old/old set) for one hour (for a total of 480 trials) prior to a transfer phase in which two blocks of all the stimuli sets were presented. Thus, participants were presented with each stimuli set (i.e., old/old, old/new, and new/new) twice during transfer. Experiments One and Two were differentiated by the keyboard employed during data collection; specifically, a novel, laser keyboard in Experiment One, and a standard QWERTY keyboard in Experiment Two. Two dependent variables were evaluated: first-keypress reaction times (RTs) and interkeystroke intervals (IKSIs). RTs were defined from the onset of the presentation of the word to the first-keypress, which evaluated "early perceptual processing and processes responsible for compiling the sequential responses prior to their execution" (Crump & Logan, 2010, p. 663).[5] IKSIs were defined by a linear slope function relating the time of each keystroke to position of the keystroke within a given word. This assessed "sequential control during online response execution" (Crump & Logan, 2010, p. 663).

[5] Reaction time defined this way equate to measuring total response time, or reaction time plus movement time (Figure 4).

The transfer pattern for Experiment One and Experiment Two (Transfer Block One only) resulted in a sloping, negative trajectory of reaction times across transfer stimuli: the fastest RTs occurring on the old/old stimuli, slower RTs during the old/new stimuli, and the slowest RTs on the new/new stimuli. However, no significant differences were noted between old/new and new/new stimuli sets in Experiment Two, Transfer Block Two. The authors concluded instance-based learning had occurred during training in both experiments; however, previous exposure to the stimuli in Experiment Two, Transfer Block One, had speeded retrieval of instance representations during Transfer Block Two. Post-hoc analyses of the ISKIs for all bigrams in all transfer stimuli sets (across both experiments) were conducted to evaluate the presence of instance-based learning at all levels of processing. Results indicated trained bigram units in words and letters were typed significantly faster than untrained bigrams. The authors attribute this finding to instance-based learning operating at all hierarchical levels of typing.

4.3.3 Critical Review

Chamberlin and Magill (1992b) and Crump and Logan (2010) evaluated their results with the same theoretical framework; however, their results were attributed to two different memory representations. One postulation for this difference is motor learning requires both types of learning, rule- *and* instance-based learning, depending upon the effectors and/or task demands required. However, significant omissions in the design, methods, and stimuli suggest the validity of the interpretations put forth by these investigators may be in question. Thus, further experimentation is required to determine the influence of rule- and/or instance-based learning during motor behavior. Although there are potential confounding variables specific to each experiment (e.g., uncontrolled word length of the Crump and Logan stimuli), two critical factors

require further evaluation for both experiments. Each of these will be described separately below.

4.3.3.1 Assessing Memory Representation

Assessment at the Intended Motor Control Level

Both sets of investigators evaluated memory representation below the level of response selection, in which the memory representation is purported to be selected in typical information processing motor control models (see Section 4.2.1). Chamberlin and Magill (1992b) assessed AE and VE, which are errors in performance evaluated at the level of motor execution (the final stage in the model of motor control). Error values in motor learning tasks describe the general pattern of performance over time (Schmidt & Lee, 2005; Schmidt & Wrisberg, 2004; Schmidt, 1988); however, AE and VE values are not sensitive at distinguishing memory representations at different hierarchical levels.

Crump and Logan (2010) employed response time as their dependent variable; however, reaction time was not differentiated from movement time. Thus, it is unknown if the response time reflected processing during response selection *or* at the response programming stage of motor control. Crump and Logan also placed emphasis on their ISKI assessments as an estimate of "sequential control during online response execution" (p. 663). Although measuring the interkeystroke intervals between letters may be useful in understanding how the response programming and execution levels of motor control interact, it is not a direct assessment of the memory representation at the response selection stage. Thus, for both sets of investigators ascription of their results to either an instance or schema memory representation is invalid as an assessment evaluating memory selection was not implemented.

The assessments conducted by Chamberlin and Magill (1992b) and Crump and Logan (2010) are not unusual within the realm of motor control research (see Section 4.2.2). However, assessments of response selection could be implemented in future studies interested in evaluating motor memory representation. Event-related potentials (ERP) evaluate temporal processing characteristics of motor commands within an information processing control hierarchy, and results suggest early ERP components (e.g., N-40) may be involved in response selection (e.g., Carbonnell et al., 2013; Vidal, Burle, Grapperon, & Hasbroucq, 2011; Vidal, Grapperon, Bonnet, & Hasbroucq, 2003). Reaction times, when differentiated from movement times, may also provide an indirect evaluation of response selection during recognition tasks. Recognition tasks include identifying a single trained item (during an old-new judgment task) or several trained items amidst foils (during a forced choice recognition task), and retrieval characteristics associated with trained stimuli provide insight into memory retrieval (Hall, 1989; Radvansky, 2006). Evaluation of reaction times across untrained stimuli varying in similarity (characteristic of instance-based learning) and motor class (characteristic of rule-based learning) in a recognition task may also provide insight into the underlying memory representation being retrieved.

Assessment of Trained Stimuli versus Prior Knowledge

It is unclear in any of the experiments reviewed what effect prior knowledge may have had on the transfer tasks. This issue is not explicitly addressed by either set of investigators. However, the two-phase design of the experiments described suggests the training phase encoded *specific* memories for the training stimuli, and these memories (not prior stored memories) were to be compared to the untrained stimuli in the transfer phase of the experiment. Unfortunately, it is unclear if the experiments were actualized in this way. For example, the untrained movements in

the Chamberlin and Magill experiments may have been familiar to participants prior to the experiment. Familiar tactile and visual components of movement may activate previously learned knowledge, which may increase the accuracy and speed of motor responses (e.g., Magill, 1998; D. L. Wright, Shea, Li, & Whitacre, 1996; D. L. Wright & Shea, 1991). Thus, the uniform transfer results noted in Chamberlin and Magill's study may have been the result of previously learned arm movements, and not based solely on movements trained during the experiment. If the trained arm movements were controlled for novelty (i.e., to differentiate them from previously learned arm movements) the results would be purer in their attribution of motor learning occurring *during* the experiment. This distinction becomes important if memory representations evolve as the result of experience and skill acquisition (e.g., J. R. Anderson et al., 2004; Logan, 1988; Tomporowski, 2003). This will be discussed further in the next section.

The results of Crump and Logan's second experiment (Transfer Block Two) may have also been modulated by prior knowledge. The only difference between Experiment One and Two was the keyboard employed during data collection. The standard QWERTY keyboard in the second experiment may have provided participants with tactile cues to initiate well-learned typing sequences, increasing generalization between Transfer Blocks One and Two. Additionally, the effect of prior knowledge on the stimuli was uncontrolled. Crump and Logan only controlled for word frequency in their stimuli; however, the effects of high frequency bi- and trigram units may have increased typing speeds (D. R. Gentner, Larochelle, & Grudin, 1988). Complete control of prior knowledge within an experiment is impossible; however, experimental controls may be instituted to decrease the overall likelihood of prior knowledge impacting results. For example, reconstruction of the different stimuli sets with novel bi- or

trigram units may have decreased potential frequency effects, and yielded a better estimate of how trained stimuli compare to untrained stimuli in Crump and Logan's experiments.

Effect of Training Amount

The amount of training required varied across experiments, with the participants in Chamberlin and Magill's (1992b) study practicing twice as much as participants in Crump and Logan (2010) study. Additionally, the distribution of practice varied: multiple sessions (Chamberlin and Magill) versus a single session (Crump and Logan). It is unknown what, if any effect, these training amounts and practice schedules may have had on the trained memory representations. Skill acquisition theories vary in how memory representations evolve with experience. For example, some schema theories suggest memory representations are refined with practice (e.g., J. R. Anderson et al., 2004; Schmidt, 1975), whereas some instance-theories suggest different aggregation strategies are employed (computational versus automatic strategies; Logan, 1988).

A dual-representation model may be another potential explanation, in which an instance-representation may be encoded early in training, but evolves into a more rule-based representation as a skill becomes well-learned. Given the differences in motor tasks in the reviewed experiments, a dual-representation model is not an appropriate explanation for the varied learning theories attributed to each investigator's results. However, it may provide a potential model to examine further in future studies. For example, evaluating recognition probes of trained and untrained stimuli at specific training intervals (e.g., 250, 500, 750, and 1000 trials) may provide insight into how the trained memory representation may be evolving with practice.

Chamberlin and Magill and Crump and Logan were not interested in the effects of training, and assumed the training phase of their experiment resulted in a well-learned baseline performance level for the trained stimuli. However, given the large amount of individual

variability noted with simple and complex motor tasks involving the limbs (as reviewed in: Magill, 2001; Schmidt & Lee, 2005), controlling for the effects of training in future studies should aid the interpretation of transfer effects. One way to control for training amount is to implement an accuracy criterion to determine the end of training versus relying on a set number of repetitions or trials. By instituting an accuracy criterion, individual variability in performance is controlled as all participants are anticipated to initiate the transfer phase of the experiment with the same "baseline" performance level based on the accuracy criterion.

4.3.3.2 Assessing Different Models of Learning

Ideally, the reviewed experiments would have attempted to signify potential memory representations of *both* learning theories in their experimental design and stimuli. This would have allowed for a true comparison of rule- and instance-based learning. However, the studies omitted critical features related to both theories. Based on the literature review in Chapters 2 and 3, instance-based learning should be evaluated based on the similarity of the trained and untrained stimuli, as the construct of similarity is critical to transfer predictions. Crump and Logan (2010) instituted a similarity gradient within their stimuli, which allowed for systematic evaluation of the similar components shared between trained and untrained stimuli during the transfer phase of their experiments. Chamberlin and Magill (1992b) did not have a similarity gradient, which decreased the likelihood of observing item-specific transfer effects associated with instance-based learning. This is most overt in Experiment One in which all the transfer movements were novel to participants. It is unknown whether the results would have changed with arm movements including portions of the trained arm trajectories, as uniform transfer results were also noted for the trained movements in Experiment Two.

Additionally, neither investigation provided a separate untrained stimuli set representing movements *outside* the class of trained movements. Motor class effects should not be observed in instance-based learning because the rule-based representation directing transfer effects for a motor class is not encoded. Although stimuli within a motor class may be similar to one another, similarity and motor class can be experimentally manipulated in separate ways (see Chapter 2, Figure 3). Thus, evaluation of motor class may provide a unique variable to further distinguish rule- and instance-based learning. For example, it is unknown whether an instance- or rule-based representation directed the generalization between Transfer Blocks One and Two in Experiment Two of Crump and Logan's study. The addition of a fourth stimulus set evaluating outside class movements may have further distinguished these potential outcomes.

4.4 SUMMARY

This chapter reviewed a common class of information-processing models of motor control, and associated assessments of reaction time, to better evaluate two investigations designed to compare rule- versus instance-based learning in motor behavior. Chamberlin and Magill (1992b) and Crump and Logan (2010) provide the only comparisons of these two theories of learning in motor control theory. The inherent flaws of these research studies generate uncertainty about the validity of the results. Both sets of investigators failed to assess the memory representation at the level of response selection; thus, interpretation of their results is difficult given the hierarchical processing influences that surely occurred prior to movement execution. Additionally, certain control variables and manipulations for both series of studies were absent, allowing for bias towards one framework of learning (e.g., Chamberlin & Magill, 1992b) or an inability to

distinguish between the two frameworks (e.g., Crump & Logan, 2010). The literature review of rule- and instance-based learning in motor behavior is limited both in scope and application. Future work in differentiating these learning theories in motor behavior will need to focus on assessing the representation at the appropriate motor control level, as well as providing experimental manipulations controlling for bias while allowing for sensitivity in detecting differences.

5.0 OVERALL SUMMARY AND DISCUSSION

Two main theories of learning have been evaluated in cognitive psychology and motor control theory to describe skill acquisition: rule- and instance-based learning. During rule-based learning, abstraction of relevant features of a stimulus is encoded, which generates a generalized, summary memory representation. Untrained stimuli sharing the same relevant features, or set of rules, are considered to be within the same class. Transfer effects are uniform across a class of behaviors as all stimuli rely on the same relevant features to direct transfer. Motor class membership is a critical variable for rule-based learning, and provides a mechanism to describe when transfer will discontinue for a given set of behaviors (i.e., poor transfer effects for outside-class behaviors).

In contrast, during instance-based learning all features (both relevant and irrelevant) are encoded into an instance representation. Thus, no particular feature of a stimulus is considered "more relevant" than another, and no underlying rules about features are generated. Transfer effects are dictated by the similarity of features between trained and untrained stimuli in a linear direction, i.e., decreased similarity between stimuli will result in decreased transfer effects. There are no assumed class effects with instance-based learning as no rules are abstracted during encoding. These theories, and their associated parameters, are contrasted in Table 2.

Table 2: Rule- and instance-based learning contrasts

	Rule-based Learning	**Instance-based Learning**
Type of representation	Summary, abstracted	Exemplar
Information discarded	Yes	No
Type of Transfer	Generalized	Item-specific
Direction of Transfer	Uniform pattern	Varies with similarity
Effect of Class on Transfer	Yes; transfer only within class	No effect
Memory process	Encoding most important	Retrieval most important

Rule-based learning, specifically motor program theory, is the dominant learning theory in limb and speech motor control models. Multiple motor program units have been postulated for speech production. Research evaluating smaller units, e.g., the phoneme, provides inconsistent evidence for motor program theory and the uniform transfer effects associated with rule-based learning (e.g., Austermann-Hula et al., 2008; Ballard et al., 1999; Knock et al., 2000). It is unclear from the literature if learning a larger motor program unit would result in rule-based transfer, or if rule-based learning theory is inappropriate in describing motor learning effects. As noted in Chapter 2, larger speech units (e.g., syllable) share characteristics with motor program theory and are rule-governed (Aichert & Ziegler, 2004; Cholin et al., 2006; Levelt et al., 1999; Maas & Mailend, 2012). In English, the frequency of occurrence for syllable stress position influences the response time in recognition and articulation of syllables and words. High-frequency syllables are responded to more quickly than low-frequency syllables, which has been postulated as a difference in memory representation for each frequency type (Aichert & Ziegler, 2004; Cholin et al., 2006; Cholin & Levelt, 2009; Laganaro, 2005, 2008; Staiger & Ziegler, 2008). Thus, differences in syllable stress frequency of occurrence may provide a viable class marker for stored motor programs of high- versus low-frequency stressed syllables.

Instance-based learning is rarely attributed to motor transfer effects despite evidence consistent with instance-based transfer. Evaluation of instance-based learning in limb studies is restricted to arm reaching and typing, and has yet to be investigated as a theoretical alternative to rule-based learning in speech motor control. Instance representation units in speech production may be similar to motor programs (e.g., phoneme, syllable, word). However, motor investigators within instance-based learning theory debate the theoretical size of motor representations (e.g., vector versus bigram). Small units of speech motor control, such as phonemes, may be an appropriate representation to investigate in initial examinations of instance-based learning for several reasons. These smaller units may be contrasted to larger proposed rule-based motor programs (e.g., syllable) in a direct comparison of rule- and instance-based learning in a speech motor control framework. Dissociation between syllable and phoneme levels of speech motor control suggests two different memory representations may be operating, and motor programs for syllable stress may dictate speech programming and execution of lower-level memory representations (i.e., phonemes). Stress patterns of trained words predict novel stress patterns in untrained words and nonwords that are controlled for shared phonemes and phonological neighborhoods (as reviewed in Guion et al., 2003). Syllable boundaries are maintained during disordered speech of individuals with AOS, with the clinical signs of consonant cluster reductions and substitutions occurring in non-syllable boundaries (i.e., onset or coda within a syllable; e.g., Aichert & Ziegler, 2004; Laganaro, 2005). Additionally, phonemes as an instance-based representation provide a unit of analysis which can be systematically manipulated based on similarity in a variety of word-like contexts. Adoption of Masson's (1986) methods within a speech motor control framework (much like Crump and Logan (2010) adopted these methods for

typing) would permit investigation of the effect of phonetic similarity of words (or nonwords) across different trained phonemes.

In Chapter 4, experiments conducted by Chamberlin and Magill (1992b) and Crump and Logan (2010) were critically reviewed. These authors were pioneers in motor control theory for evaluating rule- versus instance-based learning predictions. However, their experimental manipulations were not conducive to assessing both types of memory representations, which may have biased their results. Additionally, the assessments used by these investigators to probe memory representation were invalid, as these measurements assessed the levels of response programming and/or execution of the information processing model of motor control. Thus, within motor control theory it continues to remain undetermined which memory representation directs transfer effects, and/or how these two representations may interact with one another.

The literature review provided in this dissertation indicates more research in the area of learning theory is needed to explain the inconsistent rule-based learning effects noted in speech motor therapy (e.g., AOS treatment), as well as describe why transfer effects appear to be driven by similarity. Historically, the explanation for these results has been to propose a different motor program and continue utilizing a rule-based framework. Although an instance-based theory may be more consistent with these results and more parsimonious, this avenue of theoretical exploration has not been attempted in speech motor control theory. Thus, the proposed evaluation in this dissertation, to contrast rule- and instance-based learning within a speech motor control framework, is innovative.

There are several parameters that could be contrasted between these two theories (see Table 2); however, transfer pattern and motor class are the most prominent and distinguishing parameters. Historically, the distinct transfer patterns noted with each theory have led

investigators to postulate the underlying memory representation leading to these transfer patterns (Chamberlin & Magill, 1992a; Shanks, 1995). Evaluating transfer patterns of untrained stimuli that vary on parameters of similarity (associated with instance-based learning) and motor class (associated with rule-based learning) would allow for evaluation of *both* theoretical transfer patterns. Additionally, selecting methods that specifically target the response selection stage of motor control models (versus the response programming and/or execution stages) would provide a purer estimate of the memory representations being selected for transfer. These manipulations are described in Chapter 7, and attempt to overcome the flaws present in the early examinations of these theories by Chamberlin and Magill (1992b) and Crump and Logan (2010). The main experiment described in the upcoming chapters will provide an initial theoretical evaluation of these two learning frameworks in speech motor control theory.

6.0 SPECIFIC AIMS AND RESEARCH QUESTIONS

The primary aim of the current investigation was to examine two variables associated with rule- and instance-based learning: phonetic similarity and motor class. Phonetic similarity is operationally defined as one to two adjacent phonemes in a consonant-vowel (CV) unit that is identical in the trained and untrained stimuli. Motor class is operationally defined as syllable stress position in the trained and untrained stimuli. Trained stimuli will have syllable stress on the first or second syllable, and untrained stimuli within the same motor class will also have syllable stress in these initial positions. Stimuli outside the trained motor class will have syllable stress in the final, third position of the untrained stimuli.

To evaluate this primary aim, a study was conducted to validate the stimuli for the main experiment (Specific Aim 1). In the main experiment, phonetic similarity and motor class were investigated in an old-new judgment task (Specific Aim 2). The specific aims and research questions are listed below with the potential outcomes and interpretations for each research question described in separate tables.

Specific Aim 1: To evaluate reaction time patterns across lexical and gestural frequency levels, as well as stimuli types, in a young adult population.

A pre-experimental study was conducted to validate the stimuli for the main experiment. Kendall et al.'s (2005) study was replicated to determine if the original effects related to interphonemic

transitional gestural frequency were maintained in a young adult population (Research Question 1). This replication included evaluating real word frequency in young adults as a control variable against unusual nonword findings (e.g., error measurement; Research Question 2). Additionally, the Different-Phonemes-Different-Motor-Class (DPDC) stimuli were evaluated against Kendall et al.'s nonword stimuli, which yielded a baseline measure of how the DPDC stimuli were processed in relation to the Kendall et al. stimuli for the main experiment (Research Question 3).

Research Question #1 (Nonword Frequency): Do reaction times vary across interphonemic transitional gestural frequencies in nonwords?

Original results from Kendall et al.'s (2005) study revealed older adults had faster vocal latencies when articulating nonwords with high interphonemic transitional gestural frequency compared to moderate or low frequency nonwords. No significant difference was noted between the moderate and low frequency stimuli. If the same effects were observed in young adults, the moderate and low frequency nonwords would be combined for the main experimental stimuli set. Hypotheses, outcomes, and predictions for this research question are listed in Table 3.

Table 3: Predictions for high vs. moderate vs. low frequency nonword stimuli

Hypotheses	Potential Outcomes	Potential Interpretations
H_0: Reaction times will *not* be significantly different between different nonword stimuli varying in gestural frequency.	*Non-significant statistical difference* in reaction times will be noted across frequency categories.	This finding is highly unlikely given that RTs to high-frequency stimuli were significantly different from those to moderate- and low-frequency stimuli in Kendall et al. (2005). It is anticipated that this difference will be maintained in younger adults but that additional differences may be observed.
H_1: Reaction times *will* be significantly different between different nonword stimuli varying in gestural frequency.	*Significant statistical differences in reaction times* will be noted between all three frequency categories (high, moderate, and low).	Different processing is occurring with each type of stimuli. *Moderate-and low-frequency stimuli can NOT be collapsed into an experimental training set in the main experiment. Additional stimuli will need to be added.*
	Significant statistical differences in reaction times will be noted between high frequency nonwords in comparison to moderate- and low-frequency nonwords. No differences will be noted between moderate- and low-frequency nonwords.	This is a direct replication of Kendall et al.'s (2005) findings in older adults. Similar processing is occurring for moderate- and low-frequency stimuli. *Moderate- and low-frequency stimuli can be collapsed into an experimental training set for the main experiment.*
	Significant statistical differences in reaction times will be noted between low-frequency nonwords in relation to high- and moderate-frequency nonwords. No differences will be noted between high- and moderate-frequency nonwords.	Different processing is occurring with low-frequency stimuli. *High- and moderate-frequency stimuli can be collapsed into an experimental training set for the main experiment. Transfer stimuli will need to conform to these new training stimuli.*

<u>Research Question #2 (Real Word Frequency):</u> Do reaction times vary across lexical frequency in real words?

Kendall et al. (2005) evaluated frequency effects in real words as a control against experimental error or unusual participant performance. Replication of this effect would provide the same control in the current stimuli validation study. Hypotheses, outcomes, and predictions for this research question are listed in Table 4.

Table 4: Predictions for high vs. low frequency real words

Hypotheses	Potential Outcomes	Potential Interpretations
H_0: Reaction times will be the *same* between high- and low-frequency real words.	*Non-significant statistical difference between frequency categories in real words.*	If this hypothesis is accepted there may be experimental error or unusual participant performance influencing the results.
H_1: Reaction times will be *different* between high- and low-frequency real words.	*Significant statistical difference between frequency categories in real words.*	This is the anticipated finding given the large literature documenting lexical frequency and reaction time differences. ***Direct replication of Kendall et al.'s (2005) findings in older adults.***

<u>Research Question #3 (Stimuli Type):</u> Do reaction times vary based between the Kendall et al. stimuli sets and the DPDC stimuli?

The DPDC stimuli were constructed by the author to be as different as possible from the Kendall et al. (2005) stimuli on the variables of phonetic similarity and motor class. The DPDC stimuli were constructed with phonemes not used in the Kendall et al. stimuli, and syllable stress was placed on the final, third syllable of the nonword. These results will provide a baseline of reaction time performance for these stimuli in comparison to the Kendall et al. nonwords. Hypotheses, outcomes, and predictions for this research question are listed in Table 5.

Table 5: Predictions for Kendall et al. (2005) stimuli vs. DPDC stimuli

Hypotheses	Potential Outcomes	Potential Interpretations
H_0: Reaction times will be the *same* between Kendall et al.'s (2005) nonword stimuli and the DPDC stimuli.	*Non-significant statistical difference between stimuli categories.*	This finding suggests other, additional factors may be needed to differentiate the DPDC stimuli from trained stimuli in the main experiment.
H_1: Reaction times will be *different* between Kendall et al.'s (2005) nonword stimuli and the DPDC stimuli.	*Significant statistical difference between stimuli categories.*	This is the anticipated finding based on the construction of the DPDC stimuli. *These stimuli are different from the trained stimuli, which may be due to inherent construction of the stimuli (i.e., identical phonetic elements and syllable stress pattern).*

<u>Specific Aim 2:</u> To evaluate rule- and instance-based learning theories by examining the pattern of learning when comparing trained stimuli to transfer stimuli of similar phonetic construction and motor class.

The main experiment evaluated two learning variables associated with rule- and instance-based learning: phonetic similarity and motor class (i.e., syllable stress patterns). Participants learned to produce phonetically-complex nonwords with specific stress patterns until an accuracy criterion was met. Following training, a familiarity judgment task was administered, and reaction times from this task were analyzed. The results of the main experiment evaluated transfer patterns related to phonetic similarity (Research Question 4) and syllable stress motor class assignment (Research Question 5).

<u>Research Question 4 (Phonetic Similarity):</u> Do reaction times vary based on phonetic similarity of the untrained stimuli to the trained stimuli?

The predicted pattern of transfer for rule-based learning for within-class motor behaviors is uniform performance across stimuli (Chamberlin & Magill, 1992a, 1992b; Crump & Logan, 2010; Logan, 1988; Shanks, 1995). The trained, Same-Phonemes-Same-Motor-Class (SPSC), and Different-Phonemes-Same-Motor-Class (DPSC) stimuli were predicted to be within the same motor class even though they vary on a similarity index; therefore, in the case of rule-based learning, uniform reaction times would be noted during the judgment task (see Figure 9). Statistically, this would result in a non-significant finding between types of stimuli as reaction time does not vary by stimuli type.

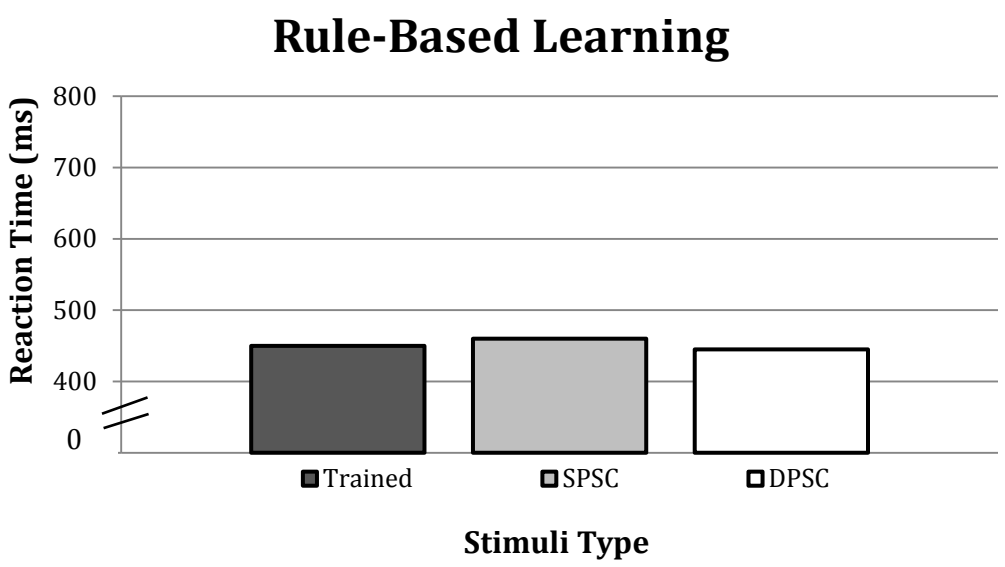

Figure 9: Rule-based learning predictions for transfer based on phonetic similarity

The predicted pattern of transfer for instance-based learning for within-class motor behaviors was a systematic slowing of reaction time as similarity between the trained and untrained stimuli decrease (Crump & Logan, 2010; Logan, 1988; Masson, 1986). Transfer is predicted across the similarity index based on stimuli type: trained stimuli would be responded to the fastest, responses to less similar stimuli (SPSC stimuli) would be slower, and responses to completely new stimuli (DPSC stimuli) would be the slowest (see Figure 10). Statistically, this will result in a significant finding between types of stimuli as reaction time would vary by phonetic similarity. Additional hypothesis, outcomes, and predictions for this research question are listed in Table 6.

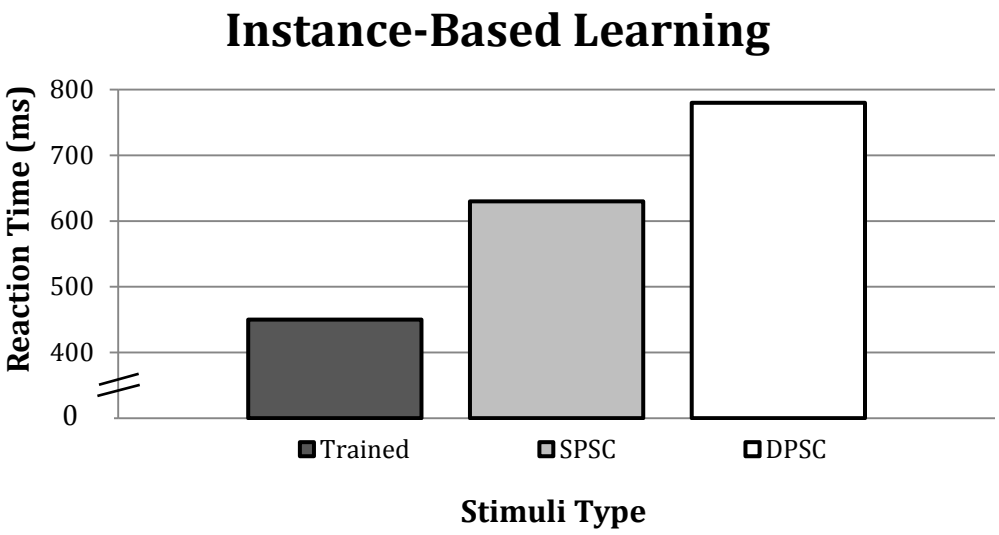

Figure 10: Instance-based learning predictions for transfer based on phonetic similarity

Table 6: Predictions for transfer pattern based on similarity

Hypotheses	Potential Outcomes	Potential Interpretations
H_0: Reaction times will *not* vary by stimuli type.	Reaction times are uniform across stimuli type. *Non-significant statistical difference between stimuli.*	Rule-based learning is occurring.
H1: Reaction times *will* vary by stimuli type.	Responses to training stimuli are fastest, to SPSC stimuli are slower, and to DPSC stimuli are slowest. S*ignificant statistical difference between stimuli type.*	Instance-based learning is occurring.
	Responses to training stimuli are slowest, to SPSC stimuli are faster, and to DPSC stimuli are fastest. S*ignificant statistical difference between stimuli type.*	Participants may be using prior knowledge to aid generalization (e.g., second language knowledge may enable generalization on portions of the nonwords). Screening assessments should rule this out (e.g., using participants who are monolingual only). Stringent control of the stimuli to assure novelty will also decrease the effects of prior knowledge.
	Reaction time patterns are random across different stimuli types. S*ignificant statistical difference between stimuli type.*	Learning may not have occurred during training. This is not anticipated given the stringent accuracy criterion required prior to the judgment task. Additionally, participants may be using prior knowledge to aid generalization. However, screening assessments should rule out this possibility.

<u>Research Question 5 (Motor Class):</u> Do reaction times vary based on motor class when similarity of the stimuli is held constant?

Rule-based learning and transfer were limited to the same class of motor behaviors (Chamberlin & Magill, 1992a, 1992b; Schmidt, 1975; Shanks, 1995). Behaviors outside of a motor class (e.g., DPDC stimuli) were not governed by the schema learned in training, resulting in slower reaction times for DPDC stimuli compared to DPSC stimuli (see Figure 11). Statistically, this would result in a significant difference in reaction times across class.

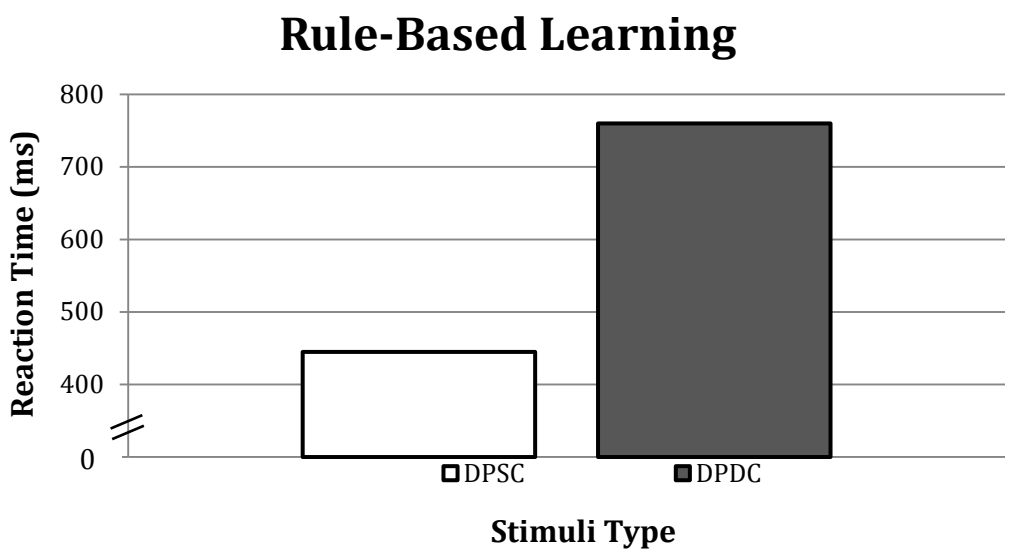

Figure 11: Rule-based learning motor class predictions

Instance-based learning and transfer were based on similarity, which may or may not be affected by motor class depending on how similar stimuli were between class types (Palmeri, 1997; Shanks, 1995). For this study, similarity of DPSC and DPDC stimuli sets was low compared to the trained stimuli; thus, transfer effects should result in slower reaction times for *both* DPSC and DPDC stimuli (see Figure 12). Statistically, this would result in a non-

significant difference in reaction times across class. Alternative hypotheses, outcomes, and predictions for this research question are listed in Table 7.

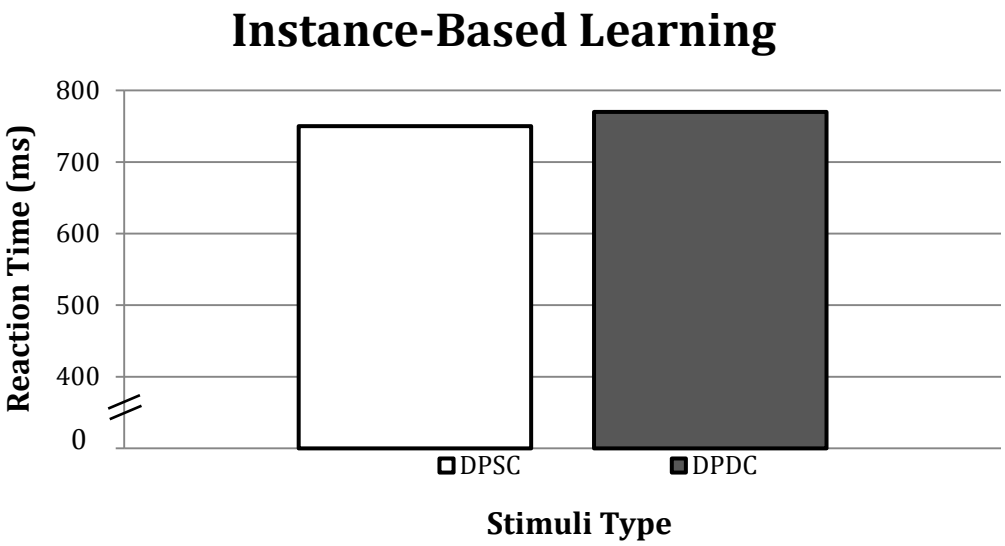

Figure 12: Instance-based learning motor class predictions

Table 7: Predictions for motor class when similarity is held constant

Hypotheses	Potential Outcomes	Potential Interpretations
H_0: Reaction times will *not* vary by class type.	Reaction times will be similarly fast for both class types. *Non-significant statistical difference between stimuli.*	The DPDC stimuli may not be outside the class of movement rules learned during training. This is unanticipated given the DPDC stimuli will have different nonwords composed of untrained phoneme consonants, as well as different syllable stress patterns not found in the training tasks.
	Reaction times will be similarly slow for both class types. *Non-significant statistical difference between stimuli.*	Instance-based learning may be occurring. Similarity between SPDC and DPDC stimuli compared to the trained stimuli is low resulting in poor transfer for both types of stimuli regardless of class.
H_1: Reaction times *will* vary by class type.	Reaction times will be slower for DPDC stimuli compared to the DPSC stimuli. S*ignificant statistical difference between stimuli type.*	Rule-based learning may be occurring. Class differences are noted as the rules/schema learned during training no longer apply to new stimuli outside the class of movements.
	Reaction times will be faster for DPDC stimuli compared to the DPSC stimuli. S*ignificant statistical difference between stimuli type.*	Participant prior knowledge or experience may cause this type of learning pattern; however, this pattern of transfer seems highly unlikely given other dissimilarities between task elements.

7.0 METHOD

This chapter is divided into two main parts. The first section describes the pre-experimental study, which validated the experimental stimuli employed in the main experiment. The second section describes the main experiment, which evaluated learning parameters associated with rule- and instance-based learning.

7.1 STIMULI VALIDATION STUDY

Prior to the implementation of the main experiment, validation of the stimuli was required to ensure the stimuli were novel at a variety of linguistic and motor levels. The application of novel, or unfamiliar, stimuli increased the likelihood participants were relying on memories encoded during experimental training and not from previous encoding experiences prior to the experimental session. Activation of previously encoded information may change anticipated learning outcomes in the main experiment. As noted in Chapter 4, the application of a standard QWERTY keyboard may have activated retrieval of well-learned motor representations for typing, increasing generalization between Transfer Blocks One and Two in Crump and Logan's (2010) experiment. Speech is a highly-practiced motor act in adults, and familiar speech-like utterances may have similar motor cueing effects that could alter the outcome of the main experiment. Other nonmotor cues (e.g., visual or tactile cues) increase accuracy and speed of

motor responses during motor learning (e.g., Crump & Logan, 2010; Magill, 1998; D. L. Wright et al., 1996; D. L. Wright & Shea, 1991). Language cues may also increase the retrieval of speech, as noted in the phonology literature in which phonologically familiar cues influence novel utterances (e.g., Ellis & Beaton, 1993; Gathercole, Frankish, Pickering, & Peaker, 1999; Kaushanskaya, Yoo, & Van Hecke, 2013; Kendall et al., 2005; Papagno & Vallar, 1992; Service & Craik, 1993; Storkel, Armbruster, & Hogan, 2006). Thus, nonword stimuli for the main study needed to control for novelty at the linguistic and motor level to control against learning effects associated with well-learned memory representations.

Kendall et al.'s (2005) nonword stimuli were employed as stimuli in the main experiment. These nonwords were controlled for novelty at the word, syllable, phoneme, and interphonemic transitional gestural level (see Table 8 for these control parameters). In particular, interphonemic gestural manipulation yielded a level of control within the stimuli not present in other studies using nonword stimuli (e.g., Dollaghan & Campbell, 1998; Roy & Chiat, 2004) by controlling the stimuli at a level of motor control, i.e., the articulatory gesture, as well as linguistic levels.

Table 8: Parameters controlled for in Kendall et al.'s (2005) nonword stimuli

Parameter	Usage
Consonant (C) and Vowel (V) structure	All nonwords were CVCVCVC.
English lexical access	Each of the three syllables did not correspond to a real word.[6]
C and V frequency	Consonant and vowels were balanced for frequency occurrence in the English language (Shriberg & Kent, 1982).
Tense and lax vowels	Tense and lax vowels were equally distributed across the nonwords.
Interphonemic transitional gestural frequency	Transitional probabilities for phoneme sequences spoken in context (Roberts, 1965) was balanced across three frequency probability categories (high, moderate, and low).

A replication and extension of Kendall et al.'s (2005) study was conducted due to differences in the research participants in the main experiment compared to the original study, as well as uncontrolled variability in stimuli duration of the original stimuli set. Kendall and colleagues' original study was conducted with older, female adults (aged 50-79 years). The main study was conducted with male and female young adult participants.

Older adults have slower reaction times compared with young adults (e.g., Craik & McDowd, 1987; Der & Deary, 2006; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994; Jevas & Yan, 2001; Luchies et al., 2002; Welford, 1977), as well as diminished and variable performance on speech production tasks (e.g., Benjamin, 1997; Gorham-Rowan & Laures-Gore, 2006; R. J. Morris & Brown Jr., 1994; Ramig & Ringel, 1983; Ryan & Burk, 1974; Torre III & Barlow, 2009). Young male and female adults perform without statistical differences on maximum performance tests of speech production (e.g., DDK rates, respiration; Kent, Kent, &

[6] This prevents lexical access of stored real word representations (Dollaghan & Campbell, 1998).

Rosenbek, 1987), as well as controlled phrases (e.g., buy bobby a puppy; A. Smith & Zelaznik, 2004; B. Walsh & Smith, 2002). Thus, age differences in reaction time and speech production skills warranted further study to evaluate if the main effects observed in the original study could be extended to a younger group of participants.

The current study also updated the original stimuli to control for mean length duration. The original high gestural frequency stimuli were longer in duration (1,173ms) than the moderate (1,080ms) or low (996ms) gestural frequency stimuli. Kendall et al.'s results indicated high gestural frequency words were produced with a lengthened final syllable, which was postulated to have occurred due to participants modeling the recorded stimuli. Post-hoc analyses indicated final syllable lengthening in participant responses was minimal (<1%), and no differences were observed between the high- and low-frequency productions in either the first or final syllable stress position. However, the potential for participants to be biased in their responses while modeling the original stimuli was still present and warranted re-recording of the stimuli for the main experiment (D. Kendall, personal communication, June 26, 2013).

7.1.1 Participant Criterion

Participants were between the ages of 18 and 35 years of age, which is the typical "young" cohort in studies evaluating normative and aging speech anatomy and physiology (e.g., Duffy, 1995; Fitch & Holbrook, 1970; Linville & Fisher, 1985; Xue, 2003). All participants were required to have an education level of at least a high school diploma or equivalent. This education requirement assured participants were able to read the sentences in the Test of Minimal Articulation (Secord, 1981), a screening tool administered to evaluate articulation competence of all phonemes in the English language. Additionally, participants needed to be

mono-lingual in English, as based on their response to a questionnaire of their native language abilities (see Appendix B). The Kendall et al. (2005) stimuli in this study controlled for novelty across a variety of parameters in English only (see Table 8). Proficiency in other languages may bias participants' responses in a faster, more efficient way if portions of the stimuli overlap with well-learned phonological representations present in the participant's native language.

If participants met this initial criterion, they were screened for normal hearing and speech production skills during the experimental session. Normal hearing was quantified as pure tone thresholds not exceeding 35 dB HL at the frequencies of 500, 1000, 2000, and 4000 Hz in at least one ear (American Speech-Language-Hearing Association, 1990). Pure tone thresholds above 35 dB HL at any of the stated frequencies in one ear disqualified the participant from the study. Additionally, participants were required to score 45/46 on List 2A (male speaker recordings) of the Northwestern University Test #6 (NU-6; Tillman & Carhart, 1966), a phonemically-balanced test approximating the phoneme distribution in English.[7] Hearing loss was not an anticipated finding for this study's sample, and the pure tone threshold and NU-6 screenings provided assurance participants were able to hear the stimuli played through an

[7] Kendall et al. (2005) also evaluated hearing ability to ensure participants would hear the recorded stimuli. However, their older population required a phoneme discrimination test weighted toward high-frequency phonemes (California Consonant Test; Owens & Schubert, 1977), which aids in the detection of high-frequency hearing loss when compared to pure tone thresholds. This type of hearing loss is frequently observed in individuals aged 60 years of age and older (e.g., A. C. Davis, 1990; Gates & Mills, 2005; Huang & Tang, 2010; Ries, 1994).

external speaker at a level within the range assessed during the hearing screening (S. Pratt, personal communication, June 27, 2013).

Auditory processing ability was also evaluated as part of the normal hearing protocol, and was assessed by the Computerized Revised Token Test (CRTT; McNeil et al., In Submission). This computer version of the Revised Token Test (Arvedson, McNeil, & West, 1985) aids diagnosis of auditory processing deficits. Participants were required to score at or above 14.14 on all subtests (first percentile) to be eligible for the study. This stringent criterion eliminated participants who may have difficulty processing the instructions and stimuli for this study.

All participants were required to present with normal anatomy and physiology of the vocal tract, which included the respiratory, laryngeal, resonatory, and articulatory systems. This was evaluated by an oral-mechanism examination (conducted by the principle investigator who is a certified and licensed speech-language pathologist), as well as evaluation of diodochokinesis rates, vowel prolongation, and observations of speech motor control during casual conversation. The oral-mechanism examination included an assessment of a) facial symmetry, b) lingual protrusion, c) mandible elevation and depression, d) labial retraction, protrusion, and closure, and e) velopharyngeal symmetry and movement. Diodochokinesis rates and vowel prolongation produced by the participant needed to be within one standard deviation of minimum normative values (Duffy, 1995). Casual conversation was assessed for any signs of motor speech or voice disorders as evaluated by the principle investigator. Any signs of anatomical or physiologic variance from normative data (Duffy, 1995), with or without a motor speech or voice disorder present, disqualified participants from the study. Finally, participants were also required to achieve 100% accuracy on the Test of Minimal Articulation Competence Screening Test (Secord, 1981), which evaluated articulation competence of all American English phonemes in

words (initial, medial, and final position) using a sentence format. A summary of all screening activities can be found in Appendix C.

7.1.2 Participant Recruitment

Participants were recruited from the University of Pittsburgh's Psychology Subject Pool. Eligibility requirements for this study were posted on the Subject Pool website, and interested participants independently selected one of the posted experimental sessions on the website. All screenings and experimental protocols were conducted by the principal investigator in a quiet lab space (room 6016 Forbes Tower) at the University of Pittsburgh. Prior to any screening or experimental protocols, written consent was obtained according to procedures outlined by the University of Pittsburgh's Institutional Review Board. Following consent, participants completed screening procedures to determine eligibility for the study as outlined above. All screening procedures took no more than 60 minutes to complete. Any unusual findings discovered during screening resulted in the principal investigator referring the participant to his or her primary care physician about the results of the screening.

7.1.3 Participant Sample Size

The proposed sample size for this study was 22 participants based on the following parameters input into the statistical power analysis program G*Power (version 3.1.7; Faul, Erdfelder, Buchner, & Lang, 2009): effect size = .26; α = .05, power = .80. This proposed participant sample allowed for a medium effect size for the proposed statistical analyses in Chapter 8 (i.e., a within-factors one-way ANOVA; J. Cohen, 1992).

7.1.4 Participant Characteristics

Twenty-six individuals participated in the consent process for this study. Three participants did not meet the screening criteria: one was excluded for being multilingual, and two did not meet criteria on the Computerized Revised Token Test (CRTT; McNeil et al., In Submission). The remaining 23 participants (10 males and 13 females) ranged in age from 18-25 years of age (*M* = 19, *SD* = 1.68), and ranged in years of education from 13-16 years (*M* = 13, *SD* = .95).

7.1.5 Experimental Procedure

After the screening protocols, participants were seated in a comfortable chair, and a headset with a unidirectional microphone (ProSeries, Pro 8Hex) was placed at approximately 2 inches mouth-to-microphone distance. A stereo speaker (Anchor, Model AN-130) was centered on the table approximately 15 inches in front of the participant. The microphone was connected to a serial response box (Psychology Software Tools; Model #200A), which was connected to a Macbook Pro utilizing Windows 7 operating system running E-prime (v. 2.0 Professional; Schneider, Eschman, & Zuccolotto, 2002).

The experimental protocol consisted of two components (training and experimental blocks), which took no more than 15 minutes to complete in a single session. The procedures conduced during the training and experimental blocks were identical. Participants listened to an auditory presentation of a stimulus item (real word or nonword) through the speaker, and then articulated the stimulus into the headset microphone. The presentation of each stimulus item was controlled by E-Prime software (version 2.0; Schneider et al., 2002), which is an experimental deployment software used to collect reaction time data with millisecond accuracy. Auditory

presentation of each trial through E-Prime consisted of a 250ms long 500 Hz warning tone, 250ms silent pause, auditory presentation of the stimulus word/nonword (mean duration: 1,121ms), 4,000ms interval of time to capture the participant's response, and a 5,000ms silent interstimulus interval prior to the next trial. E-prime captured the participant's reaction time by activating the response box using a voice key, which is a sound-activated switch permitting collection of vocal latencies detected through a microphone (Schneider et al., 2002).[8] Perceptual judgment of the participant's accuracy in producing all phonemes in the stimulus was judged by the principal investigator during data collection. The experimental session was recorded for validation of these perceptual accuracy judgments following data collection. Participants were reimbursed with course credit through the Psychology Subject Pool for their participation in this experiment.

7.1.5.1 Training Block

To ensure participants were able to trigger the voice key using the headset, a training block consisting of 10 stimuli (5 real words and 5 nonwords) was administered (see Table 9). These stimuli were taken from previous studies in the Adult Neurogenic Language Lab (under the direction of Dr. Connie Tompkins). Participants were instructed using Kendall et al.'s (2005) exact instructions: "You will hear a random series of real and made-up words. Please repeat each

[8] The response box can also collect manual response latencies via a button push. However, this study wanted to evaluate reaction time prior to *articulation* of a stimulus. Kendall and colleagues (2005) manually documented all vocal response latencies post-hoc using a KAY Computerized Speech Laboratory analysis system (Model 4300) to display the speech waveforms and wide-band spectrograms.

word after the professional speaker." Participants were required to activate the voice response key for each stimulus in the training block prior to the administration of the experimental blocks. Activation of the voice response key occurred when participants spoke directly into the microphone, and the response box switch closure mechanism was activated.[9] Sensitivity levels for voice activation were set at a level 3, which was a level sensitive to picking out voicing from respiratory exhalation.

Table 9: Training Block Stimuli

Real Words	Nonwords
Wrinkled	ɪnklɔs
Landscape	kæltiʊg
Pirate	kwɛstəs
Captain	ʤɛbo
Calories	ɛərtid

7.1.5.2 Experimental Blocks

Once the participant was able to trigger the voice key reliably, two experimental blocks were administered using identical instructions from the training block. Experimental blocks were randomized across participants to diminish potential sequencing effects. Both blocks consisted of 40 stimuli items: 20 real words and 20 nonwords. The stimuli (described below) were pseudo-

[9] The switch closure mechanism in the response box is pre-programmed by the hardware developers at PSTnet; however, a gradient of sensitivity in the amount of voicing required to trigger this switch can be manipulated with E-prime code. Sensitivity ranges from 0-9 (with 0 being less sensitive and 9 being extremely sensitive).

randomized in each list; thus, lists started and ended with real words, and no two nonword stimuli with similar phonemic structure were close to each other in the stimulus list.

7.1.6 Stimuli

7.1.6.1 Kendall et al. (2005) Nonword Stimuli

The purpose of Kendall et al.'s original study was to evaluate the effects of interphonemic transitional gestural frequency on phonetic encoding. Three groups of stimuli were created controlling for the parameters listed in Table 8, which included varying the interphonemic transitional gestural frequency across three categories: high, moderate, and low gestural frequency of occurrence (see Table 10). Participants had faster reaction times when articulating nonwords with high gestural frequencies compared to stimuli with moderate or low gestural frequencies. No statistical difference in reaction times was noted between the moderate and low gestural frequency stimuli. Replication of these reaction time effects in a younger cohort will support collapsing the moderate and low frequency stimuli into a single stimuli set for the main experiment.

Table 10: Kendall et al. (2005) nonword stimuli

High-Frequency Nonwords	Moderate-Frequency Nonwords	Low-Frequency Nonwords
zesəfis	tenærok	kozæʃɔm
zesəfin	rasæθon	fozæʃɔd
vɔsəfis	nasæθoʃ	ʃɔʤəzɔd
zirəfin	zotenav	vuzæʃɔm
disəfis	viʃədæk	dɔʤəzɔd
nʌzirəz	næθodæp	zænɔʤəθ
nʌzirəv	kæθotæs	ʤəzɔzæk
rʌzirəs	θʌrasæθ	zaʃɔʤəz
sʌzirəʤ	sæθodæk	zæʃɔʤəθ
pʌzirəʤ	sʌvenæθ	ʤʌnɔzæk

7.1.6.2 Kendall et al. (2005) Real Word Stimuli

Kendall et al. utilized real words in their original study as a control in case non-significant differences were noted between nonword frequency groups. Differences in reaction time between high- and low-frequency words were a predictable and replicated finding, and replication of frequency effects in real words would indicate negative findings in the nonword stimuli were not due to experimental error or unusual participant performance (Kendall et al., 2005). Real word stimuli, noted in Table 11, were extracted from Balota and Chumbley's (1985) lexical access study, and provided a predictable set of stimuli to evaluate lexical access frequency in real words.

Table 11: Kendall et al. (2005) real word stimuli

High-Frequency Words	Low-Frequency Words
office	nostril
mountain	trinket
student	trestle
baseball	lobster
chicken	gazelle
captain	freckle
kitchen	pebble
village	pelvis
teacher	weasel
college	rudder
machine	sequin
valley	gasket
cousin	beaker
yellow	nylon
market	banjo
forest	tunic
cotton	navel
garden	baron
coffee	ladle
window	silo

7.1.6.3 Different-Phonemes-Different-Motor-Class (DPDC) Stimuli

The DPDC stimuli were untrained stimuli created by the author for the main experiment (see Table 12). The purpose of piloting the DPDC stimuli was to provide a baseline of the reaction times associated with articulating these nonwords compared to the Kendall et al. (2005) stimuli. Descriptions about the construction of these stimuli are found in Section 7.2.7.6.

Table 12: DPDC Stimuli

ʒibʊtʃeð	gigʊðib
tʃeðugʊʒ	tʃejiwɪʒ
ʒʊgijub	bʊtʃitʃeʒ
gʊgɪðʊtʃ	tʃʊtʃubɪʒ
bɪðetʃug	gibɪðɪb

7.1.6.4 Stimuli Preparation

As noted above, the Kendall et al. (2005) stimuli were not equal in mean length duration, which potentially biased participants' reaction times on the high frequency stimuli. To decrease the potential effects of final syllable lengthening associated with the original stimuli, the Kendall et al. stimuli were re-recorded to control for mean length duration. Extensive perceptual reliability checks of the re-recorded stimuli were conducted to ensure perceptual accuracy of the nine phonemes in each nonword. Recording and perceptual ratings of phonetic accuracy will be described separately below.

Stimuli Recording

All stimuli were recorded by an individual knowledgeable in phonetic transcription and trained to maintain a slowed, natural rate of speech. All recordings were completed in a sound booth using a USB Logitech desktop microphone (Model 980186-0403) and recorded through Adobe Audition digital recording software (version 3.0) on a Macbook Pro utilizing Windows 7 operating system. Syllable stress was recorded on the first syllable for each stimulus, which is identical to Kendall et al.'s (2005) original methods.

Stimuli Perceptual Reliability

Phonetic discrimination of all nonword stimuli was conducted by three individuals trained in phonetic transcription to ensure phonetic accuracy of the recordings. Each transcriber had completed a course in descriptive phonetics and routinely transcribed in her research or clinical duties. Transcribers listened to stimuli in a quiet place while using headphones, and played the stimuli through a standard operating system player (e.g., Windows Media Player or iTunes). Several rounds of perceptual phonetic accuracy ratings were conducted in which each transcriber manually transcribed each nonword using the International Phonetic Alphabet ("IPA Chart," 2005). Two out of three transcribers needed to accurately transcribe each nonword (total of nine phonemes) for the stimulus to be accepted for stimulus validation. Stimuli not transcribed accurately by two of the three transcribers were re-recorded and further perceptual assessments were conducted.

A forced-choice questionnaire was implemented following several rounds of manual transcription and re-recording of the nonword stimuli. Transcribers listened to three nonwords and selected the correct written IPA transcription from two foils with difficult to perceive consonant contrasts. These additional ratings were deemed sensitive in detecting phonetic accuracy of all nine phonemes, especially as the foil contrasts were based on difficult to detect phonetic contrasts noted during earlier rounds of phonetic transcription attempts (M. Dickey, personal communications, September 13, 2013).

Once the nonword stimuli set was determined to be phonetically accurate, a one-way analysis of variance (ANOVA) on the mean duration as a function of frequency group (high, moderate, low) was performed. No significant differences between duration were observed

between frequency group, F (2, 27) = 2.256, p = .124; means and standard deviations for each group are listed in Table 13.

Table 13: Means and SDs for re-recorded Kendall et al. (2005) stimuli

	Mean (ms)	**SD**
High Frequency	1254	.066
Moderate Frequency	1267	.083
Low Frequency	1320	.070

7.2 RULE VERSUS INSTANCE-BASED LEARNING

The primary aim of this investigation was to evaluate two parameters of learning associated with rule- and instance-based learning theories: phonetic similarity and motor class. As noted in previous chapters, different memory representations are encoded during rule- and instance-based learning, which leads to different transfer patterns. During rule-based learning, only relevant features of stimuli are encoded, resulting in an abstracted, generalized representation of all exposures to the stimuli. These representations are referred to as motor programs in motor learning theory, in which a series of schemas direct a range of motor behaviors. Motor behaviors sharing a motor program are considered to be in the same motor class. Rule-based learning predicts uniform transfer effects within a motor class as the motor program operates similarly across behaviors. Motor class is an essential learning parameter for rule-based learning as class boundaries predict when transfer will be successful (within-class effects) versus poor (outside-class effects).

In contrast, during instance-based learning all features of the stimuli are encoded, resulting in highly-specific memory representations. In motor theory, this representation has been considered a small segment of a movement, e.g., Crump and Logan (2010)'s bigram unit in a typing task. This theory of learning has not been explored previously in speech motor control theory; however, extrapolations from Crump and Logan's experiment would suggest one or two phonemes may be a potential representational unit for instance-based learning. These highly-specific memory representations allow for direct comparison of the features between trained and untrained stimuli. Transfer occurs when similarity is maximized between trained and untrained stimuli, and decreases as the untrained stimuli become more differentiated from the trained stimuli. Due to the highly-specific construction of the memory during encoding, abstracted rules are not generated; thus, the effect of motor class is absent in instance-based learning.

To evaluate similarity and motor class parameters, two types of tasks were administered in the main experiment. First, participants engaged in a training task to build memory representations for nonword stimuli. The nonword stimuli, taken from Kendall et al. (2005), allowed for hypothesized memory representations from both learning theories to be accessible from the stimuli. For this study, instance-based learning memory representations were hypothesized to be one phoneme (or two adjacent phonemes), which is analogous to the representational units proposed by Crump and Logan (2010; see Chapter 4). Rule-based learning memory representations were hypothesized to be stressed syllable units, which also signified the separate motor classes associated with trained and untrained stimuli (see Chapter 2). Participants were required to reach an accuracy criterion to cease training, which established a baseline performance level for the trained stimuli.

Upon completion of training an old-new judgment task was administered, which required participants to make familiarity judgments while listening to different types of stimuli. Stimuli for this task included trained stimuli (highly familiar to participants) and untrained stimuli (unfamiliar to participants). All untrained stimuli varied systematically across the two learning parameters of interest: phonetic similarity and motor class. An overview of the experimental tasks and stimuli associated with each learning theory is provided in Table 14. Each portion of the experiment (both task and stimuli) is described below.

Table 14: Overview of the experimental tasks, stimuli, and learning theories

	Rule-based Learning	**Instance-Based Learning**
Training Task: What is the representation being learned?	First and second stressed syllable positions (same motor class)	1 phoneme OR 2 adjacent phonemes in a CV unit
Judgment Task: What experimental parameter is being evaluated?	*Motor class* when similarity of the stimuli is held constant	*Phonetic similarity* of adjacent phonemes within the same motor class

7.2.1 Participant Criterion

All participants were required to be between the ages of 18-35 years of age, have a minimum of a high school diploma or equivalent, be monolingual in English, and have normal speech and hearing skills as defined by the screening criterion described in the stimuli validation study (see Section 7.1.2). In addition to these criteria, participants were also informally screened for vision discrimination and physical ability to push the response buttons on the response box. Self-reports of the participant's ability to discriminate the shapes presented in the CRTT were elicited by the examiner following completion of the CRTT (McNeil et al., In Submission). This provided an

informal assessment of the participant's ability to discriminate the shapes presented in the visual feedback display during the training portion of the experiment. Participants who reported difficulty discriminating between shapes were excluded from this study even if their performance on the CRTT met inclusion criterion.

Participants were also screened to ensure their dominant hand had the strength and flexibility to push the response button on the response box. This was assessed by placing the response box in front of the participant, the examiner verbalizing either the word "yes" or "no," and the participant pushing the corresponding "yes" or "no" response button. Participants were required to achieve 100% accuracy in pushing the correct response button without reports of pain or discomfort. Any overt signs of motor difficulty noted by the examiner excluded the subject from participating in the study. These motor signs included limb apraxia; tremor of the arm, hand, or fingers; or significant joint stiffness of the hand or fingers (as assessed as decreased range of motion of the joint). See Appendix D for the screening protocol administered in the main experiment.

7.2.2 Participant Sample Size and Power Analysis

The proposed sample size for this study was 24 participants based on the following parameters input into the statistical power analysis program G*Power (version 3.1.7; Faul et al., 2009): effect size = .25; α = .05, power = .80. Crump and Logan's (2010) methods and tasks (Experiment One only) were most similar to the proposed experiment and yielded a large effect (.59; J. Cohen, 1992). However, differences in the number of within-subject measurements (three levels versus four), as well as statistical design (2 x 2 x 3 versus 1 x 4) limited the extrapolation

of this large effect size for this current study. A more conservative effect size (.25 medium effect size for a within-factors one-way ANOVA; J. Cohen, 1992) was used instead.

7.2.3 Participant Recruitment

Participant recruitment was achieved via fliers posted at the University of Pittsburgh and postings on the University of Pittsburgh's IRB-approved Clinical Translational Science Institute (CTSI) Research Subject Database. This database recruits participants from the greater Pittsburgh area who have consented to be contacted about research studies being conducted at the CTSI. Interested participants were provided with information about the study using an IRB-approved script, as well as an IRB-approved screening questionnaire about native language abilities (see Appendix B). Participants who continued to express interest and met the initial language screening criteria were scheduled for the full screening protocol.

During the experiment, participants provided consent to participate in the study using the same protocol as described for the stimuli validation experiment (see Section 7.1.2). The location of the screening and experiment were also the same (room 6016 Forbes Tower). All screening procedures took no more than 60 minutes to complete. Any unusual findings discovered during screening resulted in the principal investigator referring the participant to his or her primary care physician about the results of the screening.

7.2.4 Participant Characteristics

Twenty-nine individuals participated in the consent process for this study. Three participants did not meet screening criteria: one participant was excluded for being multilingual, one participant

did not meet criteria on the CRTT (McNeil et al., In Submission), and one participant presented with a mild dysarthria during the oral-mechanism examination. Additionally, two more participants failed to meet the accuracy criteria during the training portion of the study. These five participants were excluded from data analysis. The remaining 24 participants (11 males and 13 females) ranged in age from 18-34 years (*M* = 25.5, *SD* = 5.33), and ranged in years of education from 12-18 (*M* = 15, *SD* = 1.47).

7.2.5 Experimental Procedures and Equipment Configuration

7.2.5.1 Procedures

The main experiment consisted of a single session comprised of syllable stress training (in which nonwords with specific syllable stress patterns were encoded into memory) and an old-new judgment task (in which trained nonwords were evaluated against untrained nonwords varying in similarity and motor class). Syllable stress training was further divided into three separate tasks: perception-production training, syllable stress training, and recognition probes. Perception-production and syllable stress training were administered to encode and refine syllable stress productions associated with each nonword stimulus. For this study, the motor class variable was defined as syllable stress in either the first or second position of each nonword. The recognition probes were administered throughout training to evaluate the accuracy of the phonemes being produced during each nonword production. This ensured the accuracy of the hypothesized instance-based representations during training. Accuracy of both syllable stress production *and* phoneme production were required prior to initiating the old-new judgment task. Participants were allowed to train for up to three hours to meet syllable stress and phonetic accuracy (i.e., accuracy in articulating all nine phonemes within a nonword) during the training portion of the

experiment; however, participants were able to cease training early if the accuracy criterion was met prior to three hours.

Following syllable stress training, participants completed a 10-minute old-new judgment task. During the judgment task, participants were asked to judge the familiarity of different stimuli by pressing a button on a response box. Participants were asked to identify trained stimuli (from the syllable stress training portion of the experiment) as “old” and untrained stimuli as “new.” The untrained stimuli included three different types of stimuli systematically varied in phonetic construction and motor class from the trained stimuli. Figure 13 depicts an example of the experimental procedure for a participant who required two blocks of syllable stress training to meet accuracy criteria. Each portion of this procedure will be described below in Section 7.2.6. Participants were reimbursed $30.00 for their participation regardless of the length of the experimental protocol.

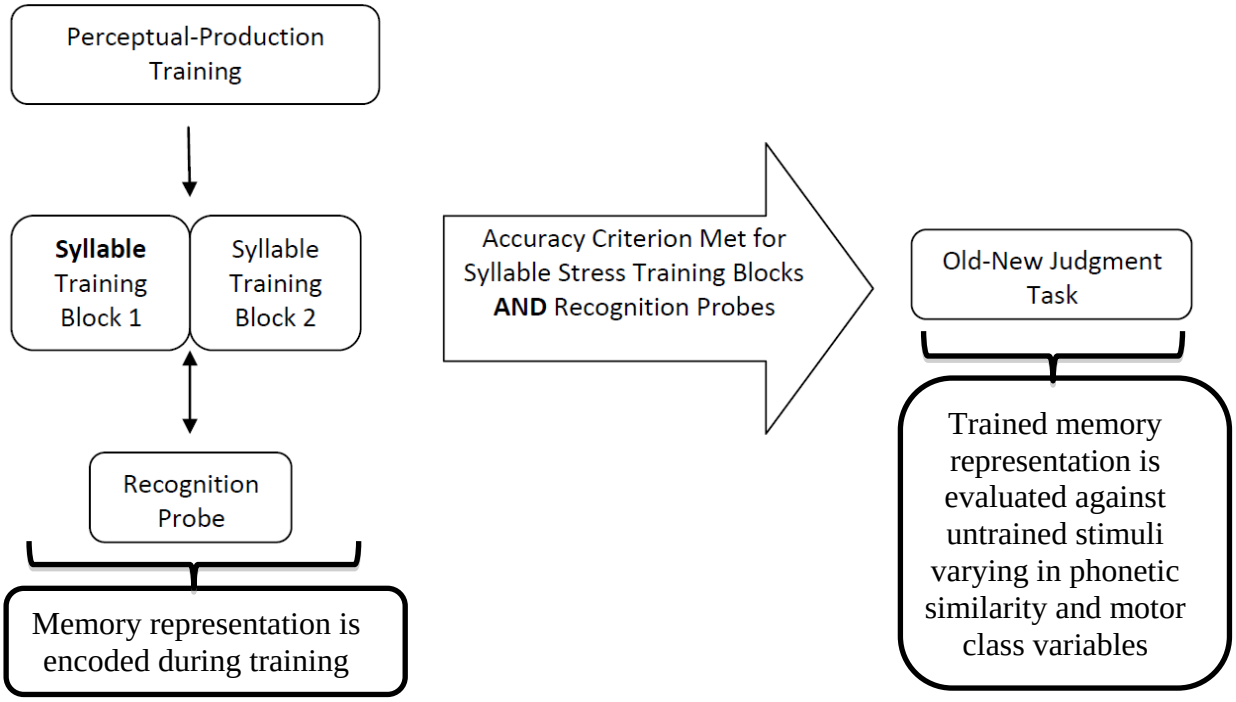

Figure 13: Experimental Protocol

7.2.5.2 Equipment Configuration

During syllable stress training, participants were seated in a comfortable chair with an external speaker (Anchor Audio, Model AN-130) positioned on a table approximately 15 inches to the right of the participant. A computer monitor was positioned to the left of the external speaker, directly in front of the participant. An USB Logitech desktop microphone (Model 980186-0403) was positioned two to three inches from the participant's mouth. The microphone was connected to a Macbook Pro utilizing Windows 7 operating system, which ran the custom training software *Stimulate*. This software and the training protocol will be described in Sections 7.2.6.1-7.2.6.4.

During the judgment task, the microphone and computer monitor were removed, and a serial response box (Psychology Software Tools; Model #200A) was placed directly in front of

the participant's dominant hand. The old-new judgment task was administered by E-Prime (v. 2.0 Professional; Schneider et al., 2002), which collected accuracy and reaction time data input from the response box. Further descriptions about the judgment task can be found in Sections 7.2.6.5-7.2.6.7.

7.2.6 Task Descriptions

7.2.6.1 Syllable Stress Training: General Overview

During perceptual-production and syllable stress training, participants articulated nonword stimuli (taken from Kendall et al., 2005) while trying to produce an increase in intensity level (as measured in decibel level) during the stressed position of the nonword production. The increase in intensity level indicated the stressed syllable in the nonword production. The nonword training stimuli for this experiment were pseudo-randomized in their order of presentation across seven different training blocks. Files containing the order of presentation for each block of nonword stimuli were entered into a custom software program "*Stimulate*" (designed specifically for this study), which controlled the timing of the experiment. *Stimulate* initiated a comprehensive acoustic analysis program, PRAAT (Boersma & Weenink, 2013), to analyze and generate visual feedback of the participant's syllable stress pattern. *Stimulate* only measured intensity level (dB) and did not measure other syllable stress markers (e.g., increases in pitch or duration of the stressed syllable) produced by the participant.

During training, *Stimulate* presented a simultaneous auditory and visual model of a single nonword. The visual model was depicted as three blue vertical bars displayed across the x-axis of the screen, in which each vertical bar represented a single syllable occurring during the articulation of the nonword (e.g., first vertical bar represented the first syllable of the nonword).

The y-axis depicted the relative amount of intensity produced by the participant, in which one vertical bar elevated in relation to the other two vertical bars indicated increased intensity and syllable stress (see Figure 14). The visual representation of the model (the vertical blue bars) was based on the intensity values of the model production. This model was recorded by a trained speaker knowledgeable in phonetic transcription as described in the stimuli validation study (see Section 7.1.6.4). The yellow circles on the display indicated the participant's intended performance for the given model.

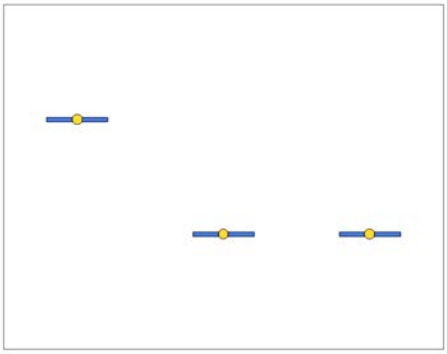

Figure 14: Visual display of a model depicting first syllable stress

Once *Stimulate* presented the auditory/visual model of the targeted nonword, the participant was prompted by *Stimulate* to articulate the stimulus. *Stimulate* recorded the participant's response, and a PRAAT script extracted the maximum intensity values produced by the participant at three pre-specified windows (first, second and third syllable stress positions).[10]

[10] Pre-specified syllable windows were generated for each of the nonword stimuli, and were based on an analysis of the model production's intensity profile. Intensity profiles were reports generated by PRAAT

Stimulate evaluated the extracted intensity data and determined which of the three syllables had the maximum intensity value. If feedback was provided on a given trial, the participant's feedback (the yellow circles) mirrored *Stimulate's* extraction and selection of the maximum intensity value (e.g., a yellow circle displayed relatively higher on the visual display compared to the other two yellow circles). The visual display of the model's production was always present during feedback, and participants are able to determine their accuracy in achieving the correct syllable stress by evaluating their feedback (yellow circles) against the model (blue vertical lines); see Figure 15 for two examples of feedback. During perception-production training feedback was provided 100% of the time (i.e., after every trial). During syllable stress training, feedback was provided 65% of the time. Regardless of feedback schedule, feedback was always displayed for 3000ms.

---

specifying the maximum intensity peaks and valleys across the duration of the nonword. Thus, the nonword model had pre-specified durations that indicated when each syllable unit should begin and end.

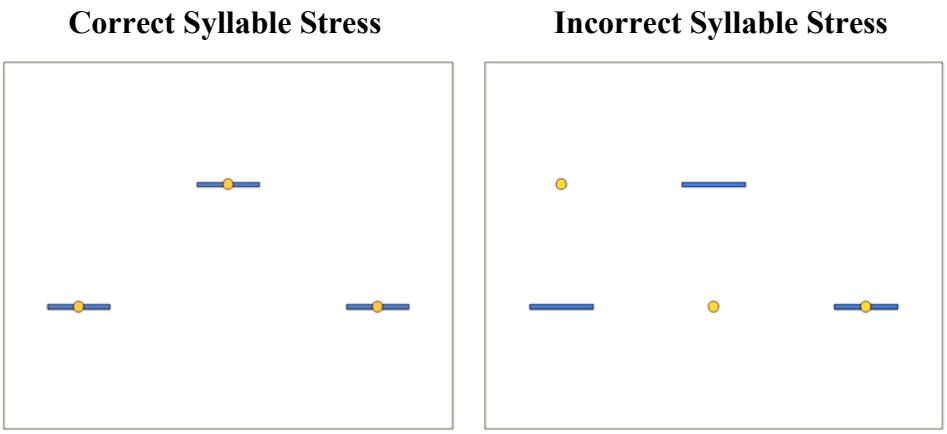

Figure 15: Stimulate feedback displays

7.2.6.2 Training Task: Perception-Production Training

Perception-production training served two purposes in this experiment. First, it allowed participants time to orient to the training procedures and *Stimulate* program. Second, it provided an opportunity for the examiner to provide remediation of participants' misarticulation of any of the nine phonemes comprising each nonword stimulus. This remediation was needed to preserve the phonetic accuracy of the trained stimuli prior to the extensive, repetitive practice incurred during the syllable stress training portion of the experiment. Repetitive practice of any incorrect phonemes during syllable stress training would encode an incorrect memory representation for the nonword stimuli and induce unwanted error into the old-new judgment task later in the experiment.

During perception-production training, participants produced each nonword twice in succession with 100% visual feedback from the *Stimulate* program. There were 30 nonword stimuli trained in this experiment (10 experimental and 20 filler nonwords), and two repetitions of each stimulus allowed the participant to self-correct mistakes involving syllable stress and/or articulation of the nonwords. All feedback about syllable stress was provided by *Stimulate*

following the completion of every trial. Feedback by the examiner about phonetic accuracy was only provided after the perceptual-production task was completed. During this summary feedback, the examiner modeled misarticulated syllables or individual phonemes.

Prior to syllable stress training, participants were required to articulate 95% of the stimuli correctly (57/60 nonwords). This ensured that participants had a correct phonetic baseline representation of the nonwords prior to the extensive practice required during syllable stress training. All accuracy ratings during perception-production training were based on the examiner's perception of the participant's production. The examiner also provided general feedback on syllable stress production (e.g., encouraging an increase in loudness to indicate syllable stress) during this portion of the experiment. All participant questions relating to the experimental procedures, *Stimulate* program, or production of the stimuli were addressed at this stage of the experiment prior to the initiation of syllable stress training.

7.2.6.3 Training Task: Syllable Stress Training

During syllable stress training, participants were provided with extensive practice in achieving the targeted syllable stress patterns for each nonword. This training was procedurally identical to perception-production training except for the number of repetitions for each stimulus item (i.e., training amount) and amount of feedback provided.

Training Amount

Training, or practice, amounts required for completion of syllable stress training were variable across participants for two reasons. First, the number of repetitions or trials required to build a rule- versus instance-based representation is comparable; evidence for extensive training is reported for each type of representation (e.g., Ashby & Gott, 1988; Doody & Zelaznik, 1988; Johansson, 2009; Masson, 1986; Ward & Churchill, 1998). Second, contrasting evidence in speech motor learning suggests varying amounts of practice are required for speech and nonspeech behaviors to be learned to a high level of accuracy (e.g., Fox, Morrison, Ramig, & Sapir, 2002; Maas, Robin, Austermann-Hula, et al., 2008; Meigh & Shaiman, 2010). Thus, participants were given a maximum of 720 repetitions to achieve an accuracy criterion during syllable stress training. These repetitions were split across 6 blocks (120 trials per block) with each stimulus presented four times per block. Participants did not require maximum training amounts to reach accuracy criterion. All but one participant was able to reach criterion following two blocks of training (total of 240 trials; see Table 15).

Table 15: Training blocks required for accuracy criterion to be met

Training Block Number	# of Participants (N = 24)
Training Block 1	11
Training Block 2	12
Training Block 3	1

Feedback

Visual feedback was provided during syllable stress training, and indicated the participant's success in achieving the desired syllable stress pattern for a given nonword. Phonetic accuracy, i.e., accuracy in articulating the correct phonemes for a given nonword, was not provided by *Stimulate* or the examiner during a training block. Instead, phonetic accuracy was assessed using recognition probes following the completion of specific syllable stress training blocks. Rationale for the administration of the recognition probes, as well as their implementation, will be described separately in Section 7.2.6.4.

Participants received visual feedback on their syllable stress productions identical to the feedback provided in the perception-production training portion of the experiment. However, visual feedback was only provided 65% of the time within a single training block (versus 100% during perception-production training). Evidence in motor learning studies suggests extrinsic feedback provided too frequently may decrease long-term learning outcomes (Schmidt & Bjork, 1992; Schmidt & Wrisberg, 2004; Schmidt, Young, Swinnen, & Shapiro, 1989; G. Wulf & Schmidt, 1989). The feedback amount utilized in this portion of the training task had been successful in enhancing learning outcomes with speech-like stimuli similar to those in this experiment (Almelaifi, 2013; Meigh & Shaiman, 2010; Shaiman et al., 2006).

Accuracy Criterion for Syllable Stress

Accuracy in achieving the correct syllable stress pattern for a nonword trial was based on the feedback provided by *Stimulate*. When participant feedback matched the intended model, participants were given credit for the trial. A syllable stress accuracy criterion was set at 90% for all experimental trained stimuli within a single training block (i.e., 36/40 trials had to have the correct syllable stress). Participants who met this criterion were considered to have the

hypothesized rule-based representation for motor class, (i.e., first and second syllable stress patterns) encoded for the trained stimuli, and were eligible to participate in the old-new judgment task.

Two issues emerged during data collection decreasing the validity of the accuracy output originating from *Stimulate*. First, participants were sometimes provided with feedback on the third syllable despite the participant's and examiner's perception of syllable stress earlier in the production. This occurred on trials in which the final consonant of the nonword stimulus was a plosive (e.g., /k/) or fricative (e.g., /s/), and an increase in air volume during the consonant release caused *Stimulate* to judge the third syllable as being stressed. Thus, phonetic structure of some stimuli increased measurement error during training.

Secondly, not all participants utilized intensity as their natural marker for syllable stress. Instructions for syllable stress training, and feedback provided during the perception-production portion of training, emphasized an increase in intensity production as the main marker for syllable stress. Despite these instructions, many participants lapsed into their natural syllable stress patterns near the end of training. Thirteen of the twenty-four participants utilized a natural syllable stress marker other than intensity during training, which resulted in poor accuracy based on *Stimulate* feedback (see Table 16). Despite poor feedback from *Stimulate* during these occurrences, perceptual ratings by the examiner and participant indicated the stress production was accurate using the participant's natural stress marker.

Table 16: Syllable stress makers produced by participants

Syllable Stress Marker	# of Participants (N = 24)
Intensity only	4
Duration only	9
Pitch only	0
Composite: Intensity + Duration	6
Composite: Duration + Pitch	4
Composite: Intensity + Duration + Pitch	1

For these two reasons, a secondary evaluation of syllable stress was adopted during data collection. In addition to noting accuracy of syllable stress production based on *Stimulate* feedback, the examiner also provided a perceptual rating for each trial within a training block. Thus, each training item was given two binary (accurate vs. inaccurate) scores: 1) accuracy based on the examiner's perception of the production and 2) accuracy based on *Stimulate* feedback. The accuracy criterion for syllable stress training was maintained with this combination of scores.

To ensure examiner bias did not influence training outcomes (e.g., stopping training prematurely due to the examiner's perceptual ratings), two undergraduate students were hired to evaluate each participant's recorded productions. Both raters were blinded to the correct syllable stress position for each stimulus item, and were instructed to rate each stimulus for first or second syllable stress. Discrepant ratings between the examiner and *Stimulate* on experimental items in each participant's final training block were evaluated by the blinded raters. Thus, raters re-evaluated up to 40 experimental items per participant to ensure accuracy criterion was actually met with the combined criteria. Reliability ratings between the examiner and the two blinded raters were 100% for 16 of the 24 participants. There was a 2-5% difference in the accuracy

ratings between the examiner and the two blinded raters for 8 of the 24 participants. For these participants, the blinded raters perceived correct syllable stress productions more frequently than the examiner. For example, participant M9 was given an accuracy score of 98% from both blinded raters, but only 93% by the examiner during data collection. Hence, the examiner's score was more likely to elicit additional syllable training than the blinded raters. Additionally, in these cases in which there was a discrepancy in reliability between the examiner and blinded raters, all participants were rated as achieving accuracy criterion. Overall, reliability measurements indicated the accuracy criterion was valid in determining when syllable stress training should cease when both the examiner's perception of syllable stress *and* the intensity feedback output from *Stimulate* were employed.

Accuracy Criterion for Articulation

Measuring phonetic accuracy *during* a speech (or nonspeech) training task is not documented in the speech production literature. Accuracy in articulating distinct and specific phonemes is typically evaluated in speech research as a post-hoc spectrogram analysis. To evaluate phonetic accuracy during training, recognition probes were administered to objectively determine if the phonetic structure of the nonwords was being encoded accurately. This ensured the hypothesized instance-based representation (one phoneme or two adjacent phonemes) was preserved in the trained representation, and could be evaluated accurately against untrained stimuli of varying similarity.

Each recognition probe consisted of 10 trials, in which each trial contained a trained experimental item and two foils. One foil varied the vowel in the stressed syllable. The second foil varied a difficult consonant contrast (e.g., /f/ vs. /θ/ in final position). All foils were recorded and reliability rated for perceptual accuracy as described in Section 7.1.6.4. Trained

experimental stimuli were the exact sound files used in training. All stimuli for the recognition probes were pseudo-randomized in their order within a single trial, as well as across trials in different probes.

During the recognition probe, participants were asked to listen to a single trial at a time, and determine the trained stimulus from the three nonwords. Participants provided the examiner with a verbal response, and were allowed to repeat a given trial an unlimited number of times prior to making a decision. The examiner recorded participants' responses and the number of repetitions required before a response was made.

Accuracy criterion for the recognition probes was 90% (i.e., 9 out of 10 experimental nonwords accurately identified). If a participant required more than one repetition on a given trial to achieve the accuracy criterion an additional recognition probe was administered following further syllable stress training. This additional criterion was instituted following informal observations by the examiner during syllable stress training. Participants who were unable to accurately produce a stress pattern for a nonword trial were also unable to quickly identify the trained item during the recognition task. These participants either required multiple repetitions of the recognition probe trial or were inaccurate in their recognition probe selection. Thus, to be certain the phonetic representation of the nonword was encoded accurately, 90% accuracy performance on the recognition probe *and* no more than one repetition per trial were required for the phonetic accuracy criteria.

A total of four probes was created for this experiment (see Figure 16), with more frequent probing occurring early in training. The actual number of probes required by participants to meet probe accuracy criterion aligned with the number of syllable stress training blocks required to meet syllable stress accuracy criterion. For example, participants who required two blocks of

syllable stress training also required two recognition probes to meet criterion for both portions of training. An old-new judgment task was administered when accuracy criteria were met for syllable stress and phonetic accuracy in the syllable training task.

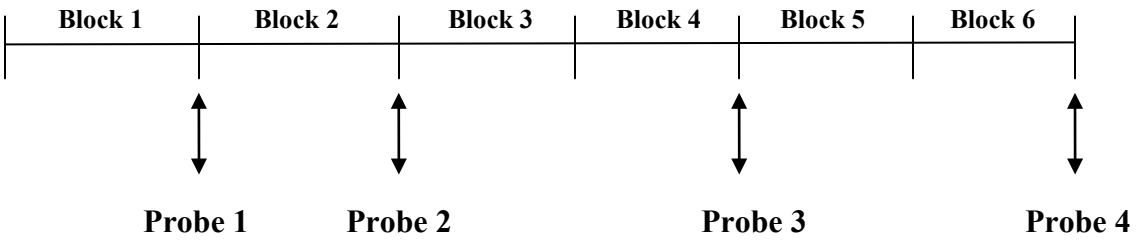

Figure 16: Recognition Probe Schedule

7.2.6.4 Old-New Judgment Task: Rationale

In motor research, evaluation of the varying hierarchical levels of processing has been assessed using reaction time measures (as noted extensively in Chapter 4; for review see Maas & Mailend, 2012). Although there is extensive research evaluating the levels of response programming, there are not reported studies or measures evaluating the response selection stage of the hierarchy. For the current study, evaluation of the response selection stage of processing is imperative given the primary research question of memory representation and learning theory in speech motor control. Thus, a new method was required for evaluating the response selection stage during this experiment.

Old-new judgment tasks are routinely administered in cognitive psychology to probe memory representations for previously encountered stimuli (e.g., Cycowicz, Friedman, Snodgrass, & Duff, 2001; Hall, 1989; Macmillan, 1991; Mulligan, Besken, & Peterson, 2010; Snodgrass & Corwin, 1988; Wallace, 1980). These tasks are highly sensitive to detecting trained stimuli even in the presence of large amounts of untrained stimuli (e.g., Nickerson, 1965;

Shepard, 1967; Standing, 1973). Although this task has not been routinely used in motor control research, it provides a means for probing the response selection level of the motor control hierarchy following syllable stress training. For this experiment, an old-new judgment task was administered to detect trained nonwords from untrained nonwords varying across the two experimental variables: phonetic similarity and motor class.

7.2.6.5 Old-New Judgment Task: General

Once participants met syllable stress accuracy criterion (90% accuracy as determined by *Stimulate* and perceptual syllable stress ratings) *and* phonetic accuracy criterion (90% accuracy on the recognition probes with limited repetitions), the old-new judgment task was administered. Participants were asked to listen to a single nonword and decide if the nonword was familiar. If the nonword was familiar (or "old"), the participant pushed the "old" button on the response box. Participants were expected to push the "old" button for all trained stimuli presented. If the nonword was unfamiliar (or "new"), the participant pushed the "new" button, which was the anticipated response for untrained stimuli.

The following instructions were spoken to the participant: "You will hear sounds similar to those you just practiced. If everything about a sound is exactly the same as one you practiced during your training session push the 'old' button. If anything about a sound is different, then press the 'new' button. Please respond as quickly but as accurately as you can."

All trained and transfer stimuli were pseudo-randomized and presented using E-Prime software (version 2.0 Professional; Schneider et al., 2002). Auditory presentation of each trial through E-Prime consisted of a 250ms long 500 Hz warning tone, 250ms silent pause, auditory presentation of a single stimulus (mean duration: 1293ms), 4,000ms interval of time to capture

the participant response, and a 3,000ms silent interstimulus interval prior to the next trial.[11] E-Prime presented this experimental cycle until all stimuli were responded to by the participant.

All accuracy and reaction time responses were collected using a response box, in which participants pushed a button with their dominant index finger to register a response. At the start of each trial participants placed their index finger in a neutral position (position 3 in Figure 17) between experimental buttons 2 and 4. Buttons 2 and 4 were randomly assigned the positions "old" and "new" for each participant; thus, an equal number of participants responded "old" using button 2 and an equal number responded "old" using button 4. Following a button push, the participant moved his or her index finger back to home base.

[11] Consistent response lag times (i.e., the time between the warning tone and the presentation of the stimulus) are frequently used in old-new judgment tasks (e.g., Mulligan, Besken, & Peterson, 2010; Nosofsky & Stanton, 2006). Comparisons of consistent versus varied response lags in these types of tasks suggests no difference in speed-accuracy judgments or over-anticipation during the response (Miller, Sproesser, & Ulrich, 2008).

Figure 17: Response box display

7.2.6.6 Old-New Judgment Task: Instructions

The terminology in the instructions was meant to be neutral, i.e., to signify the stimuli the participant should listen for without inducing any unwanted bias toward either learning theory. Empirical research in cognitive psychology suggests instructions may influence learned memory representations. Informing participants of underlying rules can bias participants to report “rules” in their description of a task when no rules are present in the experiment (e.g., McAndrews & Moscovitch, 1985; Reber, Kassin, Lewis, & Cantor, 1980; Werker & Tees, 1984). Instructions to memorize specific exemplars versus abstract rules provides some evidence of instance-based learning (e.g., Reber & Allen, 1978; Vokey & Brooks, 1992, n. experiment 2 and 3); however, variations on how to bias instructions toward instance-based learning (e.g., looking for similarity versus memorization) are lacking in the literature. Additionally, direct manipulation of instructions to induce a bias for either rule- or instance-based learning in certain cognitive

psychology tasks does not always bias learning (e.g., Dienes, Broadbent, & Berry, 1991; Mathews et al., 1989; Vokey & Brooks, 1992, n. Experiment 1).

Empirical research in auditory perception is consistent with these findings in cognitive psychology, in which instruction can bias perception of nonnative language contrasts in listeners (e.g., Bowers, Mattys, & Gage, 2009; Hyltenstam, Bylund, Abrahamsson, & Park, 2009; Oh, Au, & Jun, 2010; Singh, Liederman, Mierzejewski, & Barnes, 2011; Werker & Tees, 1984b). Werker & Tees (1984b) evaluated adult participants' abilities to perceive and process nonnative language contrasts based on different instruction sets. One set of instructions applied the terminology "speech perception" and "syllables," which was predicted to keep participants from segmenting nonnative contrasts outside their native language rules (a rule-based learning outcome). The other set of instructions applied the terminology "sound discrimination," and instructed participants to listen for "drops of water falling in a bucket." These instructions were hypothesized to promote segmented processing of the stimuli (an instance-based learning outcome). Results were consistent with both hypotheses.

The instructions for the old-new judgment task used the descriptor "sound," which does not does not imply a speech rule-based schema (e.g., syllables), but does inform participants of the type of stimuli to expect in the task. This descriptor was more neutral than "drops of water falling in a bucket," and should not have induced unwanted phonetic segmentation of the acoustic signal. These instructions were also constructed to be similar in their neutrality to other experiments employing old-new recognition tasks (e.g., Cycowicz et al., 2001; Khoe, Kroll, Yonelinas, Dobbins, & Knight, 2000; Mulligan et al., 2010).

7.2.7 Stimuli

Trained and untrained stimuli were implemented for this experiment. Trained stimuli were practiced in the syllable stress training task. Untrained stimuli were encountered *only* in the old-new judgment task. Both types of stimuli were constructed with novel phonetic combinations and specific syllable stress patterns. These two components allowed for hypothesized memory representations from both learning theories to be instantiated in the stimuli. For this study, instance-based learning memory representations were hypothesized to be one phoneme or two adjacent phonemes in a CV unit, which is analogous to the representational units proposed by Crump and Logan (2010). As noted in Chapter 5, the phoneme unit has also been contrasted to rule-based syllable units. Rule-based learning memory representations were hypothesized to be stressed syllables in the first and second position of the nonword. Syllable stress was considered a motor program, which directed transfer effects to two specific syllable stress positions (first and second position in three syllable nonwords). As both syllable positions have a high-frequency occurrence in the English, first and second syllable stress positions were considered to be within the same motor class.

7.2.7.1 Trained Stimuli: Experimental

Trained experimental stimuli were taken from a portion of Kendall et al.'s (2005) moderate and low frequency stimuli (see Table 17).

Table 17: Trained stimuli

Syllable Stress Position	Nonwords	Kendall Frequency Group
1	kæθotæs	Moderate
1	sæθodæk	Moderate
1	zotenav	Moderate
1	θʌrasæθ	Moderate
2	zaʃɔʤəz	Low
2	zæʃɔʤəθ	Low
2	ʤəzɔzæk	Low
2	zænɔʤəθ	Low
2	ʤʌnɔzæk	Low
2	tenærok	Moderate

Syllable Stress Assignment

As noted in Chapter 5, high-frequency syllables have been postulated as different motor programs from low-frequency syllables (Aichert & Ziegler, 2004; Cholin et al., 2006; Cholin & Levelt, 2009; Laganaro, 2005, 2008; Staiger & Ziegler, 2008). In English, first and second syllable stress positions are the most frequently produced stressed positions in three syllable words (Clopper, 2002). Syllable stress was systematically assigned to either the first or second syllable in each trained nonword. Training these two high-frequency positions provided a within-class variable as the motor program for syllable stress was hypothesized to be shared for these two high-frequency stress positions. Final syllable stress is uncommon in three-syllable words in English, and this syllable position was hypothesized to have a different motor program directing stress production. Thus, untrained stimuli with third syllable stress provided an outside-class variable for motor program, as the motor program learned during training (i.e., first and second syllable stress) did not aid transfer.

To enhance learning of the stressed syllable, consistent mapping between syllable stress and nonword syllables was maintained within the training list (Hall, 1989; Schneider & Shiffrin,

1977; Shiffrin & Schneider, 1977). For example, the syllable /te/ was always unstressed regardless of its syllable position in a nonword. This constraint made equal distribution of syllable stress positions across stimuli unfeasible as noted in Table 17.

7.2.7.2 Trained Stimuli: Filler

Trained filler stimuli controlled for response bias in the old-new judgment task. Twenty additional stimuli were trained with the experimental training set, ensuring an even number of anticipated "yes" and "no" responses from the participant during the judgment task. All filler stimuli were three syllable nonwords following the same CVCVCVC pattern as the training stimuli. Filler stimuli were adapted from three sources to meet this CV syllable frame: Kendall et al.'s (2005) high frequency stimuli; Roy and Chiat's (2004) three-syllable stimuli; and Dollaghan and Campbell's (1998) three-syllable stimuli. Syllable stress was randomly assigned to either the first or second syllable, and this syllable stress proportion was balanced across all filler stimuli (noted in bold font within Table 18).

Table 18: Filler stimuli

Stimuli	Source
zesəfis	Kendall et al. (2005)
ze**sə**fin	Kendall et al. (2005)
vɔsəfis	Kendall et al. (2005)
zi**rə**fin	Kendall et al. (2005)
disəfis	Kendall et al. (2005)
nʌ**zi**rəz	Kendall et al. (2005)
nʌzirəv	Kendall et al. (2005)
rʌ**zi**rəs	Kendall et al. (2005)
sʌzirəʤ	Kendall et al. (2005)
pʌ**zi**rəʤ	Kendall et al. (2005)

Stimuli	Source
sɪnədɔb	Roy & Chiat (2004)
lɑ**dɪ**heb	Roy & Chiat (2004)
nənɑbəp	Roy & Chiat (2004)
tɑ**fu**pəl	Roy & Chiat (2004)
gæzəmin	Roy & Chiat (2004)
rɪ**gə**sɛt	Roy & Chiat (2004)
tʃinɪtab	Dollaghan & Campbell (2008)
nɪ**tʃo**veb	Dollaghan & Campbell (2008)
dɪtəvæb	Dollaghan & Campbell (2008)
te**vɪ**tʃɪg	Dollaghan & Campbell (2008)

7.2.7.3 Untrained Stimuli: General Overview

The untrained stimuli were designed with a similarity index much like Crump and Logan's (2010) stimuli, in which systematic variations in similarity existed between the trained and untrained typed words. For the current study, untrained stimuli varied in similarity to the trained stimuli across two variables: phonemes and motor class. See Figure 18 for the organization of the untrained stimuli across these variables). Chamberlin and Magill (1992b) and Crump and Logan (2010) did not include a motor class variable in their stimuli design. However, rule-based learning predicts motor class effects (see Chapter 2); thus, DPDC stimuli were created to evaluate the variable of motor class. It should be noted Same-Phoneme-Different-Motor-Class (SPDC) stimuli were not included in this study because the design of the stimuli would have led to variable mapping conditions between stress and syllable. This variable mapping condition may have decreased learning outcomes during syllable stress training, and potentially biased the recognition test results.

Untrained Stimuli

	Phonemes in CV Unit	Same Motor Class (syllable stress)
SPSC	Yes	Yes
DPSC	No	Yes
DPDC	No	No

S = Same
D = Different
P = Phonemes in CV Unit
C = Motor Class

Figure 18: Legend for untrained stimuli construction

7.2.7.4 Untrained Stimuli: SPSC Stimuli

The Same-Phoneme-Same-Motor-Class (SPSC) stimuli were exactly the same as the trained experimental stimuli except the first and second syllables were exchanged, e.g., the trained stimulus "kæθotæs" was changed to "θokætæs" (see Table 19 for all other SPSC stimuli). The experimental variable of motor class was not manipulated for these untrained stimuli; both SPSC and trained stimuli shared the same motor class (first and second syllable stress positions).

Table 19: SPSC Stimuli

Syllable Stress Position	SPSC
2	θokætæs
2	θosædæk
2	tezonav
2	raθʌsæθ
1	ʃɔzadʒəz
1	ʃɔzædʒəθ
1	zɔdʒəzæk
1	nɔzædʒəθ
1	nɔdʒʌzæk
1	næterok

7.2.7.5 Untrained Stimuli: DPSC Stimuli

The Different-Phoneme-Same-Motor-Class (DPSC) stimuli were Kendall et al. (2005) moderate and low frequency stimuli not used for training (see Table 20).

Table 20: DPSC Stimuli

Syllable Stress Position	DPSC	Kendall Frequency Group
1	ʃɔdʒəzɔd	Low
1	vuzæʃɔm	Low
1	fozæʃɔd	Low
1	kozæʃɔm	Low
1	dɔdʒəzɔd	Low
1	næθodæp	Moderate
2	rasæθon	Moderate
2	sʌvenæθ	Moderate
2	nasæθoʃ	Moderate
2	viʃədæk	Moderate

These stimuli share similar, but not identical, phonemes to the trained stimuli (i.e., the phonemes were the same but not in the same CV order). Both DPSC and trained stimuli share the same motor class (first and second syllable stress positions). Consistent mapping between syllable stress and nonword syllables was maintained within this stimuli list and across other stimuli lists (trained and SPSC lists). This mapping constrained the syllable stress positions allowed for this stimuli set; see Table 21 for syllable stress proportions for the trained and DPSC stimuli sets.

Table 21: Syllable stress proportions of Trained and DPSC stimuli

	Training Set	DPSC
First Syllable Stressed	4/10 nonwords (40%)	6/10 nonwords (60%)
Second Syllable Stressed	6/10 nonwords (60%)	4/10 nonwords (40%)

7.2.7.6 Untrained Stimuli: DPDC Stimuli

The DPDC stimuli were created to be as different as possible from the trained stimuli across the two experimental variables. Phoneme similarity was minimized as much as possible while still using English phonemes. Although phonemes outside the English language may have been more dissimilar to the trained stimuli, the application of these phonemes may have altered the way participants reacted to them in the old-new judgment task. Specifically, an increase in reaction time with novel phonemes outside an individual's native language has been reported in perceptual studies, and is attributed to participants re-mapping phonemes onto a similar equivalent in their native language (Dickey, M., personal communication, May 15, 2013; also: Samuel, 1982; Savela et al., 2003). This re-mapping may lead to increased variability in individual decoding skills with novel phonemes (e.g., Best, Morrongiello, & Robson, 1981; Eisner & McQueen, 2005; Repp, 1982) and bias towards rule-based learning in which participants rely on linguistic rules in their native language to aid decoding (M. Dickey, personal communication, May 15, 2013; also: Kuhl, 1991, 2000; Samuel, 1982).

The DPDC stimuli were constructed with English phonemes not present in the Kendall et al. (2005) stimuli. Implementation of Kendall et al.'s methods was attempted during construction of the DPDC (including controlling for low frequency occurrence of interphonemic transitional gestures). Many of the phonemes not found in Kendall et al.'s (2005) study were also not listed in Roberts' (1965) phonemic transitional probability data or occurred with more than 1% relative frequency (Shriberg & Kent, 1982). These were noted as "exceptions" in Table 22, and these phonetic contexts were avoided during stimuli construction. All constructed nonwords were reviewed by a highly trained linguist for any deviations from English phonotactic constraints (see full list of stimuli in Table 23).

Table 22: Intergestural phonemic constraints for DPDC stimuli

Phoneme	What can follow it	Exceptions
b	ɪ ʊ	
g	ɪ ʊ i	
tʃ	ɪ ʊ	e u i
ʒ	e u ɪ ʊ	i
ɪ	ð w b h g j tʃ	ŋ ʒ
i	w b g j	ʒ tʃ
e	ð b g j	ŋ w ʒ tʃ
ʊ	ð w b h g j tʃ	ŋ ʒ
ð	ɪ ʊ	u
j	ɪ ʊ u i	
h	ɪ ʊ	
ŋ	ʊ ɪ i	e u
u	b g j	ŋ ʒ tʃ
w	ɪ ʊ	

Table 23: DPDC Stimuli

Syllable Stress Position	DPDC
3	ʒibʊtʃeð
3	tʃeðugʊʒ
3	ʒʊgijub
3	gʊgɪðʊtʃ
3	bɪðetʃug
3	gigʊðib
3	tʃejiwɪʒ
3	bʊtʃitʃeʒ
3	tʃʊtʃubɪʒ
3	gibɪðɪb

Syllable stress was assigned to the third syllable for each DPDC nonword, which varied from the trained motor class (first and second syllable stress). Third syllable stress position in three-syllable words was the least frequent position for stress in English (Clopper, 2002), which is hypothesized to be governed by a different motor program from high-frequency syllable positions (Aichert & Ziegler, 2004; Cholin et al., 2006; Cholin & Levelt, 2009; Laganaro, 2005,

2008; Staiger & Ziegler, 2008). Additionally, the decreased frequency of occurrence reduced the likelihood participants would rely on previously encoded representations for this stress pattern to aid their judgments during the old-new judgment task. Furthermore, the recent exposure to first and second syllable stress training prior to the judgment task should also exaggerate the differences noted between the different syllable stress positions.

7.2.8 Stimuli Preparation

In the stimuli validation study, Kendall et al.'s (2005) stimuli were assigned first syllable stress, and were validated for phonetic accuracy. These same sound files were administered for the main experiment if trained stimuli were assigned first syllable stress. All of the other stimuli were re-recorded with their assigned syllable stress (either second or third stress positions) using methods outlined in Section 7.1.6.4. Additionally, these stimuli were rigorously rated for phonetic accuracy as outlined in Section 7.1.6.4 to ensure participants perceived the nonwords correctly during the training and old-new judgment tasks.

Perceptual analysis of syllable stress was conducted for all stimuli to ensure participants perceived the correct syllable stress for each stimulus set. The same transcribers who participated in the phonetic accuracy ratings also rated syllable stress of the nonwords in a separate task. Raters listened to each nonword, and rated which syllable was perceived as the stressed syllable (first, second, or third). Agreement between two out of three transcribers was required for a stimulus to be included in the main experiment. If a stimulus did not meet criterion it was re-recorded and further perceptual assessments were conducted until criterion was met.

8.0 RESULTS

The aim of the present investigation was to evaluate learning parameters associated with rule- and instance-based learning theories following the acquisition of nonwords with varying syllable stress patterns. A stimuli validation study was conducted to validate the stimuli for the main experiment. The main experiment evaluated the variables of similarity and motor class within trained and untrained stimuli to investigate transfer patterns associated with each learning theory.

8.1 STIMULI VALIDATION

Kendall et al.'s (2005) study was replicated to determine if the original effects related to interphonemic transitional gestural frequency were maintained in a young adult population (Research Question #1). This replication included evaluating real word frequency in young adults as a control variable against unusual nonword findings (e.g., error measurement; Research Question #2). Additionally, the DPDC stimuli were evaluated against Kendall et al.'s nonword stimuli, which yielded a baseline measurement of how the DPDC stimuli were processed in relation to the Kendall et al. stimuli for the main experiment (Research Question #3).

8.1.1 Data Preparation

Prior to data analyses, data was excluded for trials produced incorrectly, i.e., misarticulated, and for reaction times slower than 50ms or less. These exclusions are reported separately below.

8.1.1.1 Accuracy Analyses

Accuracy in articulating the stimuli (real and nonword) were perceptually judged by the examiner during data collection and given a binary score (accurate or inaccurate). Participants misarticulated 1% of the real word stimuli (10 errors out of a possible 920 real word trials), and these errors were equally distributed between high and low frequency words (see Table 24). Participants misarticulated 36% of the nonword stimuli (327 errors out of a possible 920 nonword trials), and the majority of these errors were produced during articulation of the DPDC stimuli (see Table 25).

Table 24: Error analysis for real words (across participants)

Real Words	Number of Errors
High Frequency	5
Low Frequency	5

Table 25: Error analysis for nonwords (across participants)

Nonwords	Number of Errors
High Gestural Frequency	43
Moderate Gestural Frequency	65
Low Gestural Frequency	81
DPDC	138

Reliability of the examiner's perceptual judgments was conducted on a sample of 40% of the data (i.e., 9 of the 23 participants' total data). An individual trained in phonetic transcription listened to each of the nine experimental sessions (total of 164 trials), and rated each trial as correct or incorrect. Inter-rater reliability between the examiner and the independent judge was 86% (141/164 trials rated the same).

8.1.1.2 Reaction Time Data

Reaction times were excluded from the main analyses if a trial resulted in a "no response" or breath trigger. No response trials were trials produced by the participant prior to E-prime initiating the response collection command. No response trials occurred when participants initiated their articulation while the sound file was being played through the speaker. These responses were registered as 0ms by E-Prime. Breath triggers occurred when the participant exhaled prior to the onset of articulation, and this exhalation triggered the voice key in the response box. These responses were associated with reaction times of 50ms or less. Total number (and percentage) of no responses and breath triggers for real word (Table 26) and nonword (Table 27) stimuli are presented below.

Table 26: Total number and percentage of No Responses and Breath Triggers for real word stimuli

	High Frequency Real Words	**Low Frequency Real Words**
Total No Responses (% out of total 920 trials)	17 (1.85%)	15 (1.63%)
Total Breath Triggers (% out of total 920 trials)	6 (.65%)	12 (1.3%)

Table 27: Total number and percentage of No Responses and Breath Triggers for nonword stimuli

	High Gestural Frequency	Moderate Gestural Frequency	Low Gestural Frequency	DPDC
Total No Responses (% out of total 920 trials)	4 (.43%)	1 (.11%)	1 (.11%)	4 (.43%)
Total Breath Triggers (% out of total 920 trials)	3 (.33%)	3 (.33%)	3 (.33%)	2 (.22%)

8.1.2 PRIMARY OUTCOMES

8.1.2.1 Research Question 1 (RQ1): Nonword Frequency

RQ1 evaluated whether Kendall et al.'s (2005) study outcomes were replicated in a younger participant sample. A one-way within-subjects analysis of variance (ANOVA) was performed on reaction times as a function of nonword interphonemic transitional gestural frequency. Participants were measured on three frequency categories (high, moderate, and low frequencies). The assumption of sphericity was met, Mauchly's $W = .995$, $\chi^2(2) = .097$, $p = .953$. The assumption of normality was met (see Table 28). All other assumptions were met.

Table 28: Test of normality of nonword gestural frequency for each frequency category

Frequency Category	Shapiro-Wilk *W*	*df*	*p*
High Frequency	.974	23	.781
Moderate Frequency	.973	23	.758
Low Frequency	.939	23	.172

There was a significant difference in reaction times based on nonword frequency category, $F(2,44) = 3.892$, $p = .028$, $\eta^2 = .150$. In order to find the pattern of differences in reaction time based on nonword frequency category, post-hoc comparisons were performed using the Bonferroni adjustment. Participants' reaction times were significantly faster when articulating high gestural frequency nonwords compared to low gestural frequency nonwords, $p = .030$ (see Figure 19). There were no other significant differences, $ps > .468$. The means and standard deviations of the nonword gestural frequency categories are reported in Table 29.

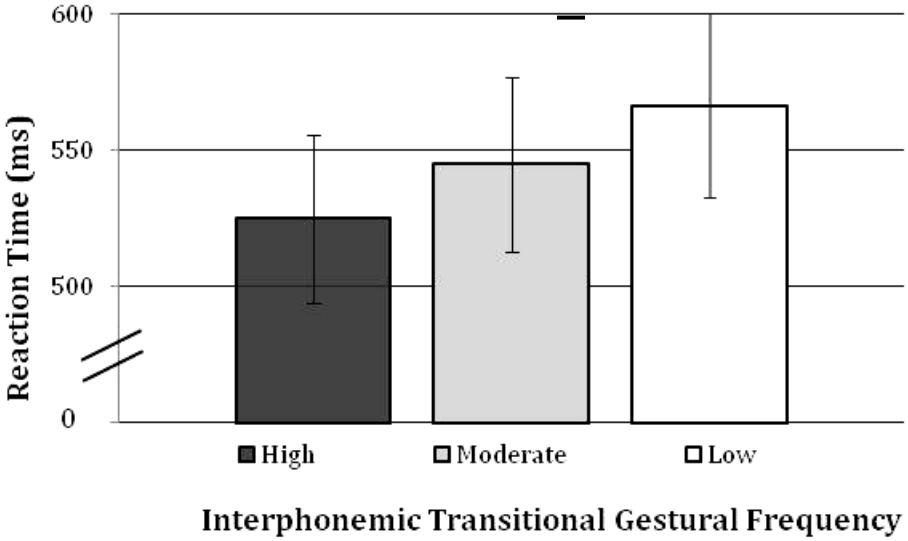

Figure 19: Mean reaction times (ms) and standard errors of interphonemic transitional gestural frequency categories.

Table 29: Mean and SD for nonword gestural frequency categories

Frequency Category	Mean (ms)	SD (ms)
High Frequency	524.82	148.84
Moderate Frequency	545.00	153.36
Low Frequency	566.41	160.87

8.1.2.2 Research Question 2 (RQ2): Real Word Frequency

RQ2 evaluated the reaction times across lexical frequencies of the real word stimuli, which yielded a control measure for evaluating the nonword results examined in RQ1. A paired-samples t-test was conducted to compare reaction times in high and low lexical frequency categories in real words. Participants reacted significantly faster to high frequency words (M = 498.08ms, SD = 139.40ms) compared to low frequency words (M = 525.06ms, SD = 135.58ms); $t(22) = -3.031$, $p = .006$ (see Figure 20).

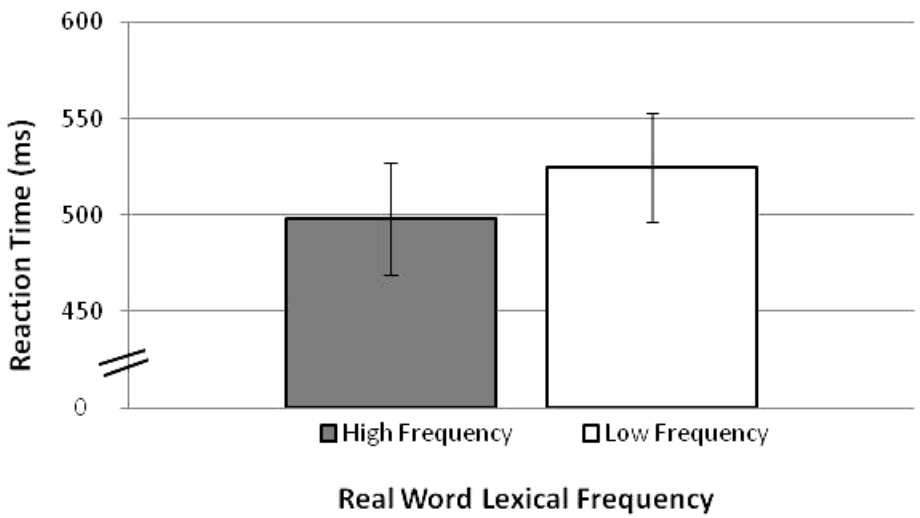

Figure 20: Mean reaction times (ms) and standard errors of real word stimuli across frequency categories

8.1.2.3 Research Question 3 (RQ3): Stimuli Type

RQ3 evaluated whether the DPDC stimuli were similar to the Kendall et al. (2005) stimuli based on reaction time. A paired-samples t-test was conducted to compare reaction times between the Kendall et al. (2005) stimuli (averaged together across gestural frequency category) and the

DPDC stimuli. Participants reacted significantly faster to the Kendall et al. stimuli (M = 545.41ms, SD = 148.82ms) compared to the DPDC stimuli (M = 602.56ms, SD = 160.81ms); $t(22) = 4.332$, $p \leq .000$ (see Figure 21).

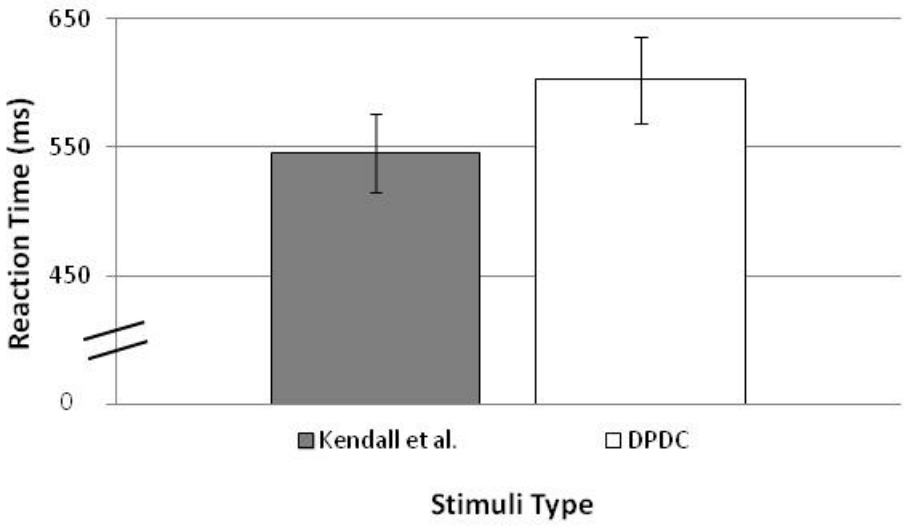

Figure 21: Mean reaction times (ms) and standard error of nonword stimuli type

8.2 RULE- VERSUS INSTANCE-BASED LEARNING

The main experiment evaluated two learning parameters associated with rule- and instance-based learning: phonetic similarity (Research Question #4) and motor class (Research Question #5). Each of these variables aligns with specific transfer predictions for each theory (see Table 30).

Table 30: Transfer variables associated with each learning theory

	Rule-based Learning	Instance-based Learning
SIMILARITY involved in transfer	No	Yes
MOTOR CLASS involved in transfer	Yes	No

8.2.1 Data Preparation

All accuracy and reaction times reported were recorded during the old-new judgment task. Results from the training portion of the experiment were reported in Chapter 7 as the sole purpose of training was to encode the trained stimuli to a high level of accuracy in memory. Only the experimental stimuli were evaluated in the main analyses; trained filler stimuli were not included. Prior to the main analyses, reaction time data were excluded for no responses. No response trials were produced by the participant prior to E-prime initiating the response collection command, and were registered as 0ms by E-Prime. Total number and percentage of no response trials across stimuli type are included in Table 31. Reactions times greater than 3 SD from the median for a given trial (1.38% of all data) were also excluded from the main analyses.

Table 31: Total number of No Responses and Outliers greater than >3SD from the median

	Trained	SPSC	DPSC	DPDC
Total No Responses (% out of total 960 trials)	8 (.83%)	2 (.21%)	4 (.42%)	8 (.83%)
Total Outliers (>3SD)	4	6	7	3

8.2.1.1 Accuracy Data: Overview of *d'* Statistic Calculations

A *d'* statistic was calculated for each participant based on their judgment of trained and untrained nonwords as old (i.e., familiar) or new (i.e., unfamiliar) using the rubric in Table 32. This statistic is common in standard "old-new" judgment experiments as an assessment of discrimination accuracy (Macmillan, 1991; Snodgrass & Corwin, 1988). Calculation of the *d'* statistic was a difference score between positive "hits" (*H)* and "false alarms" (*FA*) for each item (Corwin, 1994; Macmillan, 1991; Snodgrass & Corwin, 1988): $d' = z(H) - z(F)$, where H = Proportion ("old" responses/# of OLD stimuli) and F = Proportion ("old" responses/# of NEW stimuli).

Table 32: Anticipated and summarized data to calculate *d'* statistic

	Participant Response "OLD"	**Participant Response "NEW"**
Training Stimuli ("Old") Experimental (n=10) Filler (n=20)	Hit *(H)*	Miss
Transfer/Untrained Stimuli ("New") SPSC (n=10) DPSC (n=10) DPDC (n=10)	False Alarm *(F)*	Correct Rejection

Participants who perfectly discriminate between stimuli have *d'* scores of 4.65 (H = .99, F = .01); those who cannot discriminate accurately have *d'* scores of 0 (or an equal chance of hits and false alarms, H = F; Macmillan, 1991). Participants in this study were anticipated to have *d'* statistics ranging from 1.40 – 2.27 (see Appendix E for a description on how this range was calculated). Participants with total false alarm scores (as averaged across all trials) greater

than 50% should be excluded from all analyses; participants with high false alarm scores no longer rely on the stimuli for discrimination (Macmillan, 1991; evidence consistent with this cut-off: Khoe et al., 2000; Cycowicz et al., 2001).

A bias score was calculated across participants to evaluate participants' difficulty in classifying the SPSC stimuli (Corwin, 1994; Snodgrass & Corwin, 1988). These stimuli have syllable portions similar to the training stimuli, which may bias participants to respond "old" even though the nonword items have never been heard prior to the judgment task. According to the two-high threshold account of discrimination (Snodgrass & Corwin, 1988), there is a continuum of uncertainty and a given threshold can be calculated for when a participant will respond "old" versus "new" in uncertain conditions (Corwin, 1994; Snodgrass & Corwin, 1988). This can be calculated as $Br = F/(1-Pr)$, where $Pr = H - F$ (Corwin, 1994; Cycowicz et al., 2001; Snodgrass & Corwin, 1988). *Br* values under .50 are considered conservative in their estimation of participants providing false alarms (Corwin, 1994; Cycowicz et al., 2001; Snodgrass & Corwin, 1988).

8.2.1.2 Accuracy: *d'* Statistic and Error Analyses

All false alarm rates were under 14%, and the range of *d'* statistics was 2.56 – 4.26 (M = 3.28, SD = .44); see Table 33. This range is above the conservative estimates of anticipated *d'* statistics predicted for stimuli in this study (noted in Appendix E). The low false alarm rates and

d' statistics provided support for inclusion of all subject data in the main analyses that met the above reaction time criterion noted in the data preparation section above (Section 8.2.1).[12]

Table 33: *H, F, d'* and *Br* statistics for all participants

Participant	Hit Rate (*H*)	False Alarm (*F*)	False Alarm (*F*) Percentage	*d'*	*Br*
1	0.93	0.03	3.30	3.34	0.33
2	0.97	0.03	3.30	3.67	0.50
3	0.93	0.13	13.30	2.61	0.67
4	0.93	0.07	6.70	3.00	0.50
5	0.93	0.02	1.70	3.63	0.20
6	0.97	0.10	10.00	3.12	0.75
7	0.98	0.02	1.70	4.26	0.50
8	0.97	0.03	3.30	3.67	0.50
9	0.93	0.02	1.70	3.63	0.20
10	0.80	0.02	1.70	2.97	0.08
11	0.93	0.10	10.00	2.78	0.60
12	0.90	0.10	10.00	2.56	0.50
13	0.93	0.10	10.00	2.78	0.60
14	0.93	0.02	1.70	3.63	0.20
15	0.97	0.10	10.00	3.12	0.75
16	0.93	0.03	3.00	3.34	0.31
17	0.93	0.03	3.00	3.34	0.31
18	0.93	0.13	13.30	2.61	0.67
19	0.97	0.02	1.70	3.96	0.34
20	0.90	0.02	1.70	3.41	0.15
21	0.87	0.02	1.70	3.24	0.11
22	0.93	0.02	1.70	3.63	0.20
23	0.93	0.10	10.00	2.78	0.60
24	0.98	0.07	6.70	3.63	0.80

[12] Since the initial review and defense of this dissertation, a post-hoc analysis of the primary outcomes was re-conducted with responses times for inaccurate judgments removed. Results for this new analysis were the same as those reported in this chapter. Full detail of this post-hoc analysis is provided in Appendix F.

Participants responded incorrectly on 5.76% of the stimuli (83 errors out of a possible 1440 trials) with the majority of the judgment errors occurring on the experimental filler items (see Table 34).

Table 34: Error analysis by stimuli type (across participants)

Type of Stimuli	Number of Errors
Training: Experimental	10
Training: Filler	38
SPSC	30
DPSC	5
DPDC	0

The *Br* statistic yielded an estimate for participants' responses to the SPSC stimuli. Participants with low *Br* statistics had lower *F* (false alarm) rates and errors in judging the SPSC stimuli as "new" during this task (see Table 35).

Table 35: *Br* statistics in relation to *F* rates and accuracy scores on SPSC stimuli

	# of Participants	Total Incorrect SPSC	*Br* Mean (SD)	*Br* Range	*F* % Rate Mean (SD)
Small *Br* (< .50) *Bias to rate SPSC as "new" (correct rating)*	11	3	.22 (.09)	.08-.34	2.08% (.66)
Moderate *Br* (= .50) *Neutral bias towards SPSC judgments (either "old" or "new")*	5	7	.50 (0.00)	.50-.50	5% (3.34)
Large *Br (> .50)* *Bias to rate SPSC as "old" (incorrect rating)*	8	20	.68 (.08)	.60-.80	10% (.02)

8.2.2 PRIMARY OUTCOMES

The main analysis evaluated the effects of phonetic similarity and motor class between the trained and untrained stimuli. A one-way within-subjects ANOVA was performed on reaction times as a function of stimuli type. Participants were measured on four types of stimuli (trained, SPSC, DPSC, and DPDC stimuli). The assumption of sphericity was met, Mauchly's $W = .613$, $\chi^2(5) = 10.638$, $p = .059$; however, as the p-value was just above non-significance, the Huyn-Feldt correction is also reported.[13] The assumption of normality was met for all levels of stimuli type except the DPSC stimuli (see Table 36). All other assumptions were met.

Table 36: Test of normality across stimuli type

Stimuli Type	Shapiro-Wilk *W*	*df*	*p*
Trained	.960	24	.431
SPSC	.924	24	.072
DPSC	.904	24	.026*
DPDC	.923	24	.068

Note: 3 * significant at *p* = .05

There was a significant difference in reaction times depending on stimuli type, $F(3,69) = 18.747$, $p \leq .000$, $\eta^2 = .449$. This effect was maintained using the Hyun-Feldt correction: $F(2.582,59.390) = 18.747$ $p \leq .000$, $\eta^2 = .449$. In order to find the pattern of differences in reaction time based on transfer pattern and motor class, post-hoc comparisons were performed

[13] Violations of Mauchly's Test of Sphericity indicate an increase in Type I error may occur. The Huyn-Feldt correction modifies the degrees of freedom to provide a more valid *F*-ratio, in which the risk of a Type I error is reduced (Keselman, Algina, & Kowalchuk, 2001).

using the Bonferroni adjustment. These are described under each corresponding research question.

8.2.2.1 Research Question 4 (RQ4): Phonetic Similarity

RQ4 evaluated the effect of phonetic similarity between the trained and untrained stimuli varying by phonetic similarity only (SPSC and DPSC stimuli). The parameter of motor class was the same for these three stimuli types. Participants' reaction times were significantly slower in responding to the SPSC nonwords compared to the trained ($p = .013$) and DPSC nonwords ($p \leq .000$); see Figure 22. There were no other significant differences *ps* > .99. The means and standard deviations of the stimuli categories are reported in Table 37.

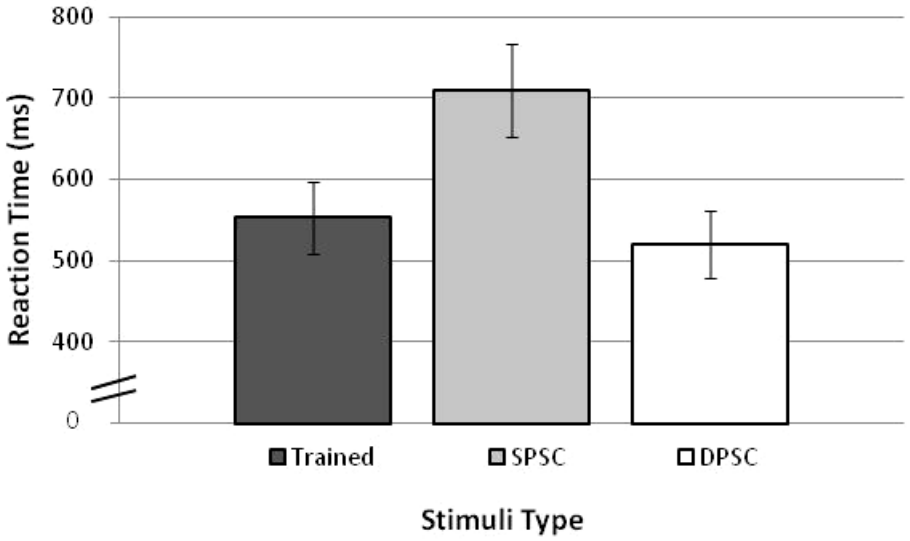

Figure 22: Mean reaction times (ms) and standard errors of stimuli varying by phonetic similarity

Table 37: Mean and SD for stimuli with varying phonetic similarity

Stimuli Type	Mean (ms)	SD (ms)
Trained	552.63	215.26
SPSC	708.66	279.34
DPSC	532.80	200.15

8.2.2.2 Research Question 5 (RQ5): Motor Class

RQ5 evaluated the effect of motor class between the DPSC and DPDC stimuli, in which the parameter of phonetic similarity was held constant. Participants' reaction times were significantly slower when responding to the DPSC nonwords (M = 532.80ms, SD = 200.15ms) compared to the DPDC nonwords (M = 452.08ms, SD = 166.50ms), p = .028; see Figure 23.

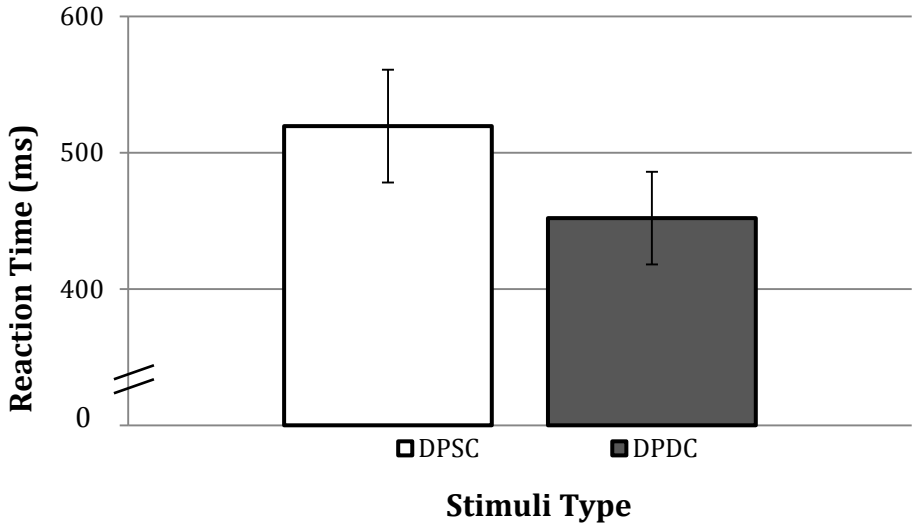

Figure 23: Mean reaction times (ms) and standard errors of stimuli varying by motor class

8.2.3 SECONDARY OUTCOMES

Several secondary analyses were conducted to evaluate other potential factors influencing the primary outcomes. These included investigating the effects of response button position (2 = old versus 4 = old), training amounts, and individual stimuli items within a stimuli set.

8.2.3.1 Effect of Response Button Position

Response button position (i.e., 2 = OLD versus 4 = OLD) was randomly assigned to participants in this study. The effect of button position assignment was evaluated to see if differences in reaction time were evident across stimuli type dependent upon response button position. A 4x2 mixed ANOVA with stimuli type (training, SPSC, DPSC, DPDC) and button group (2 = old, 4 = old) as between-subjects factor was performed on reaction times. The assumption of equality of error variance was met (see Table 38). The assumption of sphericity was met, Mauchly's W = .627, $\chi^2(5) = 9.658$, $p = .086$. The assumption of homogeneity of variance-covariance was met, Box's $M = 11.977$, $F(10, 2314) = .958$, $p = .478$. The assumption of normality was met for all levels of stimuli type across button group except for the DPSC stimuli in button group 2 = old (see Table 39). All other assumptions were met.

Table 38: Test of equality of error variances by stimuli type

Stimuli Type	Levene's Test *F*	*df*	*p*
Trained	2.067	1, 22	.165
SPSC	.000	1, 22	.997
DPSC	.252	1, 22	.621
DPDC	.286	1, 22	.598

Note: 4 significant *p* = .05

Table 39: Test of normality of stimuli type by button group

Stimuli Type	Group	Shapiro-Wilk *W*	*df*	*p*
Trained	2 = old	.876	12	.079
	4 = old	.889	12	.114
SPSC	2 = old	.910	12	.212
	4 = old	.920	12	.289
DPSC	2 = old	.816	12	.014*
	4 = old	.949	12	.616
DPDC	2 = old	.933	12	.407
	4 = old	.876	12	.079

Note: 5 * significant *p* = .05

There was a significant main effect for stimuli type, $F(3,66) = 18.594$, $p \leq .000$, $\eta^2 = .458$. No other significant effects were observed, *ps* > .491. The main effect for stimuli type was similar in pattern to the reaction time pattern presented for RQ4 (Section 8.2.2.1, Figure 22) and RQ5 (Section 8.2.2.2, Figure 23); see Figure 24 below. The means and standard deviations for stimuli type by response button position are reported in Table 40.

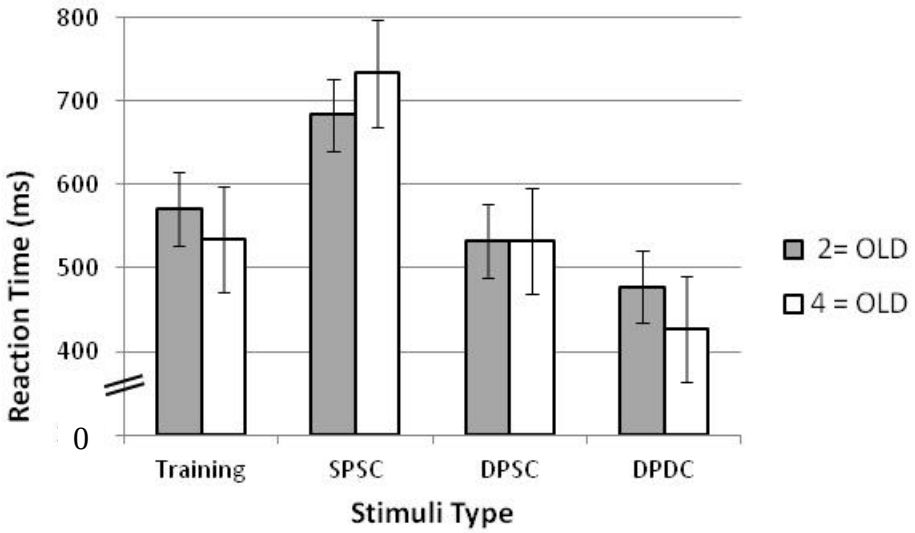

Figure 24: Mean reaction times (ms) and standard errors of stimuli type across button position groups (2 = OLD, 4 = OLD)

Table 40: Mean and SDs for stimuli type by button push group

Stimuli Type	Group	Mean (ms)	SD (ms)
Trained	2 = old	570.78	250.60
	4 = old	534.48	182.65
SPSC	2 = old	683.60	282.45
	4 = old	733.72	286.38
DPSC	2 = old	532.92	219.06
	4 = old	532.67	189.14
DPDC	2 = old	477.34	165.72
	4 = old	426.82	170.61

8.2.3.2 Effect of Training Amounts

Syllable stress training amount was examined across stimuli sets to evaluate whether training amounts affected participants' reaction times on the old-new judgment task. Only one participant required three training blocks to reach accuracy criterion during syllable stress training compared

to participants requiring one block (N = 11) or two blocks (N = 12) of training. This individual's data were collapsed with participants who had completed two blocks of training (N = 13 for the collapsed group) to further evaluate training amounts on reaction times across stimuli types.

A 4x2 mixed ANOVA with stimuli type (training, SPSC, DPSC, DPDC) and training amounts (one block, two or more blocks) as between-subjects factor was performed on reaction times. The assumption of equality of error variance was met (see Table 41). The assumption of sphericity was met, Mauchly's W = .607, $\chi^2(5)$ = 10.34, p = .066. The assumption of homogeneity of variance-covariance was met, Box's M = 12.139, $F(10, 2151)$ = .968, p = .469. The assumption of normality was met for all levels of stimuli type across training amount except for two levels: DPSC stimuli – One Block of Training and DPDC stimuli - One Block of Training (see Table 42). All other assumptions were met.

Table 41: Test of equality of error variances by stimuli type

Stimuli Type	**Levene's Test *F***	***df***	***p***
Trained	.066	1, 22	.799
SPSC	1.336	1, 22	.260
DPSC	3.429	1, 22	.078
DPDC	.569	1, 22	.459

Note: 6 significant = .05

Table 42: Test of normality of stimuli type by training amount

Stimuli Type	Group	Shapiro-Wilk *W*	*df*	*p*
Trained	1 block	.964	11	.815
	2 blocks	.936	13	.403
SPSC	1 block	.888	11	.132
	2 blocks	.886	13	.086
DPSC	1 block	.798	11	.009*
	2 blocks	.923	13	.273
DPDC	1 block	.854	11	.048*
	2 blocks	.954	13	.655

Note: 7 * significant *p* = .05

There was a significant main effect for stimuli type, $F(3,66) = 18.099$, $p \leq .000$, $\eta^2 = .451$. No other significant effects were observed, *ps* > .326. The main effect for stimuli type was similar in pattern to the reaction time pattern presented for RQ4 (Section 8.2.2.1, Figure 22) and RQ5 (Section 8.2.2.2, Figure 23); see Figure 25 below. The means and standard deviations for stimuli type by training amount are reported in Table 43.

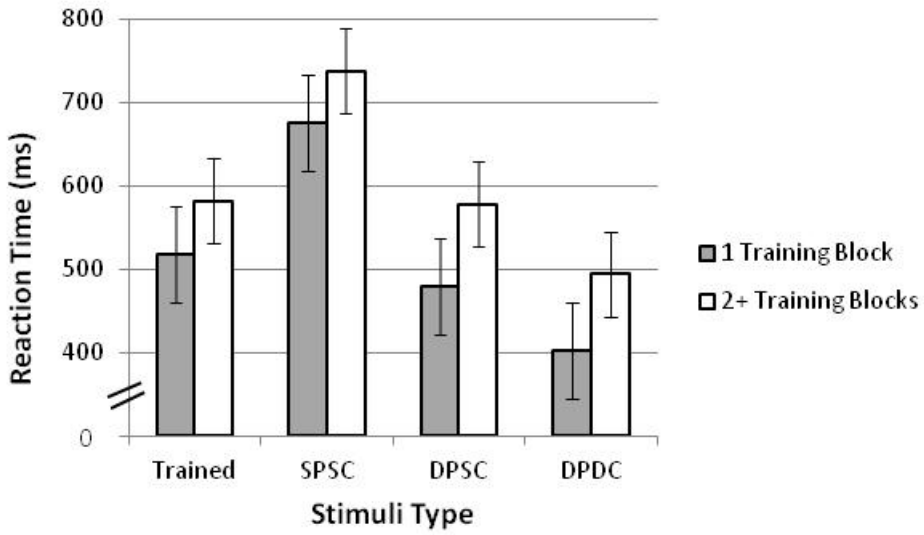

Figure 25: Mean reaction times (ms) and standard errors of stimuli type across training block groups (1 training block, 2+ training blocks)

Table 43: Mean and SDs for stimuli type by training amount

Stimuli Type	Group	Mean (ms)	SD (ms)
Trained	1 block	517.92	231.41
	2 blocks	582.00	205.29
SPSC	1 block	674.83	273.42
	2 blocks	737.28	292.11
DPSC	1 block	479.34	169.65
	2 blocks	578.03	219.00
DPDC	1 block	402.62	150.46
	2 blocks	493.93	173.58

8.2.3.3 Effect of Stimuli Type

To further investigate the transfer pattern related to similarity of the stimuli, an item analysis for each untrained stimuli set (i.e., SPSC, DPSC, and DPDC) was conducted to evaluate potential effects of stimuli construction. The item-by-item data did not meet the normality assumptions required for one-way ANOVA with repeated measurements for any of the stimuli sets; thus,

Friedman's Test was conducted as a non-parametric alternative to the one-way ANOVA with repeated measures.

SPSC Stimuli

There was a significant difference in reaction times across SPSC stimuli items, $\chi^2(9) = 53.52$, $p \leq .000$. Pairwise comparisons were performed (SPSS, 2012) with Bonferroni correction for multiple comparisons. Significant pairwise comparisons are noted in Table 44. The fastest SPSC stimuli (SPSC 1 and SPSC 3) were significantly different from the slowest SPSC stimuli (see Figure 26). The SPSC stimuli nonwords (with trained stressed syllables bolded) and median reaction times are listed in Table 45.

Table 44: Significant pairwise comparison results for SPSC stimuli

SPSC Item #1	SPSC Item #2	*p*
1	10	≤ .000
1	8	≤ .000
1	6	≤ .000
1	2	≤ .000
1	9	≤ .000
3	8	≤ .000
3	6	≤ .000
3	2	≤ .000
3	9	≤ .000

Note: 8: Significance p < .001

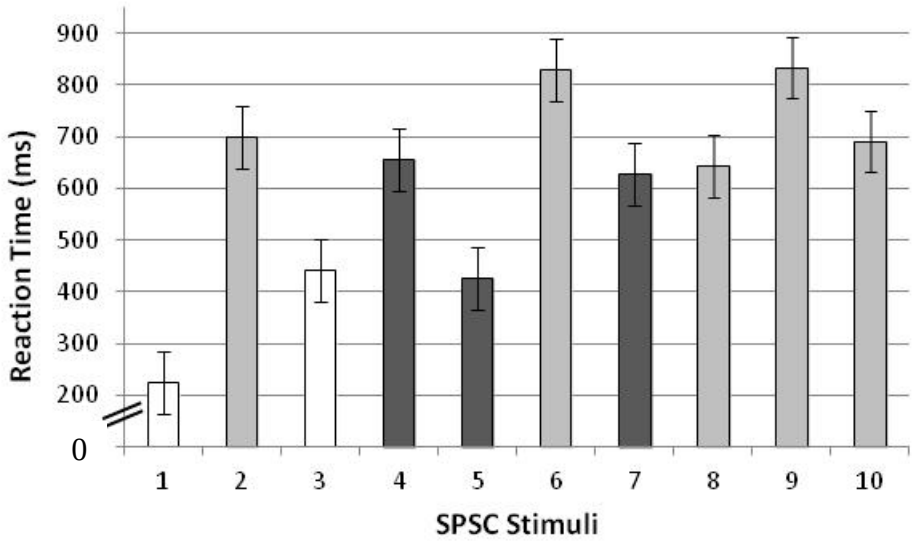

Figure 26: Mean reaction times (ms) and standard errors of individual SPSC stimuli.

Note: 9 The fastest stimuli (SPSC 1 and SPSC 3) are displayed in white compared to the slowest SPSC stimuli (displayed in gray). Any SPSC stimuli not significantly different from SPSC 1 or SPSC 3 are displayed in black.

Table 45: Median reaction times of SPSC stimuli items

SPSC Stimuli	Nonword	Median (ms)
SPSC 1	r a θ ʌ s æ θ	225
SPSC 2	n æ t e r o k	698
SPSC 3	θ o k æ t æ s	442
SPSC 4	θ o s æ d æ k	655
SPSC 5	t e z o n a v	427
SPSC 6	ʃ ɔ z a ʤ ə z	830
SPSC 7	ʃ ɔ z æ ʤ ə θ	628
SPSC 8	z ɔ ʤ ə z æ k	644
SPSC 9	n ɔ z æ ʤ ə θ	834
SPSC 10	n ɔ ʤ ʌ z æ k	691

DPSC Stimuli

There was a significant difference in reaction times across DPSC stimuli items, $\chi^2(9) = 44.07$, $p \leq .000$. Pairwise comparisons were performed (SPSS, 2012) with Bonferroni correction for multiple comparisons. Significant pairwise comparisons are noted in Table 46. The fastest DPSC stimuli (DPSC 4, 5, 6, 7, and 9) were significantly different from the slowest DPSC stimuli (see Figure 27). The DPSC stimuli nonwords (with stressed syllable bolded) and median reaction times are listed in Table 47.

Table 46: Significant pairwise comparison results for DPSC stimuli comparisons

DPSC Item #1	DPSC Item #2	*p*
4	1	≤ .000
5	2	≤ .000
6	1	≤ .000
7	1	≤ .000
9	3	≤ .000
9	1	≤ .000

Note: 10: Significance p < .001

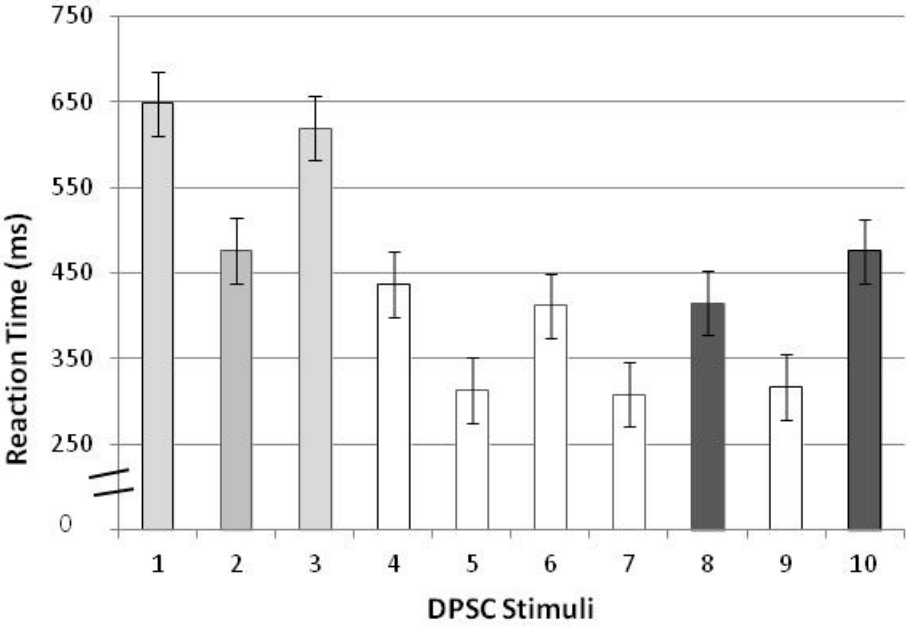

Figure 27: Mean reaction time (ms) and standard error of individual DPSC stimuli.

Note: 11 The fastest DPSC stimuli are displayed in white compared to the slowest DPSC stimuli (displayed in gray). Any DPSC stimuli not significantly different from the fastest DPSC stimuli (DPSC 4, 5, 6, 7, and 9) are displayed in black.

Table 47: Median reaction times of DPSC stimuli items

DPSC Stimuli	Nonword	Median (ms)
1	n æ θ o d æ p	648
2	r a s æ θ o n	476
3	ʃ ɔ ʤ ə z ɔ d	619
4	v u z æ ʃ ɔ m	437
5	s ʌ v e n æ θ	313
6	f o z æ ʃ ɔ d	412
7	k o z æ ʃ ɔ m	308
8	d ɔ ʤ ə z ɔ d	415
9	n a s æ θ o ʃ	317
10	v i ʃ ə d æ k	476

DPDC Stimuli

There was a significant difference in reaction times across DPDC stimuli items, $\chi^2(9) = 22.61$, $p = .007$. Pairwise comparisons were performed (SPSS, 2012) with Bonferroni correction for multiple comparisons; however, there were no significant differences noted between stimuli at $p > .001$ (see Figure 28). The DPDC stimuli nonwords (with syllable stress bolded) and median reaction times are listed in Table 48.

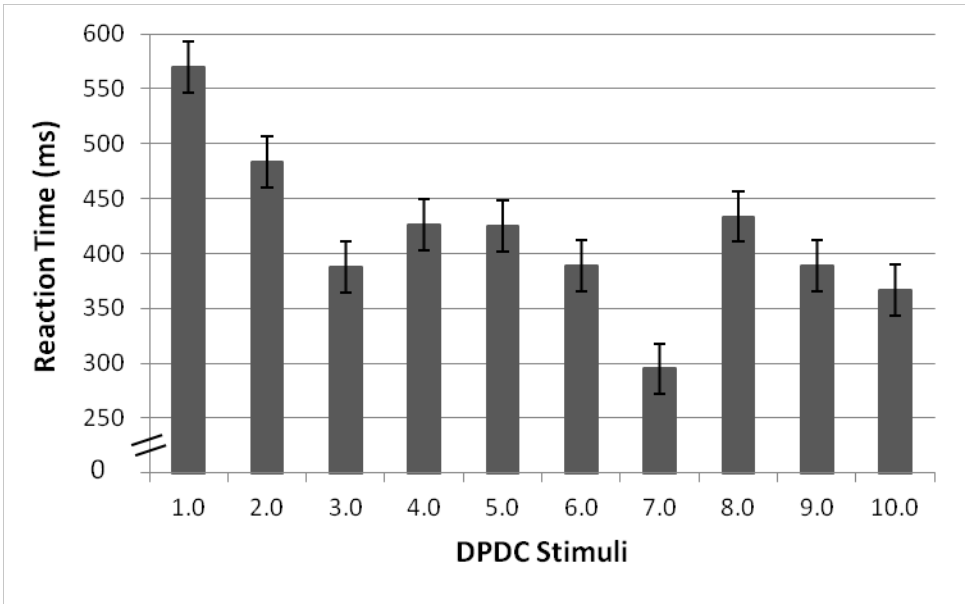

Figure 28: Mean reaction times (ms) and standard errors of individual DPDC stimuli

Table 48: Median reaction times of DPDC stimuli

DPDC Stimuli	Nonword	Median (ms)
1	g i b ɪ ð ɪ b	570
2	b ɪ ð e tʃ ʊ g	484
3	g i g ʊ ð i b	388
4	tʃ e j i w ɪ ʒ	426
5	b ʊ tʃ i tʃ e ʒ	425
6	ʒ i b ʊ tʃ e ð	389
7	tʃ e ð ʊ g ʊ ʒ	295
8	ʒ ʊ g i j ʊ b	434
9	g ʊ g ɪ ð ʊ tʃ	389
10	tʃ ʊ tʃ ʊ b ɪ ʒ	367

9.0 DISCUSSION

9.1 STIMULI VALIDATION STUDY

9.1.1 RQ1: Nonword Frequency

RQ1 of the stimuli validation study evaluated whether Kendall et al.'s (2005) study outcomes were replicated in a younger participant sample. Kendall et al.'s original results indicated older participants had faster reaction times when articulating nonwords with high frequency interphonemic gestures than with nonwords with moderate or low frequency interphonemic gestures. Younger participants in the current stimuli validation study also responded faster to the high frequency interphonemic gesture nonwords compared to the moderate or low frequency interphonemic gesture nonwords. No significant statistical differences were noted in reaction times between the moderate and low frequency interphonemic gesture nonwords in the current study; thus, these two stimuli sets were collapsed into single stimuli set in the main experiment.

9.1.2 RQ2: Real Word Frequency

RQ2 evaluated differences in high versus low lexical frequency words, which provided a control measure to evaluate non-significant differences in RQ1. Participants responded to high lexical frequency words faster than low lexical frequency words. This result is typical of lexical access

studies in young adults (e.g., Balota & Chumbley, 1985; Gahl, 2008; Kang, 2013; Strijkers, Costa, & Thierry, 2010).

9.1.3 RQ3: Stimuli Type

RQ3 evaluated how participants responded to the DPDC stimuli compared to the Kendall et al. (2005) stimuli sets. The DPDC stimuli were created by the author for the main experiment, and varied by two parameters from the Kendall et al. stimuli. First, the stimuli varied by phonetic similarity; specifically, the DPDC stimuli were created with phonemes not found in the Kendall et al. stimuli. Secondly, the stimuli varied by the position of the stressed syllable, which was considered a manipulation of the motor class marker in the main experiment. The DPDC stimuli were assigned syllable stress in the third, final position of the nonword, which is the least frequent position for syllable stress in three syllable words in English (Clopper, 2002). This is compared to the Kendall et al. stimuli in which the stressed syllable occurs in the initial position of the nonword (the most frequent position for three syllable words in English; Clopper, 2002). The manipulations of phonetic similarity and syllable stress position in the DPDC stimuli were intended to produce slower reaction times in participants.

Results indicated participants responded significantly slower to the DPDC stimuli compared to the Kendall et al. (2005) stimuli collapsed across gestural frequency group. This finding suggests the parameters of phonetic similarity and motor class were manipulated successfully in the DPDC stimuli, resulting in slower processing of the DPDC stimuli compared to the Kendall et al. (2005) stimuli. This study did not manipulate these parameters independently from one another; hence, it is unclear from these results if both phonetic similarity *and* motor class contributed to the decreased reaction time versus a single factor.

9.1.4 Summary

The results of the validation study replicated and extended Kendall et al.'s (2005) original findings in a younger sample that included both male and female participants. The reaction times for moderate and low frequency gesture stimuli were not statistically significant from one another. Thus, these stimuli were collapsed into a single stimuli set for the main experiment. Findings from RQ2 support the results from RQ1, ensuring neither the participants nor the equipment was inducing unwanted error into the stimuli validation study. The results of RQ3 provide a baseline of participant response times for the DPDC stimuli. These stimuli were constructed to be as different as possible from the Kendall et al. stimuli based on phonetic similarity and syllable stress assignment, which was predicted to result in slower reaction times. The findings of RQ3 are consistent with this prediction. In summary, the stimuli validation study support the application of Kendall et al.'s moderate and low frequency gesture stimuli, as well as the newly constructed DPDC stimuli, in the main study.

9.2 RULE- VERSUS INSTANCE-BASED LEARNING STUDY

9.2.1 Primary Outcomes

The main experiment investigated the effects of motor class (associated with rule-based learning) and phonetic similarity (associated with instance-based learning) by evaluating reaction times of trained and untrained stimuli in an old-new judgment task. The stimuli for the main experiment

have been graphed in Figure 29 to illustrate the transfer predictions between trained and untrained stimuli hypothesized for each learning theory.[14]

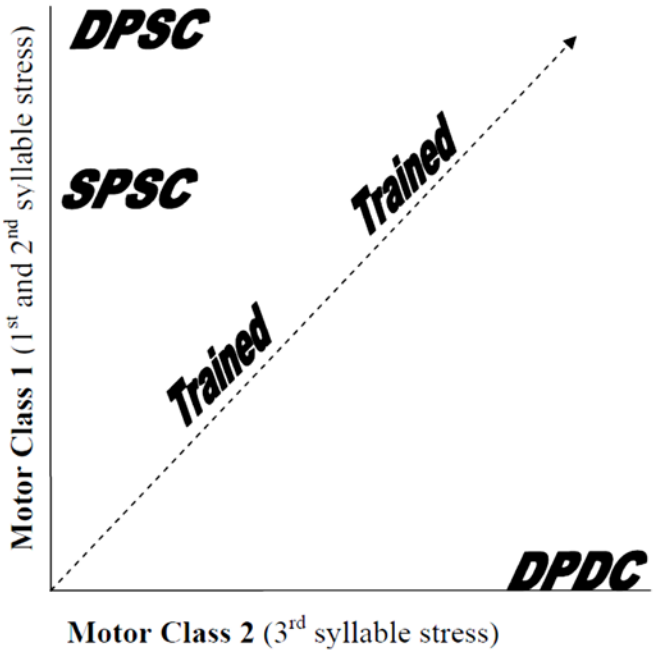

Figure 29: Transfer predictions for each learning theory

Rule-based learning relies on class boundaries to demarcate where one set of rules will end and another set of rules will begin. A class boundary distinguishing Motor Class 1 from Motor Class 2 is represented by a dotted line in Figure 29. The stimuli within Motor Class 1 (Trained, SPSC, and DPSC) all share the same motor class, i.e., first and second syllable stress.

[14]Rule- and instance-based transfer predictions were illustrated in Chapter 3 (Figure 6) using Shank's (1995) hypothetical data set. Figure 29 was graphed in a similar manner to illustrate the same predictions using the stimuli from the main experiment.

However, the DPDC stimuli have stress placement on the final syllable, which is directed by a separate motor class (Motor Class 2). Rule-based learning predicts uniform transfer *within* a motor class as all class members follow the same rules; thus, non-significant differences in reaction time would be predicted for Trained, SPSC, and DPSC stimuli. However, significant differences in reaction time would be predicted for the DPDC stimuli compared to any of the stimuli within Motor Class 1, as the rules learned during syllable stress training no longer apply. Predictions of rule-based learning (and motor class) were evaluated in RQ5.

Motor class dimensions are not relevant for instance-based learning as the abstraction of rules to govern classes of behavior does not occur during instance-based encoding. Instead, instance-based transfer is predicated on the similarity of the trained stimuli to untrained stimuli. Consider the dotted line in Figure 29 as a similarity gradient (instead of a class boundary), in which the basis of comparison for similarity is the trained stimuli. Reaction times are inversely related to similarity of the trained stimuli, i.e., reaction times increase as the similarity between the trained and untrained stimuli decrease. Consequently, the DPSC and DPDC stimuli would have the slowest reaction times compared to the trained stimuli, which is depicted in Figure 29 as an increase in distance between these stimuli from the similarity gradient. The predictions of instance-based learning (and phonetic similarity) were evaluated in RQ4.

9.2.2 RQ4: Phonetic Similarity

The results of RQ4 revealed a significant increase in reaction time on the SPSC stimuli compared to training, but no significant difference in reaction time on the DPSC stimuli compared to training. Initial interpretation of these results suggested both learning theories may be operating during the old-new judgment task based on the hypotheses put forth in Chapter 6

(and illustrated in Figure 29). The significant slowing of reaction times between the trained stimuli and the SPSC stimuli was consistent with instance-based learning. The uniform, consistent reaction times (i.e., non-significant differences) between trained and DPSC stimuli was consistent with rule-based learning. However, further investigation of the results revealed transfer patterns atypical of each learning theory.

From a rule-based learning perspective, it was unclear why the motor class marker for first and second syllable stress was not reliable in directing transfer effects across the Motor Class 1 stimuli. Both SPSC and DPSC stimuli shared the same motor class (i.e., syllable stress position) as the trained stimuli. If the rules learned during syllable stress training were not associated with syllable stress (as intended by the author), it was unclear what other parameter may have been driving the rule-based learning effect noted for the DPSC stimuli. The main difference between the SPSC and DPSC stimuli was the degree of phonetic similarity to the trained stimuli. Phonetic similarity does not seem likely as an alternative set of rules as the most similar phonemes to the trained task were in the SPSC stimuli set, which did not follow typical rule-based learning trajectories. These inquiries motivated the SPSC and DPSC item-by-item analyses.

In the same vein, it was also unclear from the data why instance-based learning was evident for only the SPSC data. As hypothesized in Chapter 6, an increase in reaction time should have been noted as the untrained stimuli systematically differed from the trained stimuli; thus, slowest reaction times should have been observed for the DPSC stimuli. Yet, the DPSC stimuli were responded to faster than the trained nonwords, which indicated participants were able to quickly identify these stimuli as being different from the trained items. Participants were also more consistent in their speed in responding to the DPSC stimuli compared to the trained or

SPSC stimuli (see Table 37), and more accurate in their responses compared to either the trained or SPSC stimuli (Table 34). Thus, these data suggest participants were certain the DPSC stimuli were not similar to the trained stimuli, while participants were less certain of how to respond to the SPSC stimuli. Evaluation of the *d'* statistic data are consistent with this claim.

The increase in *FA* (False Alarms) was greatest for the SPSC stimuli, in which participants erroneously labeled the SPSC stimuli as "old:" 30 errors for SPSC stimuli compared to 5 errors on the DPSC stimuli. The *Br* statistics yielded an estimate of how participants would respond in uncertain conditions, which included the SPSC condition in which portions of the stimuli were the same as the trained stimuli (specifically, the phonemes in a CV unit) whereas other portions were new (i.e., syllable order). As noted in Table 35, participants with high *Br* values were more likely to respond to the SPSC stimuli as "old," and these same participants also had the highest *FA* rates as well. Inquiries into the differences between participants' responses to the SPSC versus DPSC stimuli warranted secondary evaluations of the stimuli (i.e., item-by-item analysis) to examine if phonetic similarity was directing the effects noted in RQ4. Additionally, other variables such as response button position and training amount were further investigated as potential explanations for the inconsistencies noted for RQ4.

9.2.3 RQ5: Motor Class

RQ5 evaluated the effects of motor class when phonetic similarity of the nonwords was held constant. The DPSC training set shared the same motor class as the trained stimuli, and were predicted to have faster reaction times compared to the DPDC stimuli. The DPDC stimuli, on the other hand, were anticipated to have slower reaction times given their Motor Class 2 assignment, see Figure 29. The results of RQ5 indicated a significant class effect, indicative of rule-based

learning; however, this effect was in the opposite direction as would be predicted by rule-based learning. Participants responded faster to the DPDC stimuli compared to the DPSC stimuli. In fact, response times on the DPDC stimuli were faster than response times to the trained stimuli, see Table 49.

Table 49: Mean and SDs of reaction time across stimuli type

Stimuli Type	Mean (ms)	SD (ms)
Trained	552.63	215.26
SPSC	708.66	279.34
DPSC	532.80	200.15
DPDC	452.08	166.50

The predicted significant effect for motor class assumed participants processed the entire nonword (i.e., all nine phonemes) prior to pressing the response box button during the old-new judgment task. The overall faster reaction times noted for the DPDC stimuli set (versus the other nonword stimuli) suggest participants may have determined their old-new judgment prior to the motor class marker (final syllable stress position of the nonword). Participants may have decided early in the DPDC stimuli presentation the nonword was “new,” and not relied on the final syllable stress class marker for their judgments. Unfortunately, it is not feasible with the current methods to know *when* during the stimulus participants were determining their judgments. However, secondary analyses were initiated to investigate other potential factors directing judgments of the DPDC stimuli, including response button position, training amounts, and other potential variations inherent in the DPDC stimuli.

9.2.4 Secondary Outcomes

Several hypotheses were generated to explain the effects noted with RQ4 and RQ5. These included examining the effect of response button position, the amount of syllable stress training required to reach accuracy criteria, as well any potential variation within a given stimuli set.

9.2.4.1 Response Button Position

Participants in this study were randomly assigned a response button position, either response button 2 = Old or response button 4 = Old, following their enrollment in the study. During the old-new judgment task, participants then pushed buttons assigned to these positions to respond to stimuli. As both responses (old and new) were important to the primary outcomes of this study, random assignment of button position was instituted to control for any unwanted effect the response button position may have on reaction time. The results indicate the response button position did not influence reaction times during the old-new judgment task; however, the pattern of reaction times based on stimuli type was the same as the results of RQ4 and RQ5. The effect of response button position as a potential bias in the primary outcomes was eliminated with this finding.

9.2.4.2 Training Amount

The amount of repetition required for optimal learning in speech and nonspeech tasks is variable (e.g., Fox et al., 2002; Maas, Robin, Austermann-Hula, et al., 2008; Meigh & Shaiman, 2010); thus, participants were not required to practice on a certain number of trials during the syllable stress training task. Instead, participants trained to set criteria for syllable stress and phonetic accuracy prior to the administration of the old-new judgment task. The intent of the accuracy

criteria was to have all participants train to a similar memory representation (or baseline representation) for the trained nonwords. Variability in training amounts required to reach accuracy criteria was noted in this study: 11 participants required 120 trials (one block of training), 12 participants required 240 trials (two blocks of training), and one participant required 360 trials (three blocks of training) to meet the accuracy criteria. The results indicated these varying treatment amounts did not influence the reaction times on the old-new judgment task. However, the pattern of reaction times based on stimuli type were the same as the results presented for RQ4 and RQ5. The effect of training amount was eliminated as a potential biasing factor in the primary outcomes.

9.2.4.3 Stimuli Item Analysis

An item-by-item analysis was conducted for each untrained stimulus set (SPSC, DPSC, and DPDC) to evaluate other potential stimuli-dependent variables directing the pattern of results obtained for RQ4 and RQ5. In summary, these results indicated participants were not using the motor class marker (syllable stress) to influence their responses during the old-new judgment task. Instead, participants were basing their judgments on dissimilarity of the phonemes in the untrained stimuli compared to the trained stimuli. Evidence consistent with this conclusion will be described for each stimulus set below.

SPSC

The results of the SPSC item-by-item analysis revealed the fastest responses occurred on nonwords in which the trained stressed syllable occurred in the second position. The slowest responses, on the other hand, occurred in nonwords in which the trained stressed syllable

occurred in the first syllable position; see Table 50. Interpretations of this data were considered for both learning theories.

Table 50: Fastest versus slowest SPSC stimuli

FASTEST SPSC Stimuli	Nonword	Median (ms)
SPSC 1	r a **θ ʌ** s æ θ	225
SPSC 3	θ o **k æ** t æ s	442
SLOWEST SPSC Stimuli		
SPSC 2	**n æ** t e r o k	698
SPSC 6	**ʃ ɔ** z a ʤ ə z	830
SPSC 8	**z ɔ** ʤ ə z æ k	644
SPSC 9	**n ɔ** z æ ʤ ə θ	834
SPSC 10	**n ɔ** ʤ ʌ z æ k	691

Note: 12 Trained stressed syllables are bolded, underlined, and have a larger font for contrast to unstressed syllables.

Rule-based learning predicted uniform, non-significant differences between stimuli sharing the same motor class as the trained stimuli. The SPSC stimuli shared the same motor class, or syllable stress rules; thus, the trained first and second syllable stress positions (highlighted in Table 50) were predicted to result in similar reaction times for all SPSC stimuli. Yet, participants experienced significant increases in reaction time when encountering the trained motor class marker, especially when the stressed syllable occurred in the first position. Trained

stressed syllables in initial position should have signaled the familiar motor class to the participant early in the auditory presentation of the stimuli, allowing participants to ready themselves for a fast response once the stimulus had finished playing. However, the results of the SPSC item analysis revealed trained stressed syllables in initial position had significantly longer reaction times compared trained stressed syllables in second syllable position.

Initial predictions for instance-based transfer effects suggested similarity was essential for transfer to occur between trained and untrained stimuli (Ashby & Maddox, 2005; Brooks, 1978; Chamberlin & Magill, 1992b; Hintzman, 1986; Masson, 1986; Mathews et al., 1989; Palmeri, 1997; Shanks, 1995). In this study, the instance representation was defined as one or two adjacent phonemes in a CV unit, and phonetic similarity was defined as identical instance representations shared between trained and untrained stimuli. As described in Section 7.2.7.4, the SPSC stimuli were constructed with the same phonemes in a CV unit as the training stimuli, but the CV units were swapped for first and second syllable position. Thus, the SPSC stimuli systematically varied in phonetic similarity from the trained stimuli, and a significant slowing in reaction time *was* anticipated. However, it was unclear why phonetic similarity did not impact the DPSC stimuli, which were more dissimilar from the trained stimuli than the SPSC stimuli. Thus, a similarity account of instance-based learning was not consistent with the transfer effects noted across all untrained stimuli in this experiment. An alternative theory of instance-based learning, as noted in Chapter 3, postulated transfer was based on dissimilarity between trained and untrained stimuli (e.g., Ward & Churchill, 1998; R. L. Wright & Burton, 1995); thus, further investigation and reevaluation of the instance-based literature from a dissimilarities account is described below.

Several instance-based accounts of categorization and recognition interpret the construct of similarity as a measurement of distance within psychological space, in which "similarity" of two memory representations is a mathematical function of their distance to one another (e.g., Nosofsky et al., 2011; Nosofsky, 1986, 1992). The more dissimilar trained and untrained stimuli are in their features, the greater the measured psychological distance between them (Nosofsky et al., 2011; Nosofsky, 1992; Shanks, 1995). During an old-new judgment task, participants would evaluate the distance between trained and untrained stimuli, and determine if the judged distance exceeded a threshold for the category "new." If this threshold was not exceeded, the stimuli would be judged as "old" (Nosofsky et al., 2011; Nosofsky & Stanton, 2006; Nosofsky, 1992; Takane & Sergent, 1983). This version of instance-based learning predicts faster reaction times for untrained stimuli that are maximally dissimilar from trained stimuli. Specifically, untrained stimuli with distinct, novel features have a larger psychological distance in relation to trained stimuli, and the threshold for "new" can be determined quickly (Nosofsky et al., 2011; Nosofsky, 1986, 1992). Untrained and trained stimuli that share features are closer in psychological space, making the determination of "new" more difficult to discern quickly (Nosofsky et al., 2011; Nosofsky, 1992; Shanks, 1995). Evaluation of reaction time patterns in categorization (A. L. Cohen & Nosofsky, 2003; Lamberts, 1995, 1998, 2000; Nosofsky et al., 1992; Nosofsky & Palmeri, 1997; Nosofsky, 1991) and recognition tasks (Lamberts, Brockdorff, & Heit, 2003; Nosofsky et al., 2011; Nosofsky & Stanton, 2006; Nosofsky & Zaki, 2003; Nosofsky, 1988, 1991) are consistent with these predictions. Additionally, untrained stimuli with novel, untrained features result in the fastest reaction times in recognition tasks; participants are able to base their judgments solely on the novel features, and do not require additional comparisons to the trained

stimuli to exceed the “new” threshold (e.g., Johns & Mewhort, 2002, 2003; Mewhort & Johns, 2000, 2003; Nosofsky et al., 2011).

Instance-based theories of phonology provide similar evidence; however, the construct of similarity is framed within a phonological neighborhood. A phonological neighbor is a word varying from a target word by one phoneme (Grainger, Muneaux, Farioli, & Ziegler, 2005; Yates, Locker, & Simpson, 2004). A dense phonological neighborhood has many words that share phonemes, whereas sparse neighborhoods have few words that share phonemes (Luce & Pisoni, 1998; Vitevitch & Luce, 1998, 1999). Results from auditory identification studies indicate words with low-density neighborhoods are identified more quickly than words in higher density neighborhoods in the presence of noise (Goldinger, Luce, & Pisoni, 1989; Luce & Pisoni, 1998; Vitevitch & Luce, 1999). Additionally, faster reaction times are noted during auditory lexical decision tasks with words in low-density versus high-density neighborhoods (Luce & Pisoni, 1998; Vitevitch & Luce, 1999). Slower reaction times in high-density neighborhoods are hypothesized to occur as increased lexical activation increases competition between phonological neighbors, and overall discrimination between neighbors becomes more difficult (Luce & Pisoni, 1998; Vitevitch & Luce, 1998). Phonological neighborhood effects are also noted with nonwords with low probability phonological segments (e.g., Kendall et al. and the DPDC stimuli), in which faster reaction times are noted with nonwords in low-density neighborhoods (Luce & Pisoni, 1998; Vitevitch, Luce, Pisoni, & Auer, 1999; Vitevitch & Luce, 1999).

The evidence presented for a dissimilarity account of instance-based learning is consistent with the data presented in Table 50. The SPSC stimuli with the fastest responses included SPSC 1 with an initial phoneme never trained during syllable stress training. The untrained phoneme may have acted like a extraneous feature allowing participants to judge the

nonword as “new” without relying on the rest of the phonemes (cf., Johns & Mewhort, 2002, 2003; Mewhort & Johns, 2000, 2003; Nosofsky et al., 2011). Nonwords sharing only a single phoneme, such as SPSC 3, were also classified as “new” significantly faster than other SPSC stimuli. These stimuli may be considered to have sparse phonological neighborhoods, allowing for significantly faster discrimination as fewer shared phonemes were present and phonological competition was minimal (cf., Luce & Pisoni, 1998; Vitevitch et al., 1999; Vitevitch & Luce, 1999).

Evaluation of the slowest responses within the SPSC stimuli may also be explained by dissimilarity of the SPSC stimuli to the trained stimuli. The slowest SPSC stimuli have a trained stressed syllable in the initial position. These syllables may have been attended to longer than untrained syllables during syllable stress training, as the task instructions emphasized syllable stress accuracy. This additional attention may have altered the inter-stimulus similarity values (and overall distance) within psychological space of these stimuli compared to the trained stimuli. Thus, determining where the “new” threshold occurred may have required more processing time, which increased reaction time (Nosofsky et al., 2011; Nosofsky & Stanton, 2006). Within a phonological neighborhood framework, the trained syllable may have increased the number of potential neighbors, and discrimination between “old” and “new” became more difficult to discern. This interpretation is consistent with evidence from phonological window gating studies, in which slowed reaction times are noted with words in high-density neighborhoods. Increased phonological information (i.e., a larger phonological window) is required to discriminate neighbors that share initial phonemes (Bowey & Hirakis, 2006; Mainela-Arnold, Evans, & Coady, 2008). This phonological information present in the initial portion of the word (e.g., first few phonemes) filters and decreases the number of activated

words during a word recognition task, which increases the accuracy of identifying the correct word (e.g., Bowey & Hirakis, 2006; Cotton & Grosjean, 1984; Dollaghan, 1998; Grosjean, 1980; Mainela-Arnold et al., 2008; Montgomery, 1999). Words with high-frequency (or well-learned) phonetic features require significantly larger phonetic windows (i.e., more phonetic information) before accurate and reliable word identification can be determined (Cotton & Grosjean, 1984; Grosjean, 1980; Mainela-Arnold et al., 2008; Tyler, 1984).

What is unclear from the SPSC stimuli analysis is if the syllable unit or the adjacent position of the phonemes in a CV unit is important for the judgment. As noted earlier, the nature of the stress placement did not seem to aid participants' speed in making a judgment about the stimuli; however, the syllable unit is confounded with the phonetic similarity unit (i.e., two adjacent phonemes in a CV unit). This limitation will be addressed later in Section 9.3.1.

DPSC

Results of the item-by-item analysis of the DPSC stimuli are also consistent with a dissimilarity account of instance-based learning. The dissociation of syllable unit from two adjacent phonemes is more pronounced than in the SPSC stimuli, see Table 51. Of the five DPSC stimuli with the fastest response times, two of the stimuli (DPSC 4 and 6) had initial phonemes not present in the trained stimuli set. These phonemes are circled in Table 51. The remaining three fastest DPSC stimuli (DPSC 5, 7, and 9) share only the initial phoneme with the trained stimuli (also circled in Table 51). The second phoneme was novel to the participant, and would have signaled the first CV unit was not similar to the trained stimuli.

Table 51: Fastest versus slowest DPSC stimuli

FASTEST DPSC Stimuli	FASTEST DPSC Stimuli	Median (ms)
4	v u z æ ʃ ɔ m	437
5	s ʌ v e n æ θ	313
6	f o z æ ʃ ɔ d	412
7	k o z æ ʃ ɔ m	308
9	n a **s æ** θ o ʃ	317
SLOWEST DPSC Stimuli	**Nonword**	**Median (ms)**
1	**n æ** θ o d æ p	648
2	r a **s æ** θ o n	476
3	**ʃ ɔ** ʤ ə z ɔ d	619

Note: 13 Trained stressed syllables are bolded, underlined, and have a larger font for contrast to unstressed syllables. Circled phoneme(s) indicate novel, untrained phonemes. All trained stressed syllables noted in the DPSC stimuli were not trained in the same syllable position (e.g., the initial trained syllable in DPSC 1 occurred in second syllable position in the trained nonword).

The slowest DPSC stimuli had initial CV units that were trained stressed syllables. It was postulated during the SPSC item analysis that training instructions may have increased attention to the trained stressed syllable. It was not conclusive, however, whether the focus of attention during training was on the adjacent phonemes in the trained CV unit or the stressed syllable. This dissociation may be more distinct upon examination of DPSC 2. This nonword had a trained

stressed syllable in second syllable position; however, it also shared an initial untrained first syllable unit with SPSC 1 (see Table 50). It is postulated if stressed syllables (versus adjacent phonemes) were the unit directing transfer effects, response times on DPSC 2 had the potential to be twice as long as response times for the other two slow DPSC stimuli (DPSC 1 and 3). This would occur because the participant would need to process two familiar stressed syllables prior to encountering an unfamiliar syllable (i.e., the final syllable in the nonword). Yet, this nonword is the "fastest" of the three significantly slower DPSC stimuli, and has response times similar to those of some of the faster DPSC stimuli. This may have occurred because participants made their judgment of "new" following the third phoneme, i.e., when the similarity (based on adjacent phonemes) to any other stimuli was no longer present. This postulation would be consistent with instance-based learning accounts in which dissimilarity predicts transfer performance. Unfortunately, this is only speculation as this study cannot infer when the participant decided a given stimulus was old or new, or what phonemes triggered this response.

DPDC

Although a statistically significant difference was obtained between the individual stimuli in the DPDC set, the individual stimuli reaction time differences did not meet the stringent pairwise alpha level ($p > .001$) to further evaluate the reaction time pattern. The DPDC stimuli were constructed to be maximally different from the trained stimuli in phonetic similarity and motor class. The increased reaction times noted for RQ5 (motor class) were in the opposite hypothesized direction for both rule- and instance-based learning. As noted earlier, a rule-based account would have predicted slower reaction times as the syllable stress rules learned in training were no longer useful in directing recognition of the stressed syllable in final position of the DPDC nonwords. A similarity account of instance-based learning would predict slowed reaction

times as the untrained initial phonemes were not present in training. Neither of these accounts is consistent with the results of RQ5 or the DPDC item analysis. However, a dissimilarity account of instance-based learning would predict these results.

As noted in the primary outcomes for RQ5, the “no response” data suggest participants may not have been attending to the entire DPDC nonwords prior to making their judgments. This claim is further supported by the DPSC stimuli analysis listed in Table 52. All of the DPDC stimuli begin with an untrained phoneme. The associated median reaction times noted in Table 52 are within the same range of reaction times reported for the fastest DPSC stimuli, which also start with an untrained phoneme (see Table 51). These novel phonemes may have differentiated the DPDC stimuli from the trained stimuli right away, allowing the participant to make a judgment quickly regarding the familiarity of the nonword without requiring further analysis of the final syllable. However, this speculation cannot be confirmed with the current methods employed to assess participants’ responses.

Table 52: Median reaction times of DPDC stimuli

DPDC Stimuli	Nonword	Median (ms)
1	g i b ɪ **ð ɪ b**	570
2	b ɪ ð e **tʃ ʊ g**	484
3	g i g ʊ **ð i b**	388
4	tʃ e j i **w ɪ ʒ**	426
5	b ʊ tʃ i **tʃ e ʒ**	425
6	ʒ i b ʊ **tʃ e ð**	389
7	tʃ e ð u **g ʊ ʒ**	295
8	ʒ ʊ g i **j ʊ b**	434
9	g ʊ g ɪ **ð ʊ tʃ**	389
10	tʃ ʊ tʃ u **b ɪ ʒ**	367

Note: 14 Stressed syllables are bolded, underlined, and have larger font. These syllables were untrained.

9.2.5 Summary

The results of the primary and secondary analyses of the main experiment suggest instance-based learning, not rule-based learning, was a factor in the acquisition of the experimental stimuli in this study. Results of RQ4 (phonetic similarity), and the item-by-item analyses, reveal the fastest reaction times occurred when stimuli began with novel, untrained phonemes. This result is consistent with instance-based theories of categorization and recognition in which transfer is predicted based on the dissimilarity between trained and untrained stimuli (e.g., Lamberts et al., 2003; Nosofsky et al., 2011; Nosofsky & Stanton, 2006; Nosofsky & Zaki, 2003; Nosofsky, 1988, 1991), as well as theories of phonological neighborhood density (e.g., Goldinger et al.,

1989; Luce & Pisoni, 1998; Vitevitch & Luce, 1999). These accounts also provide a more cohesive explanation for the significant effect noted for RQ5 (motor class). The trained class marker, first and second syllable stress positions, did not affect reaction times in any of the untrained stimuli as predicted by rule-based learning theory. However, the “no response” data and significantly faster reaction times suggest participants may have based their recognition judgments on novel adjacent phonemes and not the motor class marker, i.e., the final syllable. The current methods can only provide speculation for this hypothesis; however, more sensitive temporal measures may be able to address this question directly in future studies.

In summary, the results of the main experiment suggest an alternative learning theory, instance-based learning, should be further investigated as a potential explanation for learning and transfer effects associated with speech production. Rule-based learning was not observed in this experiment where syllable stress was postulated as the motor program, and high-frequency syllable stress positions were considered within-class behaviors. However, other motor programs or motor classes may have altered the results. A smaller unit (e.g., phoneme) or larger unit (e.g., word) may be the motor program operating during speech production. However, the invariant features associated with Schmidt’s (1975) motor program theory would need to be specified. Further understanding of the representational forms underlying rule- and instance-based learning theories may also provide insight into potential interactions between these two theories. The next sections, Limitations and Future Directions, will address future parameters of interest in investigating the differences between rule- and instance-based learning in speech production.

9.3 LIMITATIONS

There were several limitations in the current study. The main concerns relate to confounds between CV and syllable units within the stimuli, potential instructional bias, the measurement of implicit memories, the measurement of syllable stress, and specific concerns about the syllable training program *Stimulate*. Each of these limitations will be described separately below.

9.3.1 Construction of the Stimuli: Syllable versus adjacent phonemes

The application of Kendall et al.'s (2005) stimuli decreased the likelihood participants were using well-learned representations to aid their judgments during the old-new judgment task. The controlled construction of these nonwords, however, also confounded the differentiation of the postulated rule- and instance-based representations for this experiment. The trained stressed syllable, composed of two adjacent phonemes, was proposed for the rule-based representation. One or two adjacent phonemes within a CV unit were postulated as the instance-based representation. These two representational forms were difficult to parse out in the first or second syllable position, as all the nonword stimuli were composed with the following phonetic form: CV|CV|CVC (in which the line indicates a syllable break). The current methods were not designed to differentiate these representational units, and are not sensitive in determining what portion of the nonword stimuli were influencing participants' judgments. Instead, the patterns of reaction time across the parameters of interest (phonetic similarity and motor class) were examined in an effort to distinguish these two theories of learning.

Future studies using measures designed for temporal specificity, such as event related potentials (ERP), may provide insight into which representational unit within a nonword is

influencing recognition judgments. Within language processing, the mismatch negativity (MMN) component of the ERP signal has been identified as reflecting discrimination between acoustic and phonetic properties of stimuli (Cheour, H.T. Leppänen, & Kraus, 2000; Friederici, 2005; R. Näätänen, Paavilainen, Rinne, & Alho, 2007; Risto Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). The MMN has been investigated in discrimination studies of phonemes, including consonant place of articulation (e.g., Dehaene-Lambertz & Dehaene, 1994; Shafer, Schwartz, & Kurtzberg, 2004; for a review see: Molfese, Key, Maguire, Dove, & Molfese, 2008), consonant duration (Leppänen et al., 2002), vowel differences (e.g., Cheour et al., 1998; Friederici, Friedrich, & Weber, 2002; Pihko et al., 1999), as well as differences in stressed syllable position (Dehaene-Lambertz, 1997; Honbolygó & Csépe, 2013; Weber, Hahne, Friedrich, & Friederici, 2004). A replication of the old-new judgment task while monitoring the MMN component may provide insight into the temporal location of the judgment as the stimuli are being played. Additionally, redesign of the current stimuli to differentiate the rule and instance representations, e.g., increasing the size of the first syllable to CVC, may also aid discrimination using ERP measures.

9.3.2 Instructional Bias

The old-new judgment task instructions were constructed to be neutral in describing the nonword stimuli so as to avoid biasing participants towards either learning theory (see Chapter 7, Section 7.2.6.7). Experimental manipulation of instructions to bias learning outcomes has been observed for both learning theories (as noted in Section 7.2.6.7), although these manipulations are not always successful (e.g., Dienes et al., 1991; Mathews et al., 1989; Vokey & Brooks, 1992, n. Experiment 1). It is unclear what underlying factors bias participants to perform in different

ways using a single instruction set. Instructions may direct participants to focus their attention to particular aspects of a task, which may allow for specific encoding or retrieval of memories to occur (e.g., R. E. Smith & Hunt, 2000; Sussman, Winkler, Huotilainen, Ritter, & Näätänen, 2002; Taylor & Fiske, 1978). If this is the case, instructional bias would have the potential to change some of the primary outcomes in this study.

The old-new judgment instructions in the main experiment required participants to make quick, accurate familiarity judgments. This may have biased participants to focus their attention on the *initial* portion of the nonword in an effort to increase the speed of their response. An interesting experimental manipulation using the current instructions might include placing untrained (unstressed) syllables earlier in the DPDC stimuli to see if the predicted rule-based motor class effect would be observed (RQ5). In this manipulation, participants would not need to wait for the final syllable of the DPDC stimuli to make a judgment of motor class; instead, the potential contrast between representational units would be present in the initial position. Other experimental manipulations might include creating two separate instruction sets meant to direct participants' focus of attention to a given representational unit (e.g., syllable versus initial phoneme). Application of these instructions across participants may aid our understanding of whether or not both learning theories may be operating during the same task, and whether focus of attention may modulate learning outcomes. Manipulation of rule- versus instance-specific instructions within the same participant may provide clarification as to the constraints of learning for each theory if an order effect of instruction set is observed.

Additionally, instructions during administration of the recognition probes may have directed additional focus to the instance-based representation during training, which may have altered the encoding of the trained stimuli. The training portion of the experiment was designed

to ensure both representations were encoded accurately. Feedback from the program *Stimulate* was provided to aid the encoding and refinement of the hypothesized rule-based representation, i.e., syllable stress in the first and second position of the trained stimuli. Quick and accurate identification of the trained stimuli in the recognition probe was used to estimate the phonetic accuracy of the hypothesized instance-based representation, i.e., one to two adjacent phonemes in a CV unit. Although all participants met accuracy criteria for both syllable stress training and the recognition probe prior to initiating the old-new judgment task, it is unclear if instructions for either task (syllable training or recognition probe) biased encoding. For example, the instructions to focus on phonemes in the first recognition probe may have altered participant's focus of attention during additional training blocks. Changing the recognition probes to include syllable stress foils, as well as changing the instructions to be more global (i.e., to focus on the phoneme *and* syllable stress), would decrease this potential bias in future studies. Furthermore, changing the syllable stress training instructions to include accurate production of all "sounds" (not just syllables) during syllable stress training would also decrease instructional bias during training.

9.3.3 Measuring Implicit Memory with an Explicit Task

Historically, memories for motor behavior have been classified as procedural knowledge. Procedural knowledge is defined as knowledge of how we perform actions or skills (J. R. Anderson, 1995; Reisberg, 2001; Tomporowski, 2003), and is characterized as automatic (D. Lieberman, 1990), difficult to describe (J. R. Anderson, 1980; Galotti, 1994; Hunt & Ellis, 2004; Reisberg, 2001), and encoded and retrieved from memory unconsciously (J. R. Anderson, 1995). Implicit measures of learning (e.g., priming measures) do not require verbalization or conscious processing, and are used to assess procedural knowledge (Proctor & Dutta, 1995; Seger, 1994;

Shiffrin & Dumais, 1981; Tomporowski, 2003). However, the old-new judgment task used in this experiment is an explicit measure in which participants are required to compute a response of “old” versus “new” for each trial following the auditory presentation of a nonword. Explicit measures are direct tests of memory that are consciously accessed and verbalized by an individual (J. R. Anderson, 1995; Proctor & Dutta, 1995; Reisberg, 2001; Schacter, 1987). The old-new judgment, therefore, may not reflect the underlying processing occurring during the judgment or accurately define memory retrieval during the response selection stage. Thus, the results of this study may have been more conclusive if the methods and underlying processes were aligned (e.g., implicit tasks probing implicit memories).

As noted in Chapters 4 and 7, the old-new judgment task was a novel assessment for evaluating the response selection stage of motor control as the majority of speech motor control research evaluates the motor response at the level of response programming (e.g., Deger & Ziegler, 2002; Maas & Mailend, 2012; Maas, Robin, Wright, et al., 2008; D. E. Meyer & Gordon, 1985; Rogers & Storkel, 1998). However, incorporation of an implicit measure in conjunction with the old-new recognition task would provide a more parsimonious explanation of the underlying encoding and retrieval of memories during the response selection stage. This could be achieved by evaluating temporal measures during the old-new judgment task. The use of ERP measures has already been described in the previous section, and may provide further temporal information as to when the judgment is occurring in real time prior to the summary output measured by reaction time. Other implicit measures, such as priming, during the recognition probes may also provide insight into the memory representation being encoded. This could be achieved by measuring reaction times during the recognition probes. Faster reaction times are predicted for the trained stimuli and foils with portions of the trained stimuli present

(e.g.,Cholin & Levelt, 2009; Cholin, Schiller, & Levelt, 2004; A. S. Meyer, 1990, 1991; Roelofs & Meyer, 1998; Rosenbaum et al., 1984). The foils of the recognition probes would need to reflect both memory representations, i.e., syllable (rule-based) and phoneme (instance-based), to not bias the results to one or the other learning theories.

9.3.4 Measuring Syllable Stress During Training

The proposed rule-based representation for this study was syllable stress. Syllable stress is perceptually marked by intensity, pitch, and/or duration of the syllable to indicate stress (for a review: Gay, 1978). These perceptual markers align with invariant features of motor programs (Schmidt, 1975), which are popular rule-based representations in motor speech models (See Chapter 2, Table 1). For this experiment, the main marker of syllable stress production was intensity, and participants were instructed to manipulate only this variable when producing stressed syllables during training. The application of intensity, instead of other syllable markers, was intentional, as intensity yielded the most salient visual cues for participants during feedback.[15] However, intensity was not the *natural* stress marker produced by many of the participants during training (see section 7.2.6.3, Table 16), and perceptual judgments of syllable stress accuracy by the examiner were needed to account for this unexpected outcome.

[15] A variety of syllable markers were manipulated during the programming of *Stimulate* for the main experiment. However, only intensity provided a feedback visual display that was 1) accurate in measuring syllable stress during the pre-defined syllable duration windows, and 2) provided the most visual contrast between stressed and unstressed syllables.

Further review of the literature reveals limited consensus as to which perceptual marker (or combination of markers) is required to signify syllable stress. Early investigations emphasized a single, dominant perceptual marker as the salient stress marker, e.g., increase in pitch (Fry, 1958) or duration (C. Adams & Munro, 1978; Klatt, 1976; Nakatani & Aston, 1978; Sluijter & Van Heuven, 1996). However, other research suggests individual variation and linguistic context determine which stress (or combination of stress) marker(s) are most salient (Isenberg & Gay, 1978; P. Lieberman, 1960; Morton & Jassem, 1965). From a theoretical perspective, each of these perceptual markers may provide a different experimental variable to investigate for future class effects. Although the invariant feature of force was examined by measuring intensity, a different invariant feature may provide a better estimate of class (relative timing as perceived as an increase in duration; e.g., Schmidt, 1975; Gabriele Wulf & Schmidt, 1988).

Future studies investigating syllable stress as a potential rule-based representation will need to define the invariant feature characterizing the motor class, e.g., relative timing. Then additional screening for the perceptual correlate (e.g., increased duration to signify syllable stress) may be assed prior to training. These results will allow for screening of participants for specific perceptual stress markers (e.g., intensity or duration only), or provide the information required to customize training based on individual stress markers. The application of *Stimulate* in these future studies may be problematic; however, other acoustic software and equipment may provide visualization of combined acoustic parameters (*Real-Time Pitch, Model 5121*, KayPentax).

9.3.5 Training Program *Stimulate*

Deriving intensity levels and displaying intensity information in a visual feedback display *during* training was unique to this experiment as previous studies measuring the acoustic parameters of syllable stress required all acoustic analyses to occur post data collection (as reviewed in the following: Forrest & Weismer, 1997; Harrington & Cassidy, 1999; Kent & Read, 2002). The custom software, *Stimulate,* provided an analysis of the acoustic parameters produced *during* a participant's nonword production, as well as displayed feedback regarding accuracy in meeting the targeted intensity values during syllable stress production. Both of these computational functions have the potential to be labor-intensive and time-consuming; however, *Stimulate* was able to compute this information within seconds of a participant's production. This time savings allowed for the implementation of the study design (i.e., syllable stress training *and* judgment task) within a single experimental session lasting no longer than three hours.

Despite these advantages, further modifications to *Stimulate* would be required for future studies to ensure the validity and reliability of the feedback provided to participants. As noted above in Section 9.3.3, customization of the acoustic parameters analyzed during training would be needed to incorporate individual variation in syllable stress production. Additional control in the timing of acoustic data collection and analysis during participants' productions would also be required if the Kendall et al. (2005) stimuli were to be employed again in future studies. In the current version of *Stimulate,* all three syllables were evaluated for intensity values. However, as explained in Section 7.2.6.3, the construction of some of the Kendall et al. (2005) stimuli produced an unwanted increase in air flow during the end of the nonword production, which altered the accuracy of the feedback that participants received. Alterations to *Stimulate* to evaluate only the first and second syllables would remove this potential for error, and provide

participants with valid feedback. Other additional modifications may need to be addressed in the future to incorporate other hypothesized representational forms (e.g., phonemes or words) or other motor learning variables of interest (e.g., feedback frequency).

9.4 FUTURE DIRECTIONS

Historically, rule- and instance-based learning theories have been contrasted in an effort to determine the representational form of the memories being encoded and retrieved during learning. Strong versions of each learning theory do not predict interaction effects of learning. For example, a strong version of rule-based learning would state only relevant features would be encoded during training, and item-specific transfer (based on irrelevant features) would not be feasible (see Chapter 2). However, weak versions of instance-based theory have been postulated, and may provide a viable framework for evaluating potential interaction effects in future studies of speech production.

Rawson's (2004) instance-based model of reading comprehension relies on a hierarchical organization in which higher organization levels (e.g., synthesizing sentences for content) rely on lower organization levels (e.g., letter and word decoding). Frequent, repeated exposure to letters and words at lower organization levels generate predictable patterns of memory retrieval (i.e., "rules" for aggregation of specific memories). Over time, the rules for aggregation become so speeded they are considered automatic (Rawson & Middleton, 2009; Rawson, 2004). Instance-based theories of phonology also rely on rule-based aggregation principles to retrieve instance-based phonological memories (e.g., Bybee, 2006; Goldinger, 1998; Nielsen, 2011; Pierrehumbert, 2002; Schweitzer & Möbius, 2004; M. Walsh et al., 2010). These models

postulate higher organizational language levels (e.g., word level) direct predictable patterns of phoneme retrieval at lower organization levels during word recognition (Nielsen, 2011; Pierrehumbert, 2001, 2002; M. Walsh et al., 2010). As noted in Chapter 4, skill acquisition theories vary in their description of how the memory representation evolves with practice and experience (cf. J. R. Anderson et al., 2004; Logan, 1988). These weaker versions of instance-based learning further describe skill acquisition in which aggregation principles evolve with practice, i.e., relying on computational, slowed aggregation early in learning followed by efficient rule-based aggregation with continued practice.

Current speech motor control models also rely on hierarchal levels of control (see Appendix A for examples), but rely on rule-based representations to direct processing across each level. However, using the framework provided by instance-based models of language and reading comprehension, an instance-based memory model of speech motor control could be conceptualized. Rule-based patterns of instance aggregation may emerge from experience, in which retrieval of motor commands for familiar sequences (e.g., syllables or high frequency words) occurs simultaneously. Speech development may provide the most concrete example of this hypothetical model. Early in development, children are encoding instance representations with each babbling attempt. With articulatory experience, children learn to produce specific articulatory patterns in certain contexts (e.g., saying “dada” in the presence of a father figure). As speech and language continue to develop, repeated rule-based aggregation patterns at lower organizational levels would support the coordination and construction of larger speech units being aggregated at higher organizational levels (e.g., words, phrases). Although the instance-based memory portion of this model is speculative, developmental models of speech motor control have similar physiologic hierarchical structures (e.g., B. L. Davis, MacNeilage, &

Matyear, 2002; Green & Nip, 2010; Guenther, 1994; MacNeilage & Davis, 1990). In these models, movement patterns controlled by lower organization levels (e.g., oscillation movements of the jaw) must be stabilized through experience and maturation prior to the execution of more advanced articulatory patterns, e.g. variegated babbling (e.g., Cheng, Murdoch, Goozee, & Scott, 2007; Green, Moore, Higashikawa, & Steeve, 2000; Green, 2002; A. Smith & Zelaznik, 2004; B. Walsh & Smith, 2002). Continued research investigating learning theories in speech motor control have theoretical implications for skill acquisition and learning trajectories in adults, but may also provide a more comprehensive theory of motor speech development as well.

Further development of this programmatic line of research may also inform speech motor clinical treatment goals As reviewed in earlier chapters, most theories of information processing models of motor control revolve around rule-based learning representational forms, e.g., motor program, and their associated transfer effects. This is also the case in motor-focused treatments in speech therapy (e.g., Aichert & Ziegler, 2004; Ballard, Granier, & Robin, 2000; Clark & Robin, 1998; Maas, Robin, Wright, et al., 2008). The current results suggest this model may not be appropriate, and may provide a theoretical explanation for why clinical performance using motor program theory is inconsistent (e.g., Austermann-Hula et al., 2008; Ballard et al., 1999; Knock et al., 2000). Instance-based theories, or mixed models of rule-and instance-based learning, may provide a more comprehensive explanation for successful clinical outcomes and theoretical basis for structuring new treatment goals. For example, clinical treatment may focus on maximizing retrieval cues through specific instruction instead of refining a motor program in therapy. However, before these clinical manifestations can occur, a paradigm shift would be needed within the field of speech motor control. The main experiment initiated this theoretical

step, and continuation of this programmatic line of research will allow for a reevaluation of learning in speech motor control theory.

APPENDIX A

MODELS OF MOTOR PERFORMANCE

The models presented in this appendix depict information-processing models of motor behavior in which the motor representation is selected prior to motor programming and execution. As noted in Chapter 4, each model has variation in which the four main components (i.e., cognitive input, response selection, response programming, and execution) are displayed in relation to one another, as well as where the representational unit is stored.

A.1 GENERAL MODEL OF MOTOR PERFORMANCE

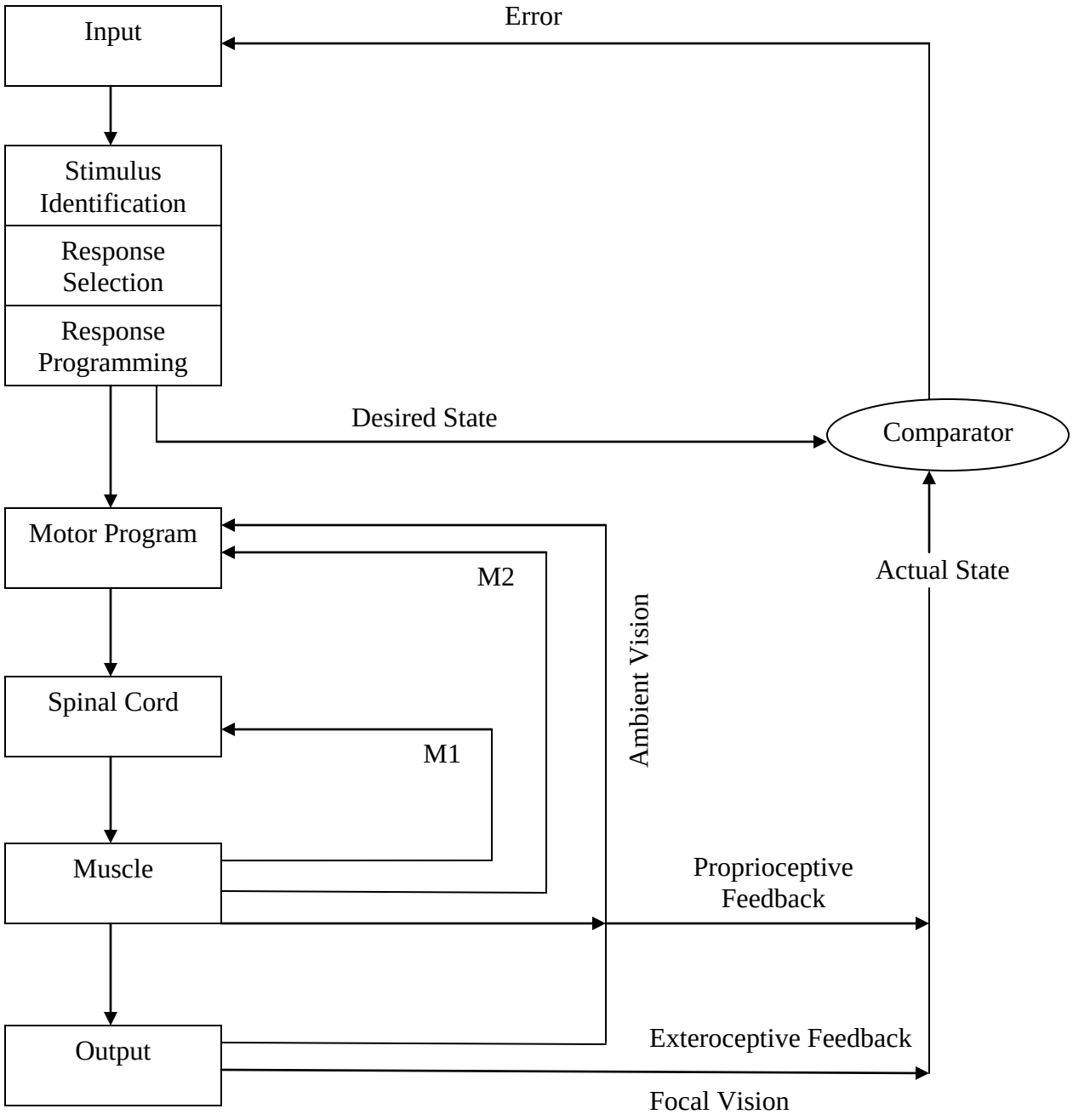

Figure 30: Information-processing model of general motor behavior

Note 14: Adapted, with permission, from R.A. Schmidt and C.A. Wrisberg, 2004, *Motor Learning and performance*, 3rd ed. (Champaign, IL: Human Kinetics), p. 134.

A.2 SPEECH MODEL OF MOTOR PERFORMANCE

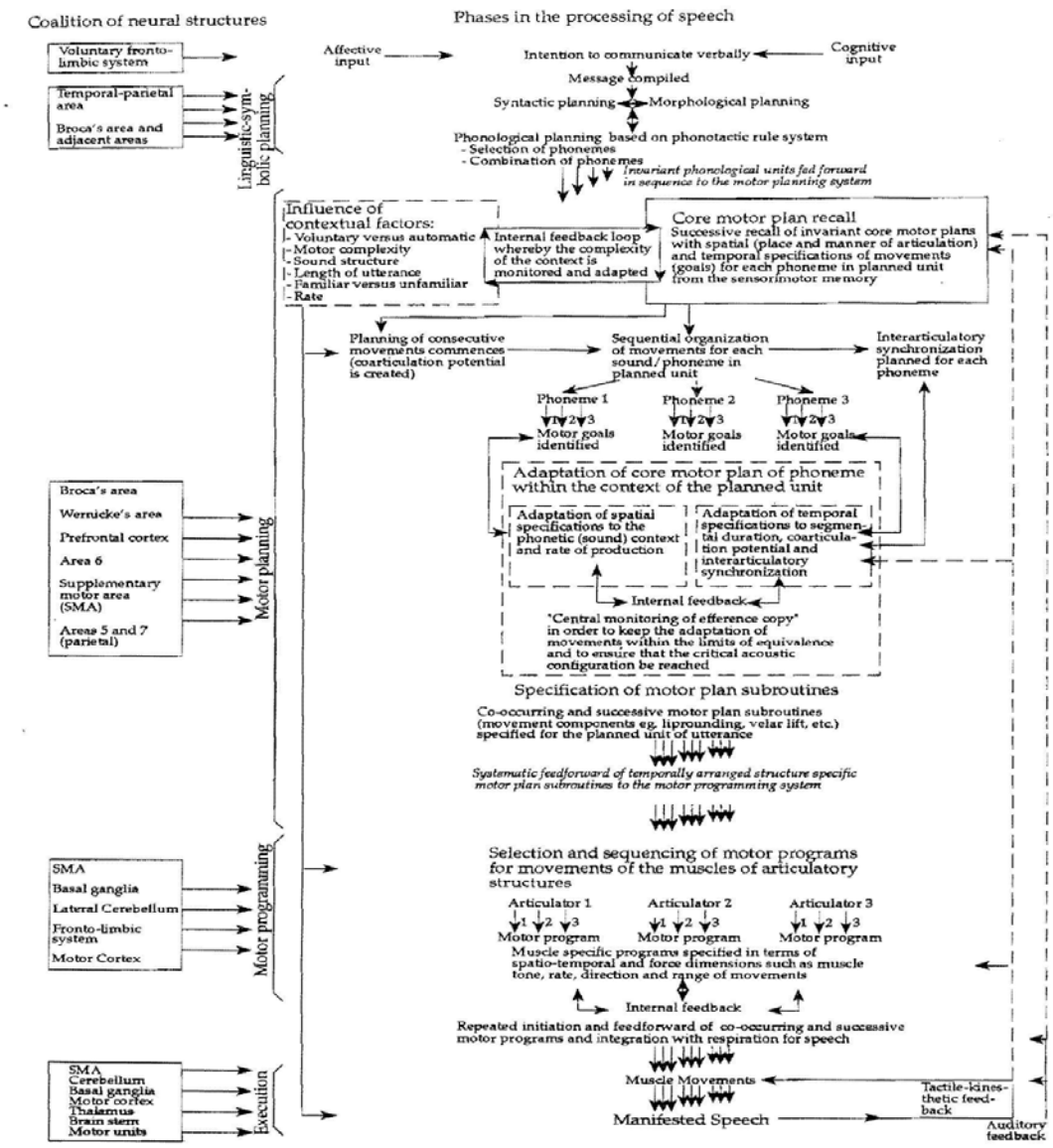

Figure 1–1. Theoretical framework of speech sensorimotor control.

Figure 31: Information-processing model of speech production

Note: 15 Van der Merwe, A., *Clinical management of sensorimotor speech disorders* (Vol. 2), M. R. McNeil (Ed.), A theoretical framework for the characterization of pathological speech sensorimotor control, 1997, Thieme, www.thieme.com (reprinted with permission)

APPENDIX B

LANGUAGE PRE-SCREENING SCRIPT

Hi Mr./Mrs./Ms.:

Thank you for contacting me about my experiment. This study will evaluate how well you can speak a novel word with varying stress patterns. You will be asked to repeat fake words while trying to achieve a particular loudness (or stress) level. You will also be asked to complete a judgment task in which you decide if a fake word is familiar to you or not. This study will take no more than 4.5 hours to complete in a single session. All research will be conducted in at Forbes Tower at the University of Pittsburgh, and you will receive $30.00 for completing this study. In order to see if you are eligible to participate in this experiment, I need to see if you are a native speaker of English with minimal knowledge of other languages. Would you consent to answering a few questions about your language abilities?

IF NO: Alright, thank you for your time. Have a nice day.

IF YES: Great. I will now ask you a few questions:

Language Screening

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

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

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

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

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

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

_____**NO**: Eligible for the study

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

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

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

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

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

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

_____**NO**: Eligible for the study

_____**NO**: Eligible for the study

DOES NOT MEET CRITERIA: I'm sorry, you are ineligible to be in this study due to your knowledge of other languages. Unfortunately this extra language knowledge may bias your responses in this experiment, and cause you to respond faster than would typically be expected. Thank you for your time and interest in my research.

***MEETS CRITERIA*:** You are eligible to be in my study. Let's look at our schedules and have you come into complete the rest of the screenings and experiment.

APPENDIX C

SCREENING FORM: STIMULI VALIDATION STUDY

SS ID: ____

DEMOGRAPHIC INFORMATION

Date of Birth: _____

Age: _____

EDUCATION

Highest Degree Earned:_____High School Degree or Equivalent

_____Undergraduate Degree

_____Graduate or Advanced Degree

Years of Education:_____

LANGUAGE[16]

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

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

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

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

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

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

_____**NO**: Eligible for the study

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

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

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

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

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

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

_____**NO**: Eligible for the study

_____**NO**: Eligible for the study

HEARING – Part I

PTA

Right Ear: 500 Hz_____ 1000 Hz _____ 2000 Hz ____ 4000 Hz _____

Pass at 35dB: Y N

Left Ear: 500 Hz_____ 1000 Hz _____ 2000 Hz ____ 4000 Hz _____

Pass at 35dB: Y N

[16] Adapted from Language Screening Tools used in Dr. Connie Tompkins' lab, University of Pittsburgh

SPEECH PRODUCTION

Oral Mechanism Exam

Facial Symmetry:_____

Lingual protrusion:_____

Mandible Elevation/Depression:_____

Labial Retraction/Protrusion/Closure:_____

Velopharyngeal Symmetry/Movement:_____

Speech Production

DDK/Vowel Prolongation:

Task	Acceptable Range-Male	Acceptable Range-Female	Participant Response	Notes
/p4/	4.6 - 5.4	4.6 - 5.4		
/t4/	4.4 - 5.2	4.4 - 5.2		
/k4/	3.8 - 5	3.8 - 5		
/p4t4k4/	3.3 – 3.9	3.3 – 3.9		
Prolonged /e/	17.1-28.1	11.1 - 19.3		

Note: 16 Values taken from Duffy (1995) pg 84-85

Conversation Observations: _____

T-MAC Screening: _____

HEARING – Part II

NU-6

List 2A Male Speaker	Accuracy Correct (1), Incorrect (0)		Accuracy Correct (1), Incorrect (0)		Accuracy Correct (1), Incorrect (0)
1. Pick		20. Young		34. Bought	
2. Room		21. Ton		35. Turn	
3. Nice		22. Keg		36. Chair	
4. Said		23. Calm		37. Lore	
5. Fail		24. Tool		38. Bite	
6. South		25. Pike		39. Haze	
7. White		21. Ton		40. Match	
8. Keep		22. Mill		41. Learn	
9. Dead		23. Hush		42. Shawl	
10. Loaf		24. Shack		43. Deep	
11. Dab		**25. Reed**		44. Gin	
12. Numb		26. Rot		45. Goal	
13. Juice		27. Hate		46. Far	
14. Chief		28. Live			
15. Merge		29. Book			
16. Wag		30. Voice			
17. Rain		31. Gaze			
18. Witch		32. Pad			
19. Soap		33. Thought			

AUDITORY PROCESSING

Subtest	I	II	III	IV	V	VI	VII	VIII	IX	X
Criterion	**14.75**	**14.59**	**14.27**	**14.29**	**14.05**	**13.78**	**13.93**	**13.84**	**12.56**	**13.04**
PARTICIPANT:										
2nd attempt										
Mean of All Subtests:										
PASS?										
Visual Discrimination Difficulty (self-report):										

APPENDIX D

SCREENING FORM: MAIN STUDY

SS ID: ____

DEMOGRAPHIC INFORMATION

Date of Birth: _____

Age: _____

EDUCATION

Highest Degree Earned:_____High School Degree or Equivalent

_____Undergraduate Degree

_____Graduate or Advanced Degree

Years of Education:_____

HEARING – Part I

PTA

Right Ear: 500 Hz_____ 1000 Hz _____ 2000 Hz ____ 4000 Hz _____

Pass at 35dB: Y N

Left Ear: 500 Hz_____ 1000 Hz _____ 2000 Hz ____ 4000 Hz _____

Pass at 35dB: Y N

SPEECH PRODUCTION

Oral Mechanism Exam

Facial Symmetry:_____

Lingual protrusion:_____

Mandible Elevation/Depression:_____

Labial Retraction/Protrusion/Closure:_____

Velopharyngeal Symmetry/Movement:_____

Speech Production

DDK/Vowel Prolongation:

Task	Acceptable Range-Male	Acceptable Range-Female	Participant Response	Notes
/p4/	4.6 - 5.4	4.6 - 5.4		
/t4/	4.4 - 5.2	4.4 - 5.2		
/k4/	3.8 - 5	3.8 - 5		
/p4t4k4/	3.3 – 3.9	3.3 – 3.9		
Prolonged /e/	17.1-28.1	11.1 - 19.3		

Note: 17 Values taken from Duffy (1995) pg 84-85

Conversation Observations: _____

T-MAC Screening: _____

HEARING – Part II

NU-6

List 2A Male Speaker	Accuracy Correct (1), Incorrect (0)		Accuracy Correct (1), Incorrect (0)		Accuracy Correct (1), Incorrect (0)
1. Pick		20. Young		34. Bought	
2. Room		21. Ton		35. Turn	
3. Nice		22. Keg		36. Chair	
4. Said		23. Calm		37. Lore	
5. Fail		24. Tool		38. Bite	
6. South		25. Pike		39. Haze	
7. White		21. Ton		40. Match	
8. Keep		22. Mill		41. Learn	
9. Dead		23. Hush		42. Shawl	
10. Loaf		24. Shack		43. Deep	
11. Dab		**25. Reed**		44. Gin	
12. Numb		26. Rot		45. Goal	
13. Juice		27. Hate		46. Far	
14. Chief		28. Live			
15. Merge		29. Book			
16. Wag		30. Voice			
17. Rain		31. Gaze			
18. Witch		32. Pad			
19. Soap		33. Thought			

AUDITORY PROCESSING

Subtest	I	II	III	IV	V	VI	VII	VIII	IX	X
Criterion	14.75	14.59	14.27	14.29	14.05	13.78	13.93	13.84	12.56	13.04
PARTICIPANT:										
2nd attempt										
Mean of All Subtests:										
PASS?										
Visual Discrimination Difficulty (self-report):										

MANUAL DEXTERITY

Reaction time test: ____/_5_ (OLD, NEW, OLD, OLD, NEW)

APPENDIX E

DATA FOR PREDICTED ACCURACY RESULTS

A highly conservative estimate of participant performance in responding to the SPSC stimuli was presented in Tables 53 and 54. Although the judgment task was estimated to yield high accuracy scores, a range of hit rates was estimated for overall performance on this task. A high "hit" rate was estimated when participants accurately classified "old" targets 29/30 times (29 "hits," Table 53), whereas a low "hit" rate was estimated when participants' accuracy was 25/30 (Table 54). All "new" participant responses were estimated the same for both high and low hit rates. DPSC and DPDC were estimated to be identified as "new" 100% of the time (i.e., these stimuli were classified as "correct rejections"). It is unclear how participants would respond to SPSC stimuli given the stimuli similarity to the training items. Thus, it was estimated all 10 SPSC stimuli would be erroneously identified as "old," which is why these 10 stimuli were classified as "misses" in the tables below.

Table 53: High hit rate

		PARTICIPANT RESPONSE	
		"OLD"	**"NEW"**
TARGET	**"OLD"**	29 (Hits)	10 (Misses)
	"NEW"	1 (False Alarms)	20 (Correct Rejections)
		Total = 30	Total = 30

Hit Proportion: 96.67%; False Alarm Proportion: 33.33%; Percent Correct: 81.67%; *d'*: 2.265
Calculated using the Online Detection Theory Calculator (Legatt, 2011) at the following website: http://www.computerpsych.com/Research_Software/NormDist/Online/Detection_Theory

Table 54: Low hit rate

		PARTICIPANT RESPONSE	
		"OLD"	**"NEW"**
TARGET	**"OLD"**	25 (Hits)	10 (Misses)
	"NEW"	5 (False Alarms)	20 (Correct Rejections)
		Total = 30	Total = 30

Hit Proportion: 83.33%; False Alarm Proportion: 33.33%; Percent Correct: 75%; *d'*: 1.398
Calculated using the Online Detection Theory Calculator (Legatt, 2011) at the following website: http://www.computerpsych.com/Research_Software/NormDist/Online/Detection_Theory.

APPENDIX F

POST-HOC ANALYSIS OF PRIMARY OUTCOMES

The inaccurate responses on the old-new judgment task (see Table 34) were removed prior to reanalysis. Analysis and results for RQ4 and RQ5 are detailed separately below.

RQ4: Phonetic Similarity

RQ4 evaluated the effect of phonetic similarity between the trained and untrained stimuli varying by phonetic similarity only (SPSC and DPSC stimuli). The parameter of motor class was the same for these three stimuli types. The data did not meet the normality assumptions required for one-way ANOVA with repeated measurements; thus, Friedman's Test was conducted as a non-parametric alternative to the one-way ANOVA with repeated measures.

There was a significant difference in reaction times across stimuli type, $\chi^2(2) = 18.58$, $p \leq .000$. Pairwise comparisons were performed (SPSS, 2012) with Bonferroni correction for multiple comparisons. The reaction times for the SPSC stimuli were significantly slower than the Trained ($p = .003$) and DPSC ($p \leq .000$) stimuli (see Figure 32). There were no other significant differences *ps* > .99. The median reaction times for the Trained, SPSC, and DPSC stimuli sets are listed in Table 55.

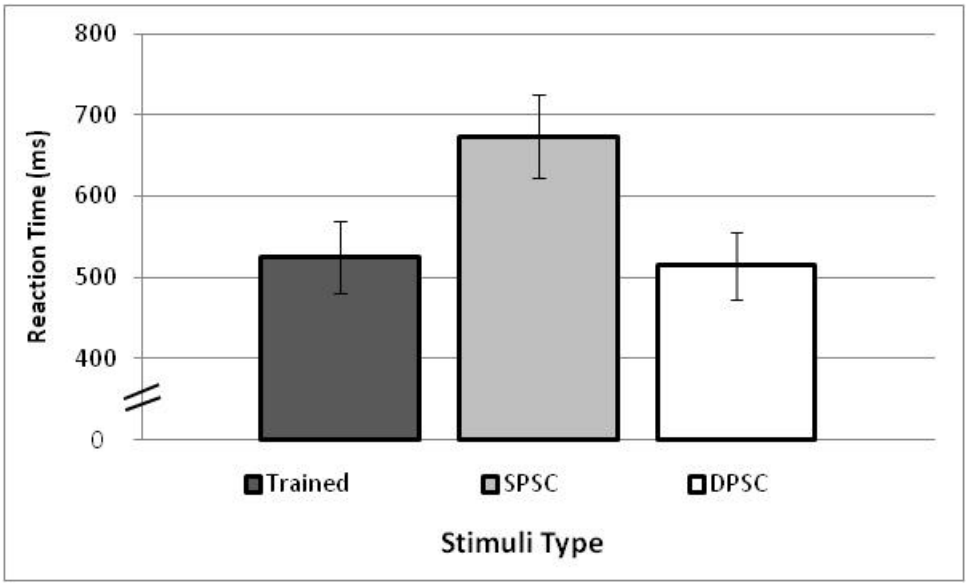

Figure 32: Mean reaction times (ms) and standard errors of Trained, SPSC, and DPSC stimuli (inaccurate responses removed)

Table 55: Median reaction times of Trained, SPSC, and DPSC stimuli

Stimuli Type	Median (ms)
Trained	520.65
SPSC	607.85
DPSC	454.39

RQ5: Motor Class

RQ5 evaluated the effect of motor class between the DPSC and DPDC stimuli, in which the parameter of phonetic similarity was held constant. The data did not meet the normality assumptions required for paired samples t-test; thus, Wilcoxon signed-rank test was conducted as a non-parametric alternative to the paired samples t-test. Participants' reaction times were

significantly slower when responding to the DPSC stimuli (*Mdn* = 454.39ms) compared to the DPDC (*Mdn* = 407.14ms), $z = -2.57$, $p = .01$; see Figure 33.

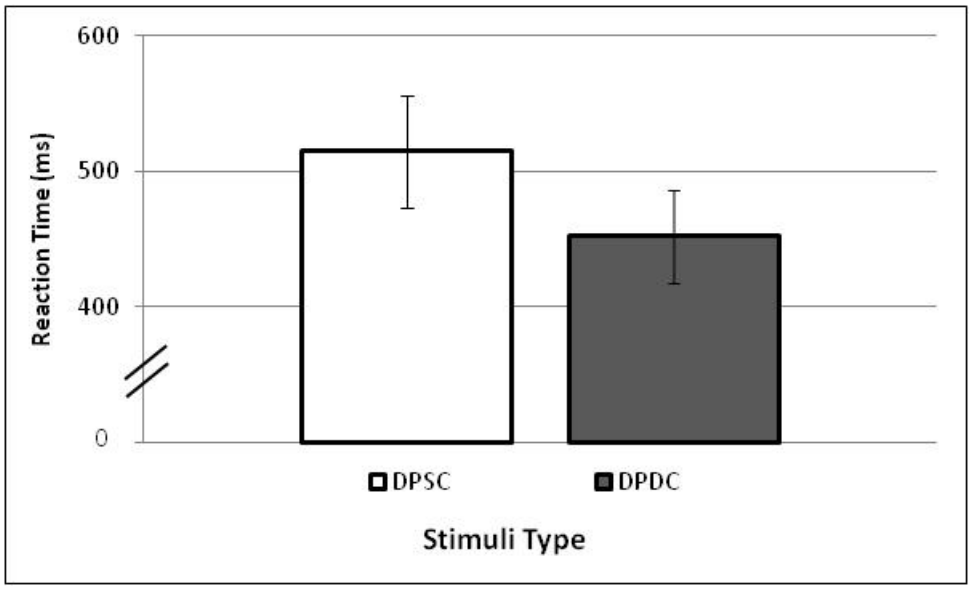

Figure 33: Mean reaction times (ms) and standard errors of stimuli varying by motor class

BIBLIOGRAPHY

Adams, C., & Munro, R. R. (1978). In search of the acoustic correlates of stress: fundamental frequency, amplitude, and duration in the connected utterance of some native and non-native speakers of English. *Phonetica*, *35*(3), 125–156.

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74.

Adams, S. G., 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.

Aha, D. W. (1997). Lazy learning: Special issue editorial. *Artificial Intelligence Review*, *11*, 7–10.

Aha, D. W., Kibler, D., & Albert, M. K. (1991). Instance-based learning algorithms. *Machine Learning*, *6*(1), 37–66.

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88*(1), 148–159.

Almelaifi, R. (2013). *The role of "focus of attention" on the learning of non-native speech sounds: English speakers learning of mandarin Chinese tones* (Dissertation). University of Pittsburgh, Pittsburgh. PA. Retrieved from http://d-scholarship.pitt.edu/16937/

Altmann, G., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 899.

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

Anderson, J. R. (1980). *Cognitive Psychology and Its Implications.* San Francisco: W.H. Freeman and Company.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, *94*, 192–210.

Anderson, J. R. (1992). Automaticity and the ACT theory. *The American Journal of Psychology*, *105*(2), 165–160.

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, *48*(1), 35–44.

Anderson, J. R. (1995). *Learning and memory: An integrated approach.* New York: John Wiley & Sons, Inc.

Anderson, J. R., Bothell, D., Bryne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. *Handbook of Reading Research*, *1*, 255–291.

Anson, J. G. (1982). Memory drum theory: Alternative tests and explanations for the complexity effects on simple reaction time. *Journal of Motor Behavior*, *14*(3), 228.

Arvedson, J. C., McNeil, M. R., & West, T. L. (1985). Prediction of Revised Token Test overall, subtest, and linguistic unit scores by two shortened versions. In Clinical Aphasiology Conference: Clinical Aphasiology Conference (1985 : 15th : Ashland, OR : June **2**-6).

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 33–53. doi:10.1037/0278-7393.14.1.33

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149–178.

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.

Ballard, K. J., Granier, J. P., & Robin, D. A. (2000). Understanding the nature of apraxia of speech: Theory, analysis, and treatment. *Aphasiology*, *14*(10), 969–995.

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

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

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*(1), 89–106.

Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Skrull & J. R. S. Wyer (Eds.), *Advances in Social Cognition* (Vol. 3). Hillsdale, NJ: Lawrence Erlbaum Assoc. Publishers.

Bartlett, F. C. (1932). *Remembering*. Cambridge, MA: Cambridge University Press.

Beek, P. J. (2000). Toward a theory of implicit learning in the perceptual-motor domain. *International Journal of Sport Psychology*, *31*, 547–554.

Benjamin, B. (1997). Speech Production of Normally Aging Adults. *Seminars in Speech and Language*, *18*(02), 135–141. doi:10.1055/s-2008-1064068

Bernstein, N. (1967). *The co-ordination and regulation of movements*. Oxford: Pergamon Press.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345–360. doi:10.1037/0096-1523.14.3.345

Best, C. T., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, *29*(3), 191–211. doi:10.3758/BF03207286

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3.53). Retrieved from www.praat.org

Boronat, C. B., & Logan, G. D. (1997). The role of attention in automatization: Does attention operate at encoding, or retrieval, or both? *Memory & Cognition*, *25*(1), 36–46. doi:10.3758/BF03197283

Bowers, J. S., Mattys, S. L., & Gage, S. H. (2009). Preserved implicit knowledge of a forgotten childhood language. *Psychological Science*, *20*(9), 1064–1069.

Bowey, J. A., & Hirakis, E. (2006). Testing the protracted lexical restructuring hypothesis: The effects of position and acoustic–phonetic clarity on sensitivity to mispronunciations in children and adults. *Journal of Experimental Child Psychology*, *95*(1), 1–17.

Brainard, R. W., Irby, T. S., Fitts, P. M., & Alluisi, E. A. (1962). Some variables influencing the rate of gain of information. *Journal of Experimental Psychology*, *63*, 105.

Bright, J. E. H., & Burton, A. M. (1994). Past midnight: Semantic processing in an implicit learning task. *The Quarterly Journal of Experimental Psychology Section A*, *47*(1), 71–89. doi:10.1080/14640749408401144

Broadbent, D. (1985). A question of levels: Comment on McClelland and Rumelhart. *Journal of Experimental Psychology. General*, *114*(2), 189–192.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. *Cognition and Categorization*, *3*, 170–211.

Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual fators in categorization.* (pp. 141–174). London/New York: Cambridge University Press.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al.(1989). *Journal of Experimental Psychology. General*, *120*(3), 316–323.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711–733.

Canic, M. J., & Franks, I. M. (1989). Response preparation and latency in patterns of tapping movements. *Human Movement Science*, *8*(2), 123–139. doi:10.1016/0167-9457(89)90013-4

Carbonnell, L., Ramdani, C., Meckler, C., Burle, B., Hasbroucq, T., & Vidal, F. (2013). The N-40: An electrophysiological marker of response selection. *Biological Psychology*, *93*(1), 231–236.

Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 517–526.

Carlton, L. G., Carlton, M. J., & Newell, K. M. (1987). Reaction time and response dynamics. *The Quarterly Journal of Experimental Psychology*, *39*(2), 337–360.

Caruso, A. J., Abbs, J. H., & Gracco, V. L. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain*, *111*(2), 439–455.

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

Chamberlin, C. J., & Magill, R. A. (1992b). The memory representation of motor skills: A test of schema theory. *Journal of Motor Behavior*, *24*(2), 309–319.

Chandler, C. C. (1991). How memory for an event is influenced by related events: Interference in modified recognition tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 115–125. doi:http://dx.doi.org/10.1037/0278-7393.17.1.115

Chandler, C. C. (1993). Accessing related events increases retroactive interference in a matching recognition test. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 967.

Cheng, H. Y., Murdoch, B. E., Goozee, J. V., & Scott, D. (2007). Physiologic development of tongue–jaw coordination from childhood to adulthood. *Journal of Speech, Language, and Hearing Research*, *50*(2), 352–360.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*(5), 351–353. doi:10.1038/1561

Cheour, M., H.T. Leppänen, P., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology*, *111*(1), 4–16. doi:10.1016/S1388-2457(99)00191-1

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

Cholin, J., Levelt, W. J., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*(2), 205–235.

Cholin, J., Schiller, N. O., & Levelt, W. J. (2004). The preparation of syllables in speech production. *Journal of Memory and Language*, *50*(1), 47–61.

Chomsky, N. (1963). Formal properties of grammars. Wiley: Hoboken, NJ.

Chomsky, N., & Halle, M. (1968). The sound pattern of English. Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED020511

Christina, R. W., & Rose, D. J. (1985). Premotor and motor reaction time as a function of response complexity. *Research Quarterly for Exercise and Sport*, *56*(4), 306–315.

Clark, H. M. (2003). Neuromuscular treatments for speech and swallowing: a tutorial. *American Journal of Speech-Language Pathology*, *12*(4), 400.

Clark, H. M., & Robin, D. A. (1998). Generalized motor programme and parameterization accuracy in apraxia of speech and conduction aphasia. *Aphasiology*, *12*(7/8), 699–713.

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

Cock, J. J., Berry, D. C., & Gaffan, E. A. (1994). New strings for old: The role of similarity processing in an incidental learning task. *The Quarterly Journal of Experimental Psychology Section A*, *47*(4), 1015–1034. doi:10.1080/14640749408401105

Cohen, A. L., & Nosofsky, R. M. (2003). An extension of the exemplar-based random-walk model to separable-dimension stimuli. *Journal of Mathematical Psychology*, *47*(2), 150–165.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Coker, C. A. (2004). *Motor learning and control for practioners*. Boston, MA: McGraw Hill.

Cooper, R., & Shallice, T. (2000). Contention Scheduling and the Control of Routine Activities. *Cognitive Neuropsychology*, *17*(4), 297–338.

Corwin, J. (1994). On measuring discrimination and response bias: Unequal numbers of targets and distractors and two classes of distractors. *Neuropsychology*, *8*(1), 110–117. doi:http://dx.doi.org.pitt.idm.oclc.org/10.1037/0894-4105.8.1.110

Cost, S., & Salzberg, S. (1993). A weighted nearest neighbor algorithm for learning with symbolic features. *Machine Learning*, *10*(1), 57–78.

Cotton, S., & Grosjean, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception & Psychophysics*, *35*(1), 41–48. doi:10.3758/BF03205923

Craik, F. I., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 474–479. doi:http://dx.doi.org.pitt.idm.oclc.org/10.1037/0278-7393.13.3.474

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

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

Curtis, C. E., Rao, V. Y., & D'Esposito, M. (2004). Maintenance of spatial and motor codes during oculomotor delayed response tasks. *The Journal of Neuroscience*, *24*(16), 3944–3952.

Cycowicz, Y. M., Friedman, D., Snodgrass, J. G., & Duff, M. (2001). Recognition and source memory for pictures in children and adults. *Neuropsychologia*, *39*(3), 255–267. doi:10.1016/S0028-3932(00)00108-1

Davis, A. C. (1990). Epidemiological profile of hearing impairments: the scale and nature of the problem with special reference to the elderly. *Acta Oto-Laryngologica, Supplement*, *476*, 23–31.

Davis, B. L., MacNeilage, P. F., & Matyear, C. L. (2002). Acquisition of Serial Complexity in Speech Production: A Comparison of Phonetic and Phonological Approaches to First Word Production. *Phonetica*, *59*(2-3), 75–107. doi:10.1159/000066065

Deger, K., & Ziegler, W. (2002). Speech motor programming in apraxia of speech. *Journal of Phonetics*, *30*(3), 321–335. doi:10.1006/jpho.2001.0163

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, *8*(4), 919–924.

Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, *370*(6487), 292–295. doi:10.1038/370292a0

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, *21*(1), 62.

Dienes, Z., Broadbent, D., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 875.

Dollaghan, C. (1998). Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics*, *19*(02), 193–207.

Dollaghan, C., & Campbell, T. F. (1998). Nonword Repetition and Child Language Impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136–1146.

Domingos, P. (1996). Unifying instance-based and rule-based induction. *Machine Learning*, *24*(2), 141–168.

Doody, S. G., & Zelaznik, H. N. (1988). Rule formation in a rapid-timing task: A test of schema theory. *Research Quarterly for Exercise and Sport*, *59*(1), 21–28.

Dopkins, S., & Gleason, T. (1997). Comparing exemplar and prototype models of categorization. *Canadian Journal of Experimental Psychology*, *51*(3), 212–230.

Duffy, J. R. (1995). *Motor speech disorders: substrates, differential diagnosis, and management*. Mosby.

Eisner, F., & McQueen, J. M. (2005). The Specificity of Perceptual Learning in Speech Processing. *Perception & Psychophysics*, *67*(2), 224–238.

Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A*, *46*(3), 533–558. doi:10.1080/14640749308401062

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Fitch, J., & Holbrook, A. (1970). Modal vocal fundamental frequency of young adults. *Archives of Otolaryngology*, *92*(4), 379–382. doi:10.1001/archotol.1970.04310040067012

Fitts, P. M., & Seeger, C. M. (1953). SR compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*(3), 199–210.

Flexser, A. J., & Bower, G. H. (1974). How frequency affects recency judgments: A model for recency discrimination. *Journal of Experimental Psychology*, *103*(4), 706.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*(3), 253–292.

Forrest, K., & Weismer, G. (1997). Acoustic analysis of motor speech disorders. In M. R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders*. New York: Thieme.

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.

Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age Differences and Changes in Reaction Time: The Baltimore Longitudinal Study of Aging. *Journal of Gerontology*, *49*(4), P179–P189. doi:10.1093/geronj/49.4.P179

Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends in Cognitive Sciences*, *9*(10), 481–488. doi:10.1016/j.tics.2005.08.008

Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. [Miscellaneous Article]. *Neuroreport July 19, 2002*, *13*(10), 1251–1254.

Fromkin, V. (1980). *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*. Academic Press.

Fry, D. B. (1958). Experiments in the Perception of Stress. *Language and Speech*, *1*(2), 126–152. doi:10.1177/002383095800100207

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474–496.

Galotti, K. M. (1994). Cognitive psychology in and out of the laboratory. Pacific Grove: Brooks/Cole Publishing Company.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), . London, England: Academic Press.

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet*, *366*(9491), 1111–1120. doi:10.1016/S0140-6736(05)67423-5

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 84–95. doi:10.1037/0278-7393.25.1.84

Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, *21*(4), 347–353.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.

Gentner, D. R., Larochelle, S., & Grudin, J. (1988). Lexical, sublexical, and peripheral effects in skilled typewriting. *Cognitive Psychology*, *20*(4), 524–548.

Goffman, L., & Smith, A. (1999). Development and phonetic differentiation of speech movement patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(3), 649–660. doi:http://dx.doi.org/10.1037/0096-1523.25.3.649

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501–518.

Gonzalez, C. (2012). Training decisions from experience with decision-making games. In P. J. Durlach & A. M. Lesgold (Eds.), *Adaptive technologies for training and education*. Cambridge, MA: Cambridge University Press.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523–551. doi:10.1037/a0024558

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-Based Learning in Dynamic Decision Making. *Cognitive Science*, *27*, 591–635.

Gordon, P. C., & Meyer, D. E. (1984). Perceptual-motor processing of phonetic features in speech. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(2), 153.

Gorham-Rowan, M. M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders*, *39*(3), 171–184. doi:10.1016/j.jcomdis.2005.11.005

Gracco, V. L. (1990). Characteristics of speech as a motor control system. In G. R. Hammond (Ed.), *Cerebral Control of Speech and Limb Movements*. North Holland, Amsterdam: Elsevier.

Gracco, V. L. (1991). *Sensorimotor mechanisms in speech motor control*. North Holland, Amsterdam: Elsevier.

Gracco, V. L. (1994). Some organizational characteristics of speech movement control. *Journal of Speech & Hearing Research*, *37*(1), 4–27.

Gracco, V. L., & Abbs, J. H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research*, *65*(1), 156–166.

Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 981–998.

Green, J. R. (2002). The Sequential Development of Jaw and Lip Control for Speech. *Journal of Speech, Language, and Hearing Research*, *45*(1), 66–79. doi:10.1044/1092-4388(2002/005)

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The Physiologic Development of Speech Motor Control: Lip and Jaw Coordination. *Journal of Speech, Language, and Hearing Research : JSLHR*, *43*(1), 239–255.

Green, J. R., & Nip, I. S. B. (2010). Some organization principles in early speech development. In B. Maassen & P. van Lieshout (Eds.), *Speech Motor Control: New Developments in Basic and Applied Research*. Oxford University Press.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*(4), 267–283.

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics*, *72*, 43–53.

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., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, *105*, 611–633.

Guion, S. G., Clark, J. J., Harada, T., & Wayland, R. P. (2003). Factors Affecting Stress Placement for English Nonwords include Syllabic Structure, Lexical Class, and Stress Patterns of Phonologically Similar Words. *Language and Speech*, *46*(4), 403–426. doi:10.1177/00238309030460040301

Hall, J. F. (1989). *Learning and Memory* (2nd ed.). Boston, MA: Allyn and Bacon.

Harrington, J., & Cassidy, S. (1999). *Techniques in Speech Acoustics*. Springer.

Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry*, 227–276.

Helie, S., & Ashby, F. G. (2009). A Neurocomputational Model of Automaticity and Maintenance of Abstract Rules. *Proceedings of the International Joint Conference on Neural Networks*, 1192–1198.

Helie, S., Waldschmidt, J. G., & Ashby, F. G. (2010). Automaticity in Rule-Based and Information-Integration Categorization. *Attention, Perception, & Psychophysics*, *72*(4), 1013–1031.

Henry, F. M. (1961). Reaction time-movement time correlations. *Perceptual and Motor Skills*, *12*(1), 63–66.

Henry, F. M., & Rogers, D. E. (1960). Increased response latency for complicated movements and a“ memory drum” theory of neuromotor reaction. *Research Quarterly of the American Association for Health, Physical Education, & Recreation*. Retrieved from http://psycnet.apa.org/?fa=main.doiLanding&uid=1962-00243-001

Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition*, *37*(1), 133–166.

Heuer, H. (1982). Binary choice reaction time as a criterion of motor equivalence. *Acta Psychologica*, *50*(1), 35–47.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), 11–26.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research Theory* (Vol. 10). New York, NY: Academic Press.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*(2), 96–101.

Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428. doi:http://dx.doi.org/10.1037/0033-295X.93.4.411

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*(4), 528–551. doi:http://dx.doi.org/10.1037/0033-295X.95.4.528

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition*, *38*(1), 102–115. doi:10.3758/MC.38.1.102

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, *88*(3), 297.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*(3), 295–355.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(6), 418–439.

Honbolygó, F., & Csépe, V. (2013). Saliency or template? ERP evidence for long-term representation of word stress. *International Journal of Psychophysiology*, *87*(2), 165–172. doi:10.1016/j.ijpsycho.2012.12.005

Huang, Q., & Tang, J. (2010). Age-related hearing loss or presbycusis. *European Archives of Oto-Rhino-Laryngology*, *267*(8), 1179–1191. doi:10.1007/s00405-010-1270-7

Hulstijn, W., & Van Galen, G. P. (1983). Programming in handwriting: Reaction time and movement time as a function of sequence length. *Acta Psychologica*, *54*(1), 23–49.

Hunt, R. R., & Ellis, H. C. (2004). *Fundamentals of Cognitive Psychology*. (7th ed.). New York: McGraw-Hill Higher Education.

Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H.-S. (2009). Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition*, *12*(02), 121–140.

Hyman, R. (1953). *Stimulus information as a determinant of reaction time*. Johns Hopkins University. Retrieved from http://www.psychology.uiowa.edu/faculty/mordkoff/InfoProc/pdfs/Hyman%201953.pdf

IPA Chart. (2005). International Phonetic Association. Retrieved from http://www.langsci.ucl.ac.uk/ipa/ipachart.html

Isenberg, D., & Gay, T. (1978). Acoustic correlates of perceived stress in an isolated synthetic disyllable. *The Journal of the Acoustical Society of America*, *64*(S1), S21–S21.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 649–667.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research Theory* (Vol. 18). New York, NY: Academic Press.

Jacoby, L. L., Marriot, M. J., & Collins, J. G. (1990). The specifics of memory and cognition. In T. K. Skrull & J. R. S. Wyer, (Eds.), *Advances in Social Cognition* (Vol. 3). Hillsdale, NJ: Lawrence Erlbaum Assoc. Publishers.

Jahnke, J. C., & Duncan, C. P. (1956). Reminiscence and forgetting in motor learning after extended rest intervals. *Journal of Experimental Psychology*, *52*(5), 273–282. doi:http://dx.doi.org/10.1037/h0043394

James, W. (1890). *Principles of psychology*. New York, NY: Holt.

Jamieson, R. K., & Mewhort, D. J. K. (2009a). Applying an exemplar model to the artificial-grammar task: Inferring grammaticality from similarity. *The Quarterly Journal of Experimental Psychology*, *62*(3), 550–575.

Jamieson, R. K., & Mewhort, D. J. K. (2009b). Applying an exemplar model to the serial reaction-time task: Anticipating from experience. *The Quarterly Journal of Experimental Psychology*, *62*(9), 1757–1783.

Jamieson, R. K., & Mewhort, D. J. K. (2010). Applying an exemplar model to the artificial-grammar task: String completion and performance on individual items. *The Quarterly Journal of Experimental Psychology*, *63*(5), 1014–1039.

Jevas, S., & Yan, J. H. (2001). The effect of aging on cognitive function: a preliminary quantitative review. *Research Quarterly for Exercise and Sport*, *72*, 49.

Johansson, T. (2009). Strengthening the case for stimulus-specificity in artificial grammar learning: No evidence for abstract representations with extended exposure. *Experimental Psychology*, *56*(3), 188–197. doi:http://dx.doi.org/10.1027/1618-3169.56.3.188

Johns, E. E., & Mewhort, D. J. K. (2002). What information underlies correct rejections in short-term recognition memory? *Memory & Cognition*, *30*(1), 46–59.

Johns, E. E., & Mewhort, D. J. K. (2003). The effect of feature frequency on short-term recognition memory. *Memory & Cognition*, *31*(2), 285–296.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136–153. doi:http://dx.doi.org/10.1037/0033-295X.93.2.136

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=_0H8gwj4a1MC&oi=fnd&pg=PR8&dq=heuristics&ots=YChd6SW3YL&sig=DZuJGuqicZWYFp_Ktpyaji7FnCk

Kang, S. (2013). Neural basis of the Word Frequency Effect and Its Relation to Lexical Processing. Retrieved from http://www.linguistics.berkeley.edu/PhonLab/annual_report/documents/2012/Kang_word_freq.pdf

Kantak, S. S., & Winstein, C. J. (2012). Learning–performance distinction and memory processes for motor skills: A focused review and perspective. *Behavioural Brain Research*, *228*(1), 219–231.

Kaushanskaya, M., Yoo, J., & Van Hecke, S. (2013). Word learning in adults with second-language experience: effects of phonological and referent familiarity. *Journal of Speech, Language, and Hearing Research : JSLHR*, *56*(2), 667.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, *70*(6, Pt.1), 387–403. doi:http://dx.doi.org/10.1037/h0026739

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 & Performance*, *31*(5), 970–978.

Kendall, D., McNeil, M. R., Shaiman, S., & Pratt, S. (2005). Phonetic encoding of infrequent articulatory phonetic transitions. *Aphasiology*, *19*(1), 39–52. doi:10.1080/02687030444000606

Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders*, *52*(4), 367.

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*. Singular/Thomson Learning.

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical and Statistical Psychology*, *54*(1), 1–20.

Khoe, W., Kroll, N. E. ., Yonelinas, A. P., Dobbins, I. G., & Knight, R. T. (2000). The contribution of recollection and familiarity to yes–no and forced-choice recognition tests in healthy subjects and amnesics. *Neuropsychologia*, *38*(10), 1333–1341. doi:10.1016/S0028-3932(00)00055-5

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.

Klapp, S. T. (1995). Motor response programming during simple and choice reaction time: the role of practice. *Journal of Experimental Psychology. Human Perception & Performance*, *21*(5), 1015–1027.

Klapp, S. T. (1996). Reaction time analysis of central motor control. In H. N. Zelaznik (Ed.), *Motor Learning and Control* (pp. 13–36). Champagne, IL: Human Kinetics.

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

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, *59*(5), 1208–1221. doi:10.1121/1.380986

Kleinman, M. (1983). *The acquisition of motor skill*. Princeton Book Co.

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

Kornblum, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review*, *76*(2), 113.

Kornblum, S. (1973). Sequential effects in choice reaction time: A tutorial review. *Attention and Performance IV*, 259–288.

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93–107. doi:10.3758/BF03212211

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 11850–11857.

Laganaro, M. (2005). Syllable frequency effect in speech production: Evidence from aphasia. *Journal of Neurolinguistics*, *18*(3), 221–235.

Laganaro, M. (2008). Is there a syllable frequency effect in aphasia or in apraxia of speech or both? *Aphasiology*, *22*(11), 1191–1200. doi:10.1080/02687030701820469

Lai, Q., & Shea, C. H. (1999). Bandwidth knowledge of results enhances generalized motor program learning. *Research Quarterly for Exercise and Sport*, *70*(1), 79+.

Lai, Q., Shea, C. H., Wulf, G., & Wright, D. L. (2000). Optimizing generalized motor program and parameter learning. *Research Quarterly for Exercise & Sport*, *71*(1), 10–24.

Lamberts, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General*, *124*(2), 161.

Lamberts, K. (1998). The time course of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 695.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*(2), 227.

Lamberts, K., Brockdorff, N., & Heit, E. (2003). Feature-sampling and random-walk models of individual-stimulus recognition. *Journal of Experimental Psychology: General*, *132*(3), 351.

Landauer, T. K. (1975). Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology*, *7*(4), 495–531.

Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990). Teaching the Use of Cost-Benefit Reasoning in Everyday Life. *Psychological Science*, *1*(6), 362–370. doi:10.1111/j.1467-9280.1990.tb00243.x

Lassaline, M. E., & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology*, *19*(3), 561.

Lee, T. D. (1988). Transfer-appropriate processing: A framework for conceptualizing practice effects. In O. G. Meijer & K. Roth (Eds.), *Complex movement behaviour: The motor-action controversy* (pp. 201–215). North Holland, Amsterdam: Elsevier Science Publishers B.V.

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In R. B. Wilberg & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control* (pp. 3–22). Amsterdam: North Holland.

Lee, T. D., & Swinnen, S. (1993). Three legacies of Bryan and Harter: Automaticity, variability and change in skilled performance. *Advances in Psychology*, *102*, 295–315.

Legatt, M. E. (2011). *Online Detection Theory Calculator*. ComputerPsych LLC.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/bdm.722/full

Leppänen, P. H. T., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., & Lyytinen, H. (2002). Brain Responses to Changes in Speech Sound Durations Differ Between Infants With and Without Familial Risk for Dyslexia. *Developmental Neuropsychology*, *22*(1), 407–422. doi:10.1207/S15326942dn2201_4

Levelt, W. J., Roelofs, A., & Meyers, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

Li, Y., & Nara, S. (2012). Solving Complex Control Tasks via Simple Rule(s): Using Chaotic Dynamics in a Recurrent Neural Network Model. In A. R. Rao & G. A. Cecchi (Eds.), *The Relevance of the Time Domain to Neural Network Models* (pp. 159–178). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4614-0724-9_9

Lieberman, D. (1990). *Learning: Behavior and Cognition.* Belmont: Wadsworth Publishing Company.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, *32*(4), 451–454.

Linville, S. E., & Fisher, H. B. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *The Journal of the Acoustical Society of America*, *78*(1), 40–48. doi:10.1121/1.392452

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology*, *39*(2), 367.

Logan, G. D. (1988). Toward an instance theory of automization. *Psychological Review*, *95*(4), 492–527.

Logan, G. D. (1995). The Weibull distribution, the power law, and the instance theory of automaticity. *Psychological Review*, *102*(4), 751–756. doi:10.1037/0033-295X.102.4.751

Logan, G. D. (2005). Attention, automaticity, and executive Control: Experimental cognitive psychology and its applications. In G. D. Logan & A. F. Healy (Eds.), *Experimental cognitive psychology and its applications, Decade of behavior* (Vol. xxii, pp. 129–139). Washington: American Psychological Association.

Logan, G. D., & Etherton, J. L. (1994). What Is Learned During Automatization?: The Role of Attention in Constructing an Instance. *Journal of Experimental Psychology: Learning, Memory and Cognition [PsycARTICLES]*, *20*(5), 1022. doi:http://dx.doi.org/10.1037/0278-7393.20.5.1022

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1.

Luchies, C. W., Schiffman, J., Richards, L. G., Thompson, M. R., Bazuin, D., & DeYoung, A. J. (2002). Effects of age, step direction, and reaction condition on the ability to step quickly. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *57*(4), M246–M249.

Ludlow, C. L., Connor, N. P., & Bassich, C. J. (1987). Speech timing in Parkinson's and Huntington's disease. *Brain and Language*, *32*(2), 195–214.

Maas, E., & Mailend, M.-L. (2012). Speech Planning Happens Before Speech Execution: Online Reaction Time Methods in the Study of Apraxia of Speech. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1523–1534. doi:10.1044/1092-4388(2012/11-0311)

Maas, E., Robin, D. A., Austermann-Hula, S. N., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, *17*, 277–298.

Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain and Language*, *106*(2), 107–118. doi:10.1016/j.bandl.2008.03.004

Macmillan, N. A. (1991). *Detection Theory: A User's Guide*. CUP Archive.

MacNeilage, P. F., & Davis, B. L. (1990). Acquisition of Speech Production: The Achievement of Segmental Independence. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 55–68). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-009-2037-8_3

Magill, R. A. (1998). Knowledge is more than we can talk about: Implicit learning in motor skill acquisition. *Research Quarterly for Exercise & Sport*, *69*(2), 104–111.

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

Magnuson, C., & Wright, D. L. (2004). Random practice can facilitate the learning of tasks that have different relative time structures. *Research Quarterly for Exercise & Sport*, *75*(2), 197.

Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2008). Lexical representations in children with SLI: Evidence from a frequency-manipulated gating task. *Journal of Speech, Language, and Hearing Research*, *51*(2), 381–393.

Malt, B. C. (1989). An On-Line Investigation of Prototype and Exemplar Strategies in Classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 539–555.

Marshall, S. P. (1995). *Schemas in Problem Solving*. Cambridge University Press.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, *79*(6), 487–509. doi:http://dx.doi.org/10.1037/h0033467

Masson, M. E. J. (1986). Identification of typographically transformed words: Instance-based skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 479–488. doi:http://dx.doi.org/10.1037/0278-7393.12.4.479

Masson, M. E. J. (1990). Cognitive theories of skill acquisition. *Human Movement Science*, *9*, 221–239.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1083–1100. doi:http://dx.doi.org/10.1037/0278-7393.15.6.1083

McAndrews, M. P., & Moscovitch, M. (1985). Rule-based and exemplar-based classification in artificial grammar learning. *Memory & Cognition*, *13*(5), 469–475.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 471.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*(2), 159–188.

McNeil, M. R., Pratt, S. R., Szuminsky, N., Sung, J. E., Fossett, T. R. D., Fassbinder, W., & Lim, K. Y. (In Submission). Description and psychometric development of the Computerized Revised Token Test (CRTT): Reliability and validity of the Computerized Revised Token Test (CRTT): Comparison of reading and listening versions in persons with and without aphasia.

Medin, D. L. (1986). Comment on“ Memory storage and retrieval processes in category learning.” *Journal of Experimental Psychology: General*, *115*, 373–381.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *8*(1), 37–50.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (Vol. 5). Hillsdale, NJ: Lawrence Erlbaum Assoc. Publishers.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238. doi:http://dx.doi.org/10.1037/0033-295X.85.3.207

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

Meulenbroek, R. G., Thomassen, A., Rosenbaum, D. A., Loukopoulos, L. D., & Vaughan, J. (1996). Adaptation of a reaching model to handwriting: How different effectors can produce the same written output, and other results. *Psychological Research*, *59*(1), 64–74.

Mewhort, D. J. K., & Johns, E. E. (2000). The extralist-feature effect: Evidence against item matching in short-term recognition memory. *Journal of Experimental Psychology: General*, *129*(2), 262.

Mewhort, D. J. K., & Johns, E. E. (2003). Sharpening the echo: An iterative-resonance model for short-term recognition memory. *Memory*, *13*(3-4), 300–307.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*(5), 524–545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, *30*(1), 69–89.

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

Miller, J., Sproesser, G., & Ulrich, R. (2008). Constant versus variable response signal delays in speed—accuracy trade-offs: Effects of advance preparation for processing time. *Perception & Psychophysics*, *70*(5), 878–886. doi:10.3758/PP.70.5.878

Miller, J., & Ulrich, R. (1998). Locus of the effect of the number of alternative responses: Evidence from the lateralized readiness potential. *Journal of Experimental Psychology: Human Perception & Performance*, *24*, 1215–1231.

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 775–799. doi:http://dx.doi.org/10.1037/0278-7393.27.3.775

Molfese, D. L., Key, A. P. F., Maguire, M. J., Dove, G. O., & Molfese, V. J. (2008). Event-related evoked potentials (ERPs) in speech perception. In D. Pisoni & R. Remez (Eds.), *The Handbook of Speech Perception* (pp. 99–121). John Wiley & Sons.

Montgomery, J. W. (1999). Recognition of Gated Words by Children With Specific Language ImpairmentAn Examination of Lexical Mapping. *Journal of Speech, Language, and Hearing Research*, *42*(3), 735–743.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533.

Morris, R. J., & Brown Jr., W. S. (1994). Age-related differences in speech variability among women. *Journal of Communication Disorders*, *27*(1), 49–64. doi:10.1016/0021-9924(94)90010-8

Morton, J., & Jassem, W. (1965). Acoustic Correlates of Stress. *Language and Speech*, *8*(3), 159–181. doi:10.1177/002383096500800303

Mulligan, N. W., Besken, M., & Peterson, D. (2010). Remember-Know and source memory instructions can qualitatively change old-new recognition accuracy: the modality-match effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 558.

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

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. doi:10.1016/j.clinph.2007.04.026

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends in Neurosciences*, *24*(5), 283–288. doi:10.1016/S0166-2236(00)01790-2

Nakatani, L. H., & Aston, C. H. (1978). *Acoustic and linguistic factors in stress perception*.

Neal, A., & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin & Review*, *4*(1), 24–37. doi:10.3758/BF03210770

Neill, W. T. (1997). Episodic retrieval in negative priming and repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(6), 1291–3105. doi:http://dx.doi.org/10.1037/0278-7393.23.6.1291

Neill, W. T., & Valdes, L. A. (1992). Persistence of negative priming: Steady state or decay? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 565–576. doi:http://dx.doi.org/10.1037/0278-7393.18.3.565

Neill, W. T., Valdes, L. A., Terry, K. M., & Gorfein, D. S. (1992). Persistence of negative priming: II. Evidence for episodic trace retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 993–1000. doi:http://dx.doi.org/10.1037/0278-7393.18.5.993

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 1–55). Hillsdale: Erlbaum.

Newell, K. M. (2003). Schema theory (1975): Retrospectives and prospectives. *Research Quarterly for Exercise and Sport*, *74*(4), 383–388.

Nickerson, R. S. (1965). Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *19*(2), 155–160. doi:http://dx.doi.org/10.1037/h0082899

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625+.

Nokes, T. . J. (2005). An investigation into adaptive shifting in knowledge transfer. In B. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 1660–1665). Mahaw, N.J.: Erlbaum.

Nokes, T. . J., & Ohlsson, S. (2001). How is abstract, generative knowledge acquired? A comparison of three learning scenarios. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 710–715). Psychology Press.

Nokes, T. J. (2004). Testing three theories of knowledge transfer. In D. Gentner, K. Forbus, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society Oriceedubg* (pp. 1029–1034). Mahaw, N.J.: Erlbaum.

Nokes, T. J. (2009). Mechanisms of knowledge transfer. *Thinking & Reasoning*, *15*(1), 1–36. doi:10.1080/13546780802490186

Nokes, T. J., & Ohlsson, S. (2003). Declarative transfer from a memory task to a problem solving task. *Cognitive Science Quarterly*, *3*, 259–296.

Norman, D. A. (1981). Categorization of Action Slips. *Psychological Review*, *88*(1), 1–15.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwartz, & D. C. Shapiro (Eds.), *Consciousness and regulation.* New York: Plenum Press.

Nosofsky, R. M. (1986). Attention, similarity, and the identification categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 700–708.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(1), 3.

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, *43*(1), 25–53.

Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 282.

Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 211–233.

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, *118*(2), 280–315. doi:10.1037/a0022494

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Nosofsky, R. M., & Stanton, R. D. (2006). Speeded old-new recognition of multidimensional perceptual stimuli: Modeling performance at the individual-participant and individual-item levels. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 314–334. doi:10.1037/0096-1523.32.2.314

Nosofsky, R. M., & Zaki, S. R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1194.

Oh, J. S., Au, T. K.-F., & Jun, S. (2010). Early childhood language memory in the speech perception of international adoptees*. *Journal of Child Language*, *37*(5), 1123.

Ohlsson, S. (1993). Abstract schemas. *Educational Psychologist*, *28*(1), 51–66.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, *103*(2), 241.

Olsson, H., Wennerholm, P., & Lyxzèn, U. (2004). Exemplars, prototypes, and the flexibility of classification models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 936.

Ono, A. (1990). Effects of different attentional strategies and practice on motor efficiency. *Perceptual and Motor Skills*, *71*, 35–43.

Osman, A., Kornblum, S., & Meyer, D. E. (1990). Does motor programming necessitate response execution? *Journal of Experimental Psychology*, *16*(1), 183–198.

Owens, E., & Schubert, E. D. (1977). Development of the California consonant test. *Journal of Speech, Language and Hearing Research*, *20*(3), 463.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 324–354.

Papagno, C., & Vallar, G. (1992). Phonological Short-term Memory and the Learning of Novel Words: The Effect of Phonological Similarity and Item Length. *The Quarterly Journal of Experimental Psychology Section A*, *44*(1), 47–67. doi:10.1080/14640749208401283

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1995). Goal-based speech motor control: A theoretical framework and some preliminary data. *Journal of Phonetics*, *23*(1–2), 23–35. doi:10.1016/S0095-4470(95)80030-1

Perruchet, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology*, *20*(1), 223.

Persson, M., & Rieskamp, J. (2009). Inferences from memory: Strategy-and exemplar-based judgment models compared. *Acta Psychologica*, *130*(1), 25–37.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee & P. J. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure*. John Benjamins Publishing.

Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory Phonology*, *7*, 101–139.

Pihko, E., Leppanen, P. H. T., Eklund, K. M., Cheour, M., Guttorm, T. K., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: I. Age effects. [Miscellaneous Article]. *Neuroreport April 6, 1999*, *10*(5), 901–905.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3p1), 353.

Posner, M. I., & Keele, S. W. (1970). Retention in abstract ideas. *Journal of Experimental Psychology*, *83*, 304–308.

Proctor, R. W., & Dutta, A. (1995). *Skill Acquisition and Human Performance.* Thousand Oaks: Sage Publications.

Radvansky, G. (2006). *Human Memory*. Boston, MA: Pearson Education Group, Inc.

Ramig, L. A., & Ringel, R. L. (1983). Effects of Physiological Aging on Selected Acoustic Characteristics of Voice. *Journal of Speech and Hearing Research*, *26*(1), 22.

Rawson, K. A. (2004). Exploring automaticity in text processing: Syntactic ambiguity as a test case. *Cognitive Psychology*, *49*(4), 333–369.

Rawson, K. A., & Middleton, E. L. (2009). Memory-Based Processing as a Mechanism of Automaticity in Text Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 353–370.

Real-Time Pitch, Model 5121. (KayPentax). Lincoln Park, NJ: KayPentax.

Reber, A. S. (1967). Implicit Learning of Artificial Grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Reber, A. S. (1989). Implicit Learning and Tacit Knowledge. *Journal of Experimental Psychology: General [PsycARTICLES]*, *118*(3), 219. doi:http://dx.doi.org/10.1037/0096-3445.118.3.219

Reber, A. S., & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition*, *6*(3), 189–221. doi:10.1016/0010-0277(78)90013-6

Reisberg, D. (2001). *Cognition: Exploring the science of the world.* (2nd ed.). New York: W.W.Norton & Company.

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, *92*(1), 81–110. doi:10.1037/0033-2909.92.1.81

Rhodes, B. J., Bullock, D., Verwey, W. B., Averbeck, B. B., & Page, M. P. A. (2004). Learning and production of movement sequences: Behavioral, neurophysiological, and modeling perspectives. *Human Movement Science*, *23*(5), 699–746. doi:10.1016/j.humov.2004.10.008

Ries, P. W. (1994). Prevalence and characteristics of persons with hearing trouble: United States, 1990-91. *Vital and Health Statistics. Series 10, Data from the National Health Survey*, (188), 1–75.

Roberts, A. H. (1965). *A statistical linguistic analysis of American English*. London, England: Mouton and Co.

Rodrigues, P. M., & Murre, J. M. (2007). Rules-plus-exception tasks: A problem for exemplar models? *Psychonomic Bulletin & Review*, *14*(4), 640–646.

Roelofs, A., & Meyer, A. S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(4), 922.

Rogers, M. A., & Storkel, H. L. (1998). Reprogramming phonologically similar utterances: The role of... *Journal of Speech, Language & Hearing Research*, *41*(2), 258.

Rogers, M. A., & Storkel, H. L. (1999). Planning speech one syllable at a time: The reduced buffer capacity hypothesis in apraxia of speech. *Aphasiology*, *13*(9-11), 793–805.

Rosalie, S. M., & Muller, S. (2012). A model for the transfer of perceptual-motor skill learning in human behaviors. *Research Quarterly for Exercise and Sport*, *83*(3), 413+.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 491.

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, *109*(4), 444–474. doi:http://dx.doi.org/10.1037/0096-3445.109.4.444

Rosenbaum, D. A. (1990). On choosing between movement sequences: Comments on Rose (1988). *Journal of Experimental Psychology. Human Perception and Performance*, *16*(2), 439.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology*, *52*, 453–470.

Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Psychology: General*, *113*(3), 372.

Rosenbaum, D. A., Loukopoulos, L. D., Meulenbroek, R. G. J., Vaughan, J., & Engelbrecht, S. E. (1995). Planning reaches by evaluating stored postures. *Psychological Review*, *102*(1), 28–67. doi:http://dx.doi.org/10.1037/0033-295X.102.1.28

Rosenbaum, D. A., Meulenbroek, R. J., Vaughan, J., & Jansen, C. (2001). Posture-based motion planning: Applications to grasping. *Psychological Review*, *108*(4), 709–734. doi:http://dx.doi.org/10.1037/0033-295X.108.4.709

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*(3), 371–416.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 629–639. doi:http://dx.doi.org/10.1037/0278-7393.13.4.629

Ross, B. H., Taylor, E. G., Middleton, E. L., & Nokes, T. J. (2008). Concept and category learning in humans. In J. Byrne & H. L. Roediger 3rd (Eds.), *Learning and Memory: A comprehensive reference* (Vol. Cognitive psychology of memory, pp. 535–557). Oxford: Elsevier Science, Ltd.

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2-to 4-year-olds: evidence from typically developing children. *Journal of Speech, Language and Hearing Research*, *47*(1), 223.

Rumelhart, D. E. (1989). The architecture of mind: A connectionist approach. In J. Haugeland (Ed.), *Mind Design II*. MIT Press.

Rumelhart, D. E., & Ortony, A. (1977). Representation of knowledge. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ryan, W. J., & Burk, K. W. (1974). Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorders*, *7*(2), 181–192. doi:10.1016/0021-9924(74)90030-6

Sakamoto, Y., Matsuka, T., & Love, B. C. (2004). Dimension-wide vs. exemplar-specific attention in category learning and recognition. In M. Lovett, C. Schunn, C. Lebiere, & P. Munro (Eds.), *Proceedings of the 6th International Conference on Cognitive Modeling* (pp. 261–266). Mahwah, NJ: Erlbaum.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382.

Samuel, A. G. (1982). Phonetic prototypes. *Perception & Psychophysics*, *31*(4), 307–314.

Savela, J., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). The mismatch negativity and reaction time as indices of the perceptual distance between the corresponding vowels of two related languages. *Cognitive Brain Research*, *16*(2), 250–256.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 501–518.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hove, UK: Lawrence Erlbaum Associates.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*(4), 225.

Schmidt, R. A. (1983). On the Underlying Structure of Well-Learned Motor Responses A Discussion of Namikas and Schneider and Fisk. *Advances in Psychology*, *12*, 145–165.

Schmidt, R. A. (1988). *Motor control and learning: A behavioral emphsis* (2nd ed.). Champagne, IL: Human Kinetics.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest a new concepts for training. *Psychological Science*, *3*(207), 207–217.

Schmidt, R. A., & Lee, T. (2005). *Motor Control and Learning: A Behavioral Emphasis* (4th ed.). Champagne, IL: Human Kinematics.

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

Schmidt, R. A., Young, D. E., Swinnen, W., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 352–59.

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

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Schweitzer, A., & Möbius, B. (2004). Exemplar-based production of prosody: Evidence from segment and syllable durations. In *Speech Prosody 2004, International Conference*. Retrieved from http://www.isca-speech.org/archive_open/sp2004/sp04_459.html

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

Seger, C. A. (1994). Implicit Learning. *Psychological Bulletin*, *115*(2), 163–196.

Service, E., & Craik, F. I. M. (1993). Differences Between Young and Older Adults in Learning A Foreign Vocabulary. *Journal of Memory and Language*, *32*(5), 608–623. doi:10.1006/jmla.1993.1031

Sevald, C. A., Dell, G. S., & Cole, J. S. (1995). Syllable Structure in Speech Production: Are Syllables Chunks or Schemas? *Journal of Memory and Language*, *34*(6), 807–820. doi:10.1006/jmla.1995.1035

Shafer, V. L., Schwartz, R. G., & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research*, *18*(3), 242–254. doi:10.1016/j.cogbrainres.2003.10.007

Shaiman, S., McNeil, M. R., Szuminsky, N. J., Meigh, K. M., & Kotler, J. B. (2006). Motor learning of volitional nonspeech oral movements: Intraoral pressure and articulatory kinematics. Presented at the Conference on Motor Speech.

Shanks, D. R. (1991). Some parallels between associative learning and object classification. In J. A. Meyer & S. W. Wilson (Eds.), *From animals to animats: proceeding of the first international conference on simulations of adaptive behavior* (pp. 337–343). Cambridge, MA: MIT Press.

Shanks, D. R. (1995). *The psychology of associative learning*. Cambridge, MA: Cambridge University Press.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning-systems. *Behavioral Brain Science*, *17*(3), 367–395.

Shapiro, D. C., & Schmidt, R. A. (1982). The schema theory: Recent evidence and developmental implications. In J. A. S. Kelso & J. E. Clark (Eds.), *The development of movement control and co-ordination.* (pp. 113–150). New York: John Wiley.

Shea, C. H., & Wulf, G. (2005). Schema theory: A critical appraisal and reevaluation. *Journal of Motor Behavior*, *37*(2), 85–101.

Shea, J. B., & Wright, D. L. (1991). When Forgetting Benefits Motor Retention. *Research Quarterly for Exercise and Sport*, *62*(3), 293–301. doi:10.1080/02701367.1991.10608726

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156–163.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42.

Shiffrin, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of human memory*. New York, NY: Academic Press.

Shiffrin, R. M. (1976). Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Memory processes* (Vol. 4). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Dumais, S. (1981). The development of automatism. In J. A. Anderson (Ed.), *Cognitive Skills and Their Acquistion* (pp. 111–140). Hillsdale: Lawrence Erlbaum Associates.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127.

Shriberg, L. D., & Kent, R. D. (1982). Distributional, structural, and proportional occurrence for American English sounds, syllables, and words. In *Clinical Phonetics* (p. 429). New York, NY: John Wiley & Sons.

Sidaway, B., Sekiya, H., & Fairweather, M. (1995). Movement variability as a function of accuracy demand in programmed serial aiming responses. *Journal of Motor Behavior*, *27*(1), 67–76.

Singh, L., Liederman, J., Mierzejewski, R., & Barnes, J. (2011). Rapid reacquisition of native phoneme contrasts after disuse: You do not always lose what you do not use. *Developmental Science*, *14*(5), 949–959.

Singley, M. K., & Anderson, J. R. (1989). *The Transfer of Cognitive Skill.* Cambridge: Harvard University Press.

Skemp, R. R. (1987). *The Psychology of Learning Mathematics: Expanded American Edition*. Psychology Press.

Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, *100*(4), 2471–2485.

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research*, *104*(3), 493–501.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology*, *45*(1), 22–33.

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science*, *16*(1), 1–40.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*(1), 3–21. doi:http://dx.doi.org/10.1037/0033-295X.99.1.3

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1411–1436. doi:http://dx.doi.org/10.1037/0278-7393.24.6.1411

Smith, R. E., & Hunt, R. R. (2000). The Effects of Distinctiveness Require Reinstatement of Organization: The Importance of Intentional Memory Instructions. *Journal of Memory and Language*, *43*(3), 431–446. doi:10.1006/jmla.2000.2707

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34.

SPSS. (2012). *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Corp.

Staiger, A., & Ziegler, W. (2008). Syllable frequency and syllable structure in the spontaneous speech production of patients with apraxia of speech. *Aphasiology*, *22*(11), 1201–1215. doi:10.1080/02687030701820584

Standing, L. (1973). Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. doi:10.1080/14640747308400340

Sternberg, S., Knoll, R. L., Monsell, S., & Wright, C. E. (1988). Motor programs and hierarchical organization in the control of rapid speech. *Phonetica*, *45*(2-4), 175–197.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. *Information Processing in Motor Control and Learning*, 117–152.

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating Phonotactic Probability and Neighborhood Density in Adult Word Learning. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1175–1192. doi:10.1044/1092-4388(2006/085)

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking Lexical Access in Speech Production: Electrophysiological Correlates of Word Frequency and Cognate Effects. *Cerebral Cortex*, *20*(4), 912–928. doi:10.1093/cercor/bhp153

Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., & Näätänen, R. (2002). Top-down effects can modify the initially stimulus-driven auditory organization. *Cognitive Brain Research*, *13*(3), 393–405.

Sweegers, C. C. G., Takashima, A., Fernández, G., & Talamini, L. M. (2013). Neural mechanisms supporting the extraction of general knowledge across episodic memories. *NeuroImage*. doi:10.1016/j.neuroimage.2013.10.063

Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika*, *48*(3), 393–423. doi:10.1007/BF02293683

Taylor, S. E., & Fiske, S. T. (1978). Salience, Attention, and Attribution: Top of the Head Phenomena. In Leonard Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. Volume 11, pp. 249–288). Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S006526010860009X

Terry, W. S. (2000). *Learning and Memory: Basic Principles, Processes, and Procedures.* Boston: Allyn and Bacon.

Thorndike, E. L. (1903). *Educational psychology*. New York: Lemke and Buechner.

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

Tomporowski, P. D. (2003). *The psychology of skill: A life-span approaches*. Westport: Praeger.

Torre III, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, *42*(5), 324–333. doi:10.1016/j.jcomdis.2009.03.001

Touron, D. R., Hooyer, W. J., & Cerella, J. (2001). Cognitive Skill Acquisition and Transfer in Younger and Older Adults. *Psychology and Aging*, *16*(4), 555–563.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Tyler, L. K. (1984). The structure of the initial cohort: Evidence from gating. *Perception & Psychophysics*, *36*(5), 417–427. doi:10.3758/BF03207496

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

Van Rossum, J. H. (1987). *Motor Development and Practice: The Variability of Practice Hypothesis in Perspective*. Free University Press.

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,< i> Brain and Language</i>, 88, 148–159, 2004. *Brain and Language*, *96*(2), 235–239.

Verwey, W. B. (1995). A Forthcoming Key Press Can Be Selected While Earlier Ones Are Executed. *Journal of Motor Behavior*, *27*(3), 275–284. doi:10.1080/00222895.1995.9941717

Verwey, W. B. (1999). Evidence for a Multistage Model of Practice in a Sequential Movement Task. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1693–1708.

Verwey, W. B. (2001). Concatenating familiar movement sequences: the versatile cognitive processor. *Acta Psychologica*, *106*(1–2), 69–95. doi:10.1016/S0001-6918(00)00027-5

Vidal, F., Burle, B., Grapperon, J., & Hasbroucq, T. (2011). An ERP study of cognitive architecture and the insertion of mental processes: Donders revisited. *Psychophysiology*, *48*(9), 1242–1251.

Vidal, F., Grapperon, J., Bonnet, M., & Hasbroucq, T. (2003). The nature of unilateral motor commands in between-hand choice tasks as revealed by surface Laplacian estimation. *Psychophysiology*, *40*(5), 796–805.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325–329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of Memory and Language*, *40*(3), 374–408. doi:10.1006/jmla.1998.2618

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1), 306–311.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 328–344.

Wadman, W. J., Denier van der Gon, J. J., Geuze, R. H., & Mol, C. R. (1979). Control of fast goal-directed arm movements. *J Hum Mov Stud*, *5*(1), 3–17.

Wallace, W. P. (1980). On the use of distractors for testing recognition memory. *Psychological Bulletin*, *88*(3), 696–704. doi:10.1037/0033-2909.88.3.696

Walsh, B., & Smith, A. (2002). Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language and Hearing Research*, *45*(6), 1119.

Walsh, M., Möbius, B., Wade, T., & Schütze, H. (2010). Multilevel exemplar theory. *Cognitive Science*, *34*(4), 537–582.

Ward, G., & Churchill, E. F. (1998). Two tests of instance-based and abstract rule-based accounts of invariant learning. *Acta Psychologica*, *99*(3), 235–253. doi:10.1016/S0001-6918(98)00014-6

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence. *Cognitive Brain Research*, *18*(2), 149–161. doi:10.1016/j.cogbrainres.2003.10.001

Welford, A. T. (1977). Motor performance. *Handbook of the Psychology of Aging*, *2*(450), 95.

Wells, J. E. (1974). Strength theory and judgments of recency and frequency. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 378–392.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 349–355.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*(5), 672–683. doi:10.1037/0012-1649.24.5.672

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*(1), 35–44.

Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63.

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America*, *75*(6), 1866–1878. doi:10.1121/1.390988

Whitacre, C. A., & Shea, C. H. (2000). Performance and Learning of Generalized Motor Programs: Relative (GMP) and Absolute (Parameter) Errors. *Journal of Motor Behavior*, *32*(2), 163–175. doi:10.1080/00222890009601368

Whittlesea, B. W. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 3–17. doi:http://dx.doi.org/10.1037/0278-7393.13.1.3

Whittlesea, B. W., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General*, *122*(2), 227.

Wilberg, R. B., & Guay, M. (1985). Long-term memory for a single movement. In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control*. North Holland, Amsterdam: Elsevier Science Publishers.

Witt, A., & Vinter, A. (2012). Artificial grammar learning in children: abstraction of rules or sensitivity to perceptual features? *Psychological Research*, *76*(1), 97–110. doi:10.1007/s00426-011-0328-5

Wright, D. L., & Shea, C. H. (1991). Contextual dependencies in motor skills. *Memory & Cognition*, *19*(4), 361–370. doi:10.3758/BF03197140

Wright, D. L., & Shea, C. H. (2001). Manipulating generalized motor program difficulty during blocked and random practice does not affect parameter learning. *Research Quarterly for Exercise & Sport*, *72*(1), 32–8.

Wright, D. L., Shea, C. H., Li, Y., & Whitacre, C. (1996). Contextual Dependencies During Perceptual-motor Skill Acquisition: Gone But Not Forgotten! *Memory*, *4*(1), 91–108. doi:10.1080/096582196389112

Wright, R. L., & Burton, A. M. (1995). Implicit Learning of an Invariant: Just Say No. *The Quarterly Journal of Experimental Psychology Section A*, *48*(3), 783–796. doi:10.1080/14640749508401415

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

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*(4), 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*(5), 1134.

Xue, S. A. (2003). Changes in the Human Vocal Tract Due to Aging and the Acoustic Correlates of Speech Production: A Pilot Study. *Journal of Speech, Language, and Hearing Research*, *46*(3), 689–701. doi:10.1044/1092-4388(2003/054)

Yap, K. S., Lim, C. P., & Au, M. T. (2011). Improved GART Neural Network Model for Pattern Classification and Rule Extraction With Application to Power Systems. *IEEE Transactions on Neural Networks*, *22*(12), 2310–2323. doi:10.1109/TNN.2011.2173502

Yates, M., Locker, L., & Simpson, G. B. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review*, *11*(3), 452–457.

Youmans, G., Youmans, S. R., & Hancock, A. A. (2011). Script training treatment for adults with apraxia of speech. *American Journal of Speech-Language Pathology*, *20*(1), 23–37.

Zelaznik, H. N. (1977). Transfer in rapid timing tasks: An examination of the role of variability in practice. In R. W. Christina & D. M. Landers (Eds.), *Psychology of motor behavior and sports*. Champagne, IL: Human Kinetics.